

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

QuoVadis Research @ FEI



ročník 5, číslo 1
Február 2022

ISSN 2585-9587

QuoVadis Research @ FEI je vedecký časopis, ktorého obsahom sú prezentácie výsledkov vedeckého bádania v oblasti informatických a elektrotechnických vied na Fakulte elektrotechniky a informatiky Technickej univerzity v Košiciach (FEI TUKE). Časopis obsahuje plnotextové vedecké články, ktorých obsahom sú výsledky dizertačných prác doktorandov a ich školiteľov obhájených v aktuálnom období na Fakulte elektrotechniky a informatiky Technickej univerzity v Košiciach. Vedecké články prispievajú k rozšíreniu vedeckých poznatkov v odboroch informatiky a elektrotechniky. Ide o podporu a rozvoj takých odborov ako je výpočtová technika a informatika, počítačové siete, kyberbezpečnosť, mechatronika a automatizácia, kybernetika a umelá inteligencia, multimediálna telekomunikačná technika, aplikovaná informatika, počítačové modelovanie, vybrané oblasti matematiky, hospodárska informatika, elektronika, mikroelektronika a optoelektronika, spracovanie signálov, teoretická elektrotechnika, elektroenergetika a technika vysokých napätí, silnoprúdová elektrotechnika, meracie systémy v elektronike, riadenie elektrotechnickej výroby, materiály a technológie elektroniky, štruktúra a vlastnosti moderných materiálov. Prijaté rukopisy uvádzajú relevantnú medzinárodnú literatúru a popisujú nové objavy, aplikácie nových techník a metód, použitie existujúcej metodológie na nový problém, čím prispievajú k rozširovaniu základných vedeckých poznatkov v danej vednej disciplíne.

Redakčná rada

Šéfredaktor časopisu:

prof. Ing. Liberios Vokorokos, PhD.

Výkonný redaktor:

doc. Ing. Anton Baláž, PhD.

Členovia:

prof. Ing. Jaroslav Porubán, PhD.

odbor: informatika

prof. Ing. Liberios Vokorokos, PhD.

odbor: informatika, kyberbezpečnosť

prof. Ing. Peter Sinčák, CSc.

odbor: kybernetika, umelá inteligencia, inteligentné systémy, aplikovaná informatika

prof. Ing. Ján Paralič, PhD.

odbor: hospodárska informatika

prof. RNDr. Ján Plavka, CSc.

odbor: aplikovaná matematika, počítačové modelovanie

prof. Ing. Jozef Juhár, CSc.

odbor: multimediálne telekomunikácie, počítačové siete

prof. Ing. Michal Kolcun, PhD.

odbor: elektroenergetika a technika vysokých napätí

prof. Ing. Dobroslav Kováč, CSc.

odbor: teoretická elektrotechnika, automatizácia

prof. Ing. Daniela Perduková, PhD.

odbor: elektrotechnika, mechatronika, riadenie elektrotechnickej výroby

prof. Ing. Alena Pietriková, CSc.

odbor elektrotechnológie a materiály

prof. Ing. Ján Šaliga, CSc.

odbor: elektronika, spracovanie signálov, optoelektronika, meranie

prof. RNDr. Ján Ziman, CSc.

odbor: štruktúra a vlastnosti moderných materiálov

Adresa vydavateľa a redakcie: Redakcia časopisu QuoVadis Research @ FEI, Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky, Letná 9, 042 00 Košice, Slovenská republika, email:quovadis@fei.tuke.sk

Obsah

Tepelný model činnosti CNC frézy	
<i>Branislav Fecko, Tibor Vince</i>	4-13
Optimalizácia výkonových strát v smart sieťach	
<i>Daniel Pál, Ľubomír Beňa</i>	13-22
Review Of Feature Extraction Techniques From Still-3D Images For Facial Emotion Recognition	
<i>Fouzia Adjailia, Peter Sinčák</i>	22-29
Reinforcement Learning-Based Study Scheduling for Optimal Learning	
<i>Ján Magyar, Peter Sinčák</i>	29-41
Návrh metodiky pre modelovanie, analýzu a simuláciu efektívne kráčajúcich robotických systémov	
<i>Lukáš Koska, Anna Jadlovská, Slávka Jadlovská</i>	41-61
Smart grid research	
<i>Maksym Oliinyk, Jaroslav Džmura</i>	61-68
Virtuálno-realistné technológie a inteligentné používateľské rozhrania	
<i>Martin Sivý, Branislav Sobota</i>	68-76
Metódy strojového učenia na rozpoznávanie objektov v počítačovom videní	
<i>Martin Štancel, Liberios Vokorokos</i>	76-89
Lung Ultrasound Analysis with Deep Learning	
<i>Miroslav Jascur, Marek Bundzel</i>	89-99
Výskum komunikačných architektúr v prostredí internetu vecí	
<i>Rastislav Petija, František Jakab</i>	99-105
Metódy hlbokého učenia v analytických úlohách a rozhodovaní	
<i>Viera Maslej Krešňáková, Peter Butka</i>	105-112

Tepelný model činnosti CNC frézy

Ing. Branislav Fecko

Katedra teoretickej a priemyselnej elektrotechniky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovensko
branislav.fecko@tuke.sk

doc. Ing. Tibor Vince, PhD.

Katedra teoretickej a priemyselnej elektrotechniky
Fakulta elektrotechniky a informatiky, TUKE
Košice, Slovensko
tiber.vince@tuke.sk

Abstrakt — Cieľom publikácie je dokumentácia realizácie tepelného modelu CNC frézy. V prvej časti sa nachádza analýza strát pohonov a tepelných zdrojov CNC frézy. Pre monitorovanie CNC zariadenia sa vyvinul monitorovací systém. Jeho následným využitím sa získali potrebné údaje na vytvorenie tepelného modelu CNC frézy a jeho verifikáciu. Následná časť práce obsahuje simulácie použitých pohonov vyšetrovanej CNC frézy. Vytvoril sa model jednosmerného motora s permanentnými magnetmi a model krokového motora. Modely jednotlivých motorov obsahujú simuláciu elektrotechnických a tepelných vlastností zariadenia. Parciálne simulácie sa využili na vytvorenie kompletného tepelného modelu. Využitím tepelnej simulácie v kombinácii s monitorovacím systémom sa vytvoril systém na predikciu ohrievania jednotlivých častí, s praktickým využitím v prediktívnej údržbe.

Kľúčové slová — CNC fréзка; jednosmerný motor; krokový motor; straty; tepelný model

Abstract — The aim of the publication is to document the implementation of the thermal model of a CNC milling machine. The first part is an analysis of drive losses and heat sources of the CNC milling machine. A monitoring system has been developed for monitoring CNC equipment. Its subsequent use provided the necessary data to create a thermal model of a CNC milling machine and its verification. The following part of the work contains simulations of used drives of the investigated CNC milling machine. A model of a DC motor with permanent magnets and a model of a stepper motor were created. Models of individual motors contain a simulation of electrical and thermal properties of the device. Partial simulations were used to create a complete thermal model. Using thermal simulation in combination with a monitoring system, a system was created to predict the heating of individual parts, with practical use in predictive maintenance.

Keywords— CNC milling machine; DC motor; stepper motor; losses; thermal model

I. ÚVOD

Hlavným dôvodom prečo sa zameriavame na modelovanie CNC systému, je jeho komplexnosť a kombinácia elektrotechnického a strojárského inžinierstva. Modelovanie tak komplexných systémov ako CNC fréзка je malo preskúmaná oblasť z dôvodu náročnosti riešenia. Až prvky Industry 4.0 nám prinášajú možnosti analyzovať tak dynamické systémy.

Každé CNC zariadenie ma svoju špecifickú úlohu, na ktorú je navrhnuté. Je potreba aby CNC stroj spracoval s určitou presnosťou a spoľahlivosťou. Práve preto medzi často vyskytujúce problémy týkajúce sa CNC strojov patrí ich presnosť, životnosť a chladenie, ktoré sú skúmané a ošetrované rôznymi spôsobmi. Všetky spomenuté problémy vieme čiastočne prepojiť s teplom vznikajúcim v CNC stroji počas jeho prevádzky. Presnosť zariadenia je ovplyvnená tepelnou rozťažnosťou konštrukčných materiálov stroja a jeho iných častí. Životnosť CNC frézy priamo súvisí so životnosťou jej pohonov, v prípade elektrických pohonov je známe, že zvýšenie teploty motorov nad prevádzkovú hodnotu znižuje životnosť motora. Samotné chladenie pri CNC frézkach je priamo spojené s teplom, vtedy hovoríme o chladení obrábacieho nástroja alebo o chladení pohonov stroja, ktoré pri svojej prevádzke vytvárajú teplo. V prípade vhodného využitia možnosti predikcie vytvoreného tepla CNC frézy je možné minimalizovať vplyv teploty na spomínané problémy alebo predikovať ich vzniknutie. Z týchto dôvodov sa táto práca zaoberá skúmaním tepelného modelu CNC frézy.

Model CNC stroja, monitorovací systém a implementácia tepelného modelu bude aplikovaná na CNC zariadenie, ktoré bolo vyvíjané v priebehu riešenia bakalárskej aj diplomovej práce [1].

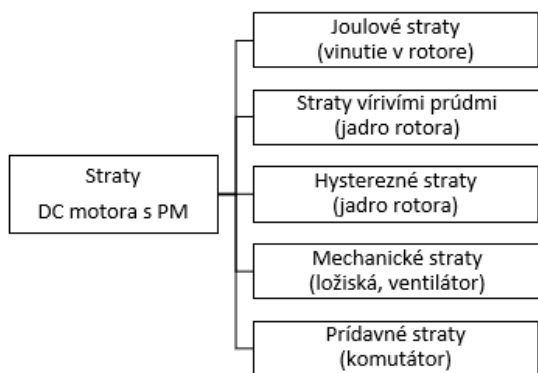
II. TEPELNÉ ZDROJE CNC FRÉZKY

Elektrické motory slúžia na premenu elektrickej energie na mechanickú. Počas tejto premeny nie je premenená celá časť elektrickej energie na mechanickú, z dôvodu vzniku strát. Straty motora sa premieňajú na tepelnú energiu, preto pri výskume tepelnej činnosti CNC frézy je potrebné sa zaoberať vzniknutými stratami v pohonoch frézy [2]. Jednotlivé straty budú opísané v nasledujúcich podkapitolách. Ich správne definovanie je potrebné pre vytvorenie tepelného modelu. Tepelné zdroje CNC frézy je možné rozdeliť na teplo generované elektrickými pohonmi, trením konštrukčných častí stroja a teplo generované v procese obrábania.

A. Rozdelenie Strát v Jednosmernom Motore s Permanentnými Magnetmi

Jednosmerný motor s permanentnými magnetmi (DC motor s PM) sa využíva ako vreteno nášho CNC stroja. Nasledujúci Obr. 1 zobrazuje rozdelenie jednotlivých strát a miesto ich pôsobenia. Vznik strát závisí od konštrukčného

návrhu pohonu. Cieľom každého návrhu pohonu je čo najviac eliminovať vznik strát.

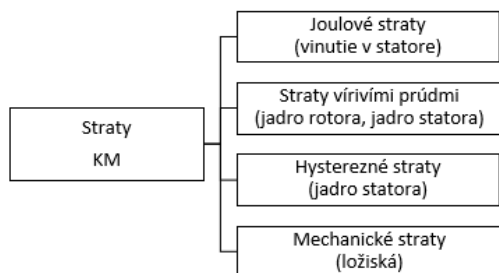


Obr. 1 Rozdelenie a umiestnenie vzniknutých strát v DC motore s PM

B. Rozdelenie strát v 2-fázovom hybridnom krokovom motore

Rozdelenie tepelných strát krokového motora znázorňuje

Obr. 2. Na základe iného konštrukčného usporiadania prvkov a typu činnosti motora môžeme vidieť rozdiel oproti rozdeleniu zobrazujúceho tepelné zdroje DC motora s PM.



Obr. 2 Rozdelenie a umiestnenie vzniknutých strát v krokovom motore

Na základe vyššie zobrazených rozdelení strát sa vytvorili modely pohonov v Simulinku, ktoré sa využili pri vytváraní tepelného modelu CNC fréžky.

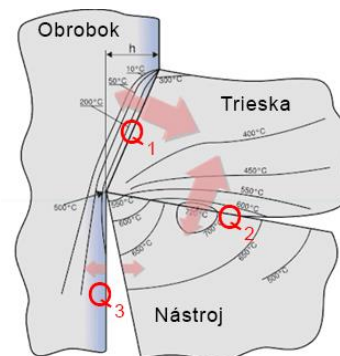
C. Proces obrábania CNC fréžkou

Spôsob obrábania materiálu CNC fréžkou je fréžovanie. Počas procesu obrábania materiálu fréžovaním vzniká teplo. Obrábaním sa od 95% do 98% mechanickej práce premení na teplo. Mechanická práca, ktorá sa pri fréžovaní premení na teplo pozostáva z plastickej deformácie a trenia nástroja, obrobku a triesky. Vzniknuté teplo urýchľuje opotrebenie nástroja a ovplyvňuje mechanické vlastnosti obrobku. V oblasti rezu vznikajú tri tepelné zdroje, medzi ktoré patrí:

- primárna plastickej deformácia (Q_1)

- kontakt čela nástroja s trieskou a sekundárna plastickej deformácia (Q_2)
- trenie medzi plochami obrobku a nástroja (Q_3)

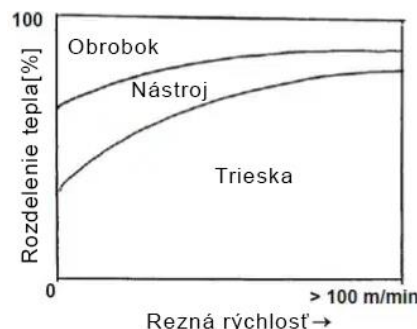
Teplo sa z vymenovaných zdrojov následne šíri do obrobku, nástroja, triesky a prostredia [3]. Trieska z oblasti rezu odoberá najväčšie množstvo tepla Q_1 a Q_2 . Nižšiu hodnotu tepla Q_2 a Q_3 absorbuje nástroj. Obrobok prijíma najnižšiu časť tepla absorbuje Q_1 a Q_3 . Obr. 3 obsahuje grafické znázornenie smeru šírenia tepla v procese obrábania [4].



Obr. 3 Tepelné zdroje v procese obrábania (obrábanie ocele)

Obrábanie fréžovaním vytvára v zóne rezania deformáciu materiálu a trenie, čo je príčinou zvýšenia zaťaženia CNC stroja. Sila, ktorú je potrebné vykonať na oddelenie triesky od obrobku priamo ovplyvňuje zaťaženie vretena a pohonov lineárnych vedení, ktoré je zodpovedné za ďalší nárast teploty CNC komponentov. Medzi parametre, ktoré najviac ovplyvňujú zaťaženie CNC fréžky patrí rýchlosť posuvu, otáčky vretena za minútu, geometria nástroja, materiál obrobku, stratégia fréžovania a objem odoberaného materiálu.

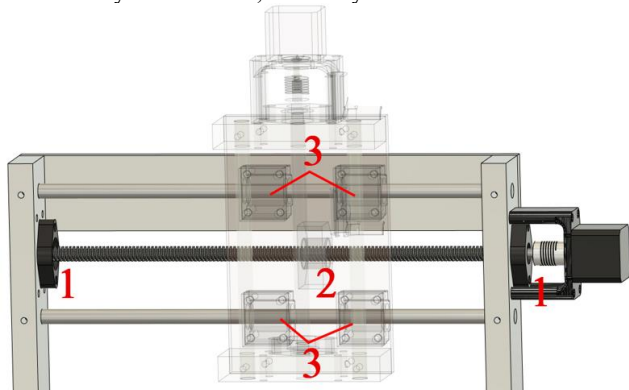
Vymenované vlastnosti neovplyvňujú iba množstvo vytvoreného tepla ale aj jeho rozdelenie. Napríklad pri obrábaní materiálov so zlou tepelnou vodivosťou sa viac tepla presunie do nástroja, v prípade zvyšovania rýchlosti posuvu sa zväčšuje objem tepla odoberaný trieskou. Ďalší parameter, ktorý mení rozdelenie vytvoreného tepla je rýchlosť fréžovania. Vplyv reznej rýchlosti na rozdelenie tepla medzi obrobok, nástroj a triesku zobrazuje Obr. 4 [5][6].



Obr. 4 Rozdelenie tepla v procese fréžovania

D. Trenie konštrukčných častí stroja

Ďalšie tepelné zdroje, ktoré vznikajú na CNC frézke môžeme zhrnúť do skupiny tepelných zdrojov vznikajúcich trením konštrukčných častí stroja. CNC frézka sa skladá s troch lineárnych vedení. Obr. 5 zobrazuje lineárne vedenie osi X, všetky vedenia na skúmanej CNC frézke sú konštruované analogicky, preto sme si zvolili os X na znázornenie potenciálnych zdrojov tepla CNC frézky. Číslo 1 označuje umiestnenie guľôčkových ložísk, ktoré slúžia na upevnenie závitovej tyče. V našom prípade využívame trapézovú závitovú tyč v kombinácii s maticou. Matica je označená číslom 2 je zdrojom klzného trenia. Poslednou konštrukčnou časťou sú lineárne vedenia, ktoré sú tvorené 12 mm nerezovou tyčou a lineárnymi ložiskami, označenými číslom 3.



Obr. 5 os X CNC frézky

Vo zvýraznených konštrukčných prvkoch vzniká valivé a klzne trenie, ktoré sa premení na teplo. Počas bezporuchovej prevádzky stroja je objem vygenerovaného tepla vzhľadom k ostatým tepelným stratám CNC stroja zanedbateľný. Preto pri vytváraní základného tepelného modelu CNC frézky opisované straty zanedbávame. Opisované konštrukčné prvky sú navrhnuté tak, aby v bezporuchovej prevádzke zabezpečili čo najmenšie trenie. Väčšina porúch daných komponentov sa prejaví zvýšením ich trenia, čím sa zvýši aj tvorba tepla. Presne s týmto stavom budeme uvažovať pri praktickom využití nášho tepelného modelu CNC frézky. Za bezporuchového chodu CNC frézky sa teplota predikcie a teplota reálneho systému bude s určitou presnosťou zhodovať, vznik väčších odchýlok ako je presnosť modelu v porovnaní systémov, bude slúžiť na signalizovanie vzniku porúch v CNC frézke

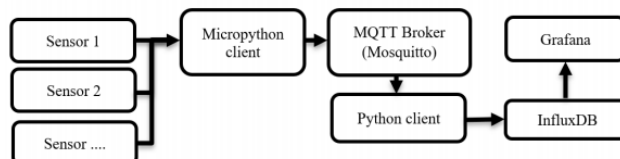
III. MONITOROVACÍ SYSTÉM CNC FRÉZKY

Hlavný dôvod vzniku monitorovacieho systému bola potreba zaznamenávania teplôt pohonov CNC frézky a teploty jej okolia. Manuálne odčítavanie údajov z meracích prístrojov prichádza do úvahy iba v prípade neopakujúcich sa meraní. Analýza nášho systému si vyžaduje hodiny nepretržitého merania a zaznamenávania množstva premenných, preto úsilie bolo zamerané na automatizáciu merania. Z danej potreby vznikla myšlienka vytvorenia monitorovacieho systému celej CNC frézky. Pri návrhu bol zvolený čo najefektívnejší prístup k zberu dát a ich následnému spracovaniu, až po vytvorenie

simulačných modelov. Ďalšou vlastnosťou, na ktorú sme sa zamerali je modulárnosť systému, ktorá umožňuje jednoducho pridávať ďalšie senzory na zaznamenávanie potrebných údajov. V nasledujúcich podkapitolách sa nachádza podrobnejší opis monitorovacieho systému a použitých komponentov na jeho realizáciu.

A. Koncept systému

Prvý návrh vznikol z potreby automatizovaného merania teplôt častí CNC stroja. Na meranie teploty sa využil senzor Dallas 18B20 [7]. Na Obr. 6 je zobrazená bloková schéma návrhu merania a vizualizácie teploty, ktorá zobrazuje prepojenie medzi jednotlivými uzlami monitorovacieho systému.



Obr. 6 Bloková schéma návrhu monitorovacieho systému

Monitorovací systém je založený na databáze InfluxDB, komunikačnom protokole MQTT, programovacím jazyku Python a vizualizačnom nástroji Grafana. Pre podrobnejšie vysvetlenie fungovania monitorovacieho systému je potrebné definovať jednotlivé nástroje, ktoré boli využité na jeho realizáciu.

InfluxDB je open-source databáza prispôbená na ukladanie časových radov a na zvládanie vysokej záťaže počas zápisu a dopytovania. Využíva sa na ukladanie časovo závislých údajov ako sú údaje senzorov IoT, analýza v reálnom čase, monitorovanie a metriky procesov. V našom systéme ju využívame na ukladanie všetkých nameraných údajov [8].

Grafana je webová platforma, ktorá umožňuje vizualizovať údaje vytváraním grafov, dashboardov a ďalších vizualizačných prvkov. Na zobrazenie údajov sa využíva priame prepojenie na databázy, kde sú uložené potrebné údaje. Ďalšou vlastnosťou aplikácie Grafana je možnosť definovať výstrahy. To znamená, že po prekročení sledovanej hodnoty sa vykoná určená úloha. V prípade merania teploty to môžeme využiť v systémoch, kde nastavíme spúšťač výstrahy na konkrétnu kritickú teplotu a po jej prekročení sa výstraha aktivuje a vykonajú sa definované úlohy [9].

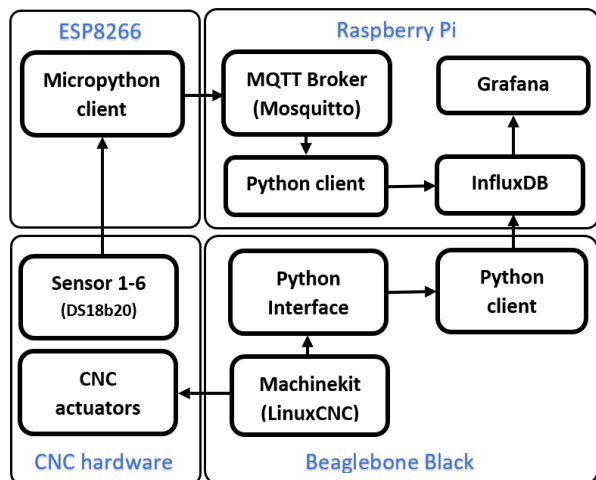
MQTT je skratkou pre MQ Telemetry Transport, využíva sa ako štandardným komunikačným protokolom pre oblasť IoT. Medzi hlavne výhody patrí nízke zaťaženie siete a vysoká spoľahlivosť. Základný komunikačný model protokolu je definovaný ako publikovať/odoberať (*publish/subscribe*) spôsob komunikácie. Ak chceme docieľiť prenos informácie z uzla A do uzla B, tak uzol A publikuje správu s konkrétnou témou (*topic*) do hlavného centra (*MQTT broker*). Na to aby sa správa dostala do uzla B. Uzol B musí odoberať tému, s ktorou bola publikovaná správa z uzla A. V prípade odoberania danej témy, *MQTT Broker* prepošle danú správu do uzla B [10].

Python je programovací jazyk so širokou škálou využitia. V našom systéme bol vybraný na zabezpečenie komunikácie medzi jednotlivými uzlami systému. Využíva sa na získavanie údajov zo senzorov a ich následne odosielanie do databázy. V budúcnosti ho plánujeme využiť aj na vytváranie simulačných modelov. Python verzia pre mikrokontroléry sa nazýva Micropython.

Komunikácia v systéme prebieha nasledovne, Micropython client(zariadenie ESP8266) získa údaje zo senzorov Dallas 18B20. Údaje sú publikované s témou „CNC/Temperature“ do MQTT Brokera. Python client odoberá danú tému, po prijatí správy s nameranými hodnotami sú údaje spracované a odoslané do databázy InfluxDB. Poslednou časťou monitorovacieho systému je vizualizačný nástroj, ktorý ma vytvorené priame prepojenie s databázou InfluxDB. To znamená, že je schopný sa dopytovať po dátach uložených v databáze. Grafana ponúka široké možnosti zobrazenia údajov, závisí na koncovom užívateľovi systému, aká možnosť vizualizácie údajov bude zvolená. Pre naše účely sme si nakonfigurovali vizualizáciu prispôbenú pre potreby tepelného modelu CNC fréžky, jej podrobný opis sa nachádza v podkapitole 2.4. Modulárnosť monitorovacieho systému spočíva v neobmedzenom pridávaní zdrojov dát, alebo v umožnení monitorovania viacerých zariadení súčasne. Ďalší zdroj dát, ktorý chceme pridať musí spĺňať podmienku schopnosti komunikovať MQTT protokolom alebo musí byť schopný priamo zapisovať údaje do databázy InfluxDB.

B. Rozšírenie monitorovacieho systému

V predchádzajúcich častiach sme opisovali základný návrh monitorovacieho systému, jedným z hlavných vlastností systému bola modulárnosť. V tejto podkapitole sa nachádza opis prvého rozšírenia systému. V prvotnom vlákne sme sa zamerali na meranie teploty. Na to aby sme získali potrebné informácie na vytvorenie tepelného modelu potrebujeme zaznamenávať ďalšie údaje CNC fréžky. Preto bol systém rozšírený o monitorovanie dostupných informácií CNC fréžky. Bloková schéma kompletného monitorovacieho systému CNC fréžky je zobrazená na Obr. 7.

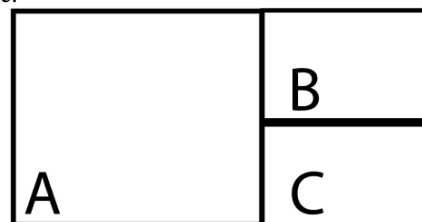


Obr. 7 Bloková schéma monitorovacieho systému

Riadiaci systém CNC fréžky sa nazýva Machinekit, čo je upravená verzia softvéru LinuxCNC pre mikrokontroléry. Machinekit umožňuje pomocou Python knižníc načítať spustenú konfiguráciu stroja získať základné informácie a vykonať základné príkazy CNC stroja. Preto sa ako ďalšia vetva monitorovacieho systému vytvoril Python skript. Jeho úlohou je získať vybrané údaje o CNC stroji a odoslať ich priamo do databázy. Následne sme nakonfigurovali nástroj Grafana na zobrazovanie nameraných údajov.

C. Grafická vizualizácia údajov

Táto podkapitola sa zameriava na zobrazenie finálnej konfigurácie monitorovacieho systému CNC fréžky. Obr. 8 zobrazuje rozloženie jednotlivých prvkov vizualizácie. Tento typ interpretácie bol zvolený z dôvodu jej veľkosti. Jednotlivé časti budú prezentované samostatne, v praktickom využití sú zobrazované všetky zobrazovacie prvky spolu na jednej obrazovke.



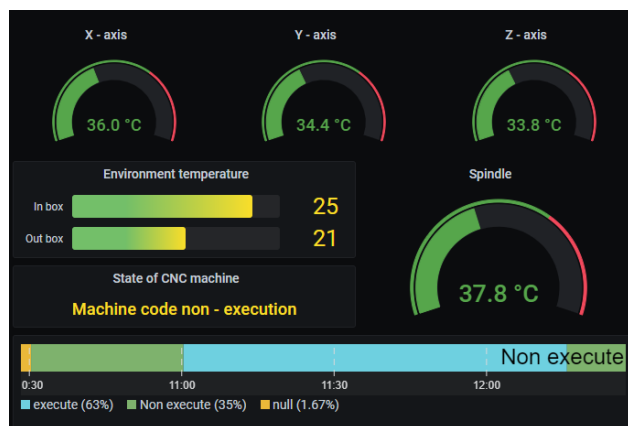
Obr. 8 Základný koncept rozloženia vizualizácie v rozhraní Grafana

Časť vizualizácie „A“ zobrazuje grafy teplôt a informácie o aktuálnej rýchlosti pohonov CNC fréžky, je zobrazená na Obr. 9. Prvý graf vľavo hore zobrazuje aktuálne rýchlosti krokových motorov. Rýchlosť je zaznamenávaná frekvenciou 10 Hz. Graf vľavo v strede zobrazuje teploty krokovým motorov, ktorá je zaznamenávaná každých 5 sekúnd. Graf vľavo dole zobrazuje teplotu vretena a jeho aktuálnu rýchlosť. V pravej časti Obr. 9 je umiestnená tabuľka s vykonanými inštrukciami CNC fréžky. Zobrazovaný obsah vizualizácie a zaznamenávané dáta do databázy sú prispôbené potrebám návrhu tepelného modelu



Obr. 9 Grafy teplôt, rýchlosti pohonov a tabuľka G-kódu

Na nasledujúcom Obr. 10 je znázornená časť vizualizácie „B“, ktorá zobrazuje vizualizáciu aktuálnych teplôt pohonov a okolia. V spodnej časti je zobrazená prezentácia stavov CNC stroja, ktorého úlohou je definovať kedy bol vykonávaný obrábací program a kedy bol stroj mimo prevádzky.



Obr. 10 Vizualizácia teplôt častí stroja a informácie o stave zariadenia (časť B)

Na Obr. 11 je zobrazený úsek vizualizácie C. Posledná časť vizualizácie zobrazuje informácie o obrábacom procese. Medzi hlavné zaznamenávané informácie, ktoré ovplyvňujú generovanie tepla patrí materiál obrobku, obrábací nástroj a stratégia frézovania. Ďalšími zobrazovanými parametrami sú kalibračné informácie zariadenia, ktoré ovplyvňujú maximálnu možnú rýchlosť, akceleráciu pohonov CNC zariadenia a generovanie riadiacich signálov.

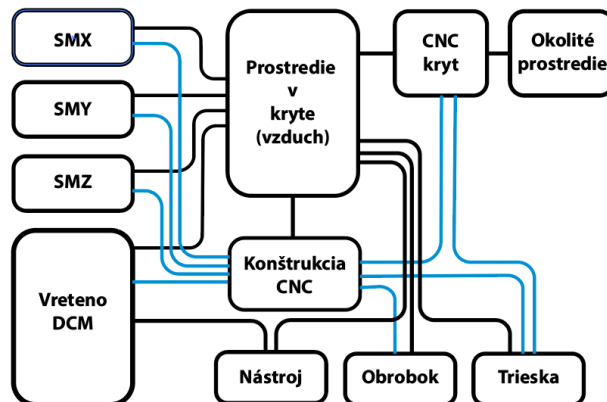
Milling information							
tool	workpiece material	action info					
none	none	pocket-milling					
machine name	servo period	default velocity	max linear velocity				
CNC Mill-M KTPF	1000000	20	25				
Stepper motors calibration informations							
axis	max_vel	max_acc	scale	dirsetup	dirhold	steplen	stepspace
AXIS_0	60	30	67	200	200	1000	1000
AXIS_1	60	30	67	200	200	1000	1000
AXIS_2	60	30	67	200	200	1000	1000

Obr. 11 Informácie o obrábaní a kalibrácii zariadenia (časť C)

IV. TEPELNÝ MODEL ČINNOSTI CNC FRÉZKY

Pri návrhu tepelného modelu bolo potrebné uvažovať s všetkými prebiehajúcimi procesmi počas procesu frézovania. Obr. 12 zobrazuje blokovú schému tepelného modelu činnosti CNC frézky. V blokovej schéme sú prezentované hlavné časti systému a ich vzájomná interakcia. Modré a čierne čiary predstavujú tepelný prenos medzi spojenými blokmi. Tepelný prenos je tvorený prenosom energie kondukciami, konvekciou a radiáciou. Vo väčšine prípadov jeden z typov prenosu

prevláda nad ostatnými do takej miery, že zvyšné typy prenosu z dôvodu zjednodušenia modelu je možné zanedbať. Určenie konkrétnych spôsobov prenosu tepla medzi časťami CNC frézky sa nachádza v opise vytvoreného tepelného modelu.



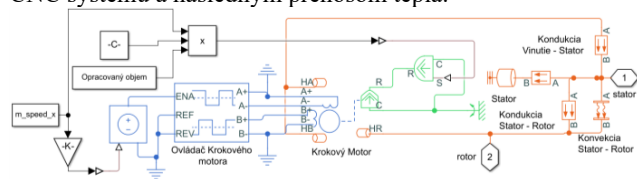
Obr. 12 Bloková schéma tepelného modelu

Blokovú schému zobrazenú na Obr. 12 sme využili na vytvorenie tepelného modelu. Daná bloková schéma opisuje nasledovné vzájomné pôsobenie systému. V priebehu frézovania nastáva výmena tepla z krokových motorov do prostredia v kryte a konštrukcie CNC stroja. Teplo vygenerované vretenom je odovzdávané konštrukcii a prostrediu. V priebehu frézovania sa až 95% spotrebovaného výkonu premení na teplo. Teplo vytvorené v priebehu rezania materiálu je rozdelené medzi nástroj, obrobok a triesku. Časť vytvoreného tepla v nástroji sa presúva do vretena a prostredia. Najväčšia časť tepla vytvoreného v priebehu frézovania absorbuje trieska, najmenšiu časť absorbuje obrobok. Teplo sa z obrobku a triesky šíry do prostredia, konštrukcie CNC frézky a krytu stroja. CNC stroj je umiestnený v kryte, preto prostredie v kryte je v interakcii s každou časťou CNC stroja. Vplyv okolitého prostredia na časti CNC stroja je oslabený krytom CNC zariadenia.

A. Model krokového motora

V prvej časti je opísaná simulácia krokového motora. Model sa skladá z riadiaceho vstupu m_speed_x , ovládača krokového motora a modulu krokového motora. Modul Krokový Motor je hlavnou časťou opisovanej simulácie. V nastaveniach modulu je možné zvoliť simuláciu tepelných strát. Dané nastavenia umožnia simulovať straty rozdelené medzi straty železa a Joulové straty. Tepelný model krokového motora je doplnený o blok tepelnej hmoty, ktorý zastupuje materiál tvorený statorom. Vinutie krokového motora je umiestnené v statore. Preto medzi statorom a vinutím je vložený blok prenosu tepla kondukciami. Rotor motora je zdrojom tepelných strát v železe. Prestup tepla medzi statorom a rotorom je simulovaný kondukciami a konvekciou. Prestup tepla prúdením je zavedený z dôvodu vzduchovej medzery medzi rotorom a statorom. Prestup tepla vedením zohľadňuje kontakt ložísk, na ktorých je umiestnený rotor. Výstup z podsystemu krokového motora je teplota

statora a teplota rotora z dôvodu kontaktu s ďalšími časťami CNC systému a následným prenosom tepla.

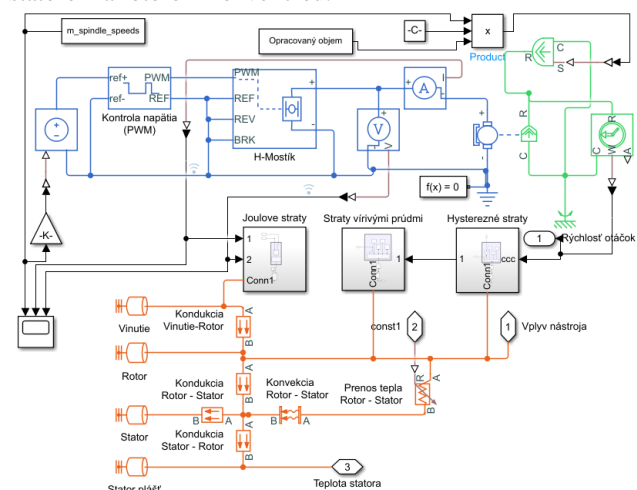


Obr. 13 Tepelný model krokového motora

Ovládač a blok motora je v režime *Averaged*. Daný režim umožňuje riadenie motora definovaním rýchlosti otáčania. Blok *m_speed_x*, načíta údaje z Matlab súboru, ktoré definujú pohyb krokového motora na základe analýzy G_kódu. Ďalším načítaným údajom je signál *Opracovaný objem*, ktorý definuje aktuálne zaťaženie pohonu. *Opracovaný objem* ovplyvňuje silu, ktorú musí krokový motor vynaložiť pri pohybe v priebehu frézovania. Záťaž krokového motora je simulovaná zdrojom ideálneho točivého momentu, ktorý vytvára silu pôsobiacu proti momentu otáčania krokového motora. Veľkosť vytvoreného momentu je závislá od rýchlosti pohybu, záťaže krokového motora a konštanty obrobiteľnosti materiálu.

B. Model vretena CNC fréžky

Simulácia jednosmerného motora s permanentnými magnetmi obsahuje blok PWM modulácie, h-mostík, DC motor, záťaž a bloky reprezentujúce tepelné pomery pohonu. Model simuluje Joulove straty vytvárané vo vinutí, hysterézne straty a straty vytvárané vírivými prúdmi, ktoré sú zdrojmi tepla v rotore. Vinutie a rotor sú prezentované ako samostatné tepelné hmoty, medzi ktorými nastáva prenos tepla kondukciou. Statorová časť motora je rozdelená na dve časti. Rozdelenie bolo zadefinované z dôvodu vplyvu ventilátora na ochladzovanie motora. Medzi vnútornou časťou statora a vonkajšou časťou statora je zadefinovaný tepelný prenos kondukciou. Analogicky ako pri krokovom motore, stator je umiestnený v ložiskách čo zabezpečuje prenos tepla medzi statorom a rotorom konvekciou.

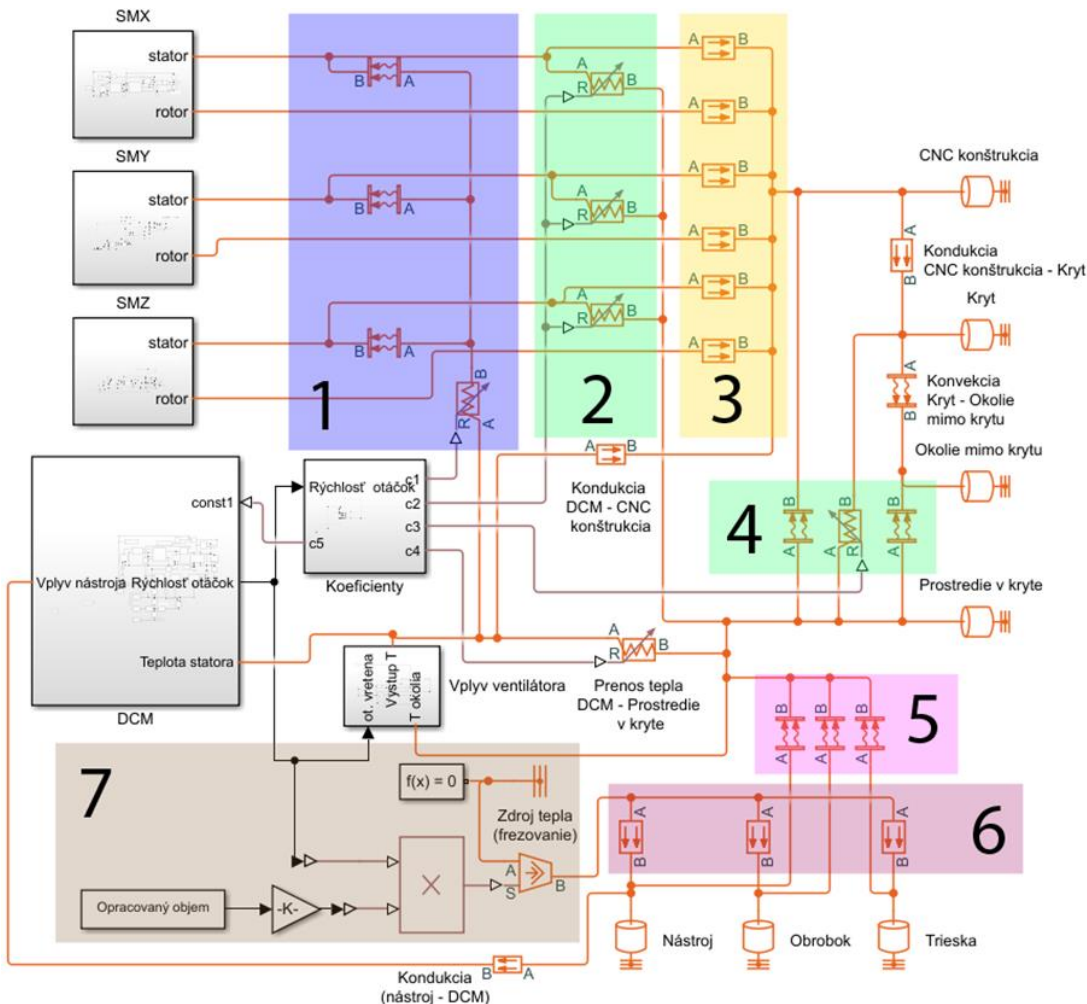


Obr. 14 Tepelný model vretena CNC fréžky

Pri zvýšení otáčok vretena sa znižuje hodnota odporu, čím sa zabezpečí väčší prenos tepla. Prenos tepla variabilným rezistorom nahrádza prenos tepla konvekciou, kondukciou a radiáciou. Blok *m_spindle_speed* importuje riadiaci signál vretena. Ďalší importovaný riadiaci signál predstavuje *Opracovaný objem*, ktorý ovplyvňuje aktuálny výkon vynaložený pri frézovaní.

C. Kompletný tepelný model činnosti CNC fréžky

Obr. 155 zobrazuje vytvorený tepelný model činnosti CNC fréžky. Podsystemy SMX,SMY a SMY predstavujú bloky krokových motorov jednotlivých osí. Podsystem DCM obsahuje simuláciu vretena. Ďalšie doplnkové podsystemy v simulácii obsahujú vplyv ventilátora a výpočet koeficientov pre ovládanie premenlivého tepelného odporu. Prenos tepla medzi časťami systému je definovaný spôsobom, ktorý sa najviac podieľa na danom prenose. Zvyšné spôsoby prenosu sú buď zahrnuté v primárnom spôsobe prenosu tepla, alebo sú nahradené blokom premenlivého tepelného odporu. V simulácii sa bloky variabilného tepelného odporu využili z dôvodu simulovania vplyvu ventilátora vretena. Koeficienty boli získane estimáciou parametrov. Každý tepelný prenos simulovaný tepelným odporom si vyžadoval vlastný koeficient, lebo vplyv ventilátora na jednotlivé komponenty bol rôzny. Vo využívanom spôsobe simulácie sa predpokladá, že jednotlivé časti CNC fréžky sú tvorené homogénnymi telesami s rovnakými tepelnými vlastnosťami v celom telese. Pri vytváraní modelu je potrebné brať ohľad na dané zjednodušenie a prispôbiť model tak, aby vzhľadom na požiadavky čo najlepšie opisoval simulovaný systém. Opis využitia konkrétnych blokov modelu je nasledovný. Časť vzájomného ovplyvňovania pohonov je simulovaná blokmi umiestnenými v oblasti 1. Dané bloky simulujú prenos tepla žiarením a vírením vzduchu, ktoré nespadá do predpokladu homogénneho prostredia v kryte CNC fréžky. Variabilné tepelné odpory sprostredkujú výmenu tepla medzi krokovými motormi a prostredím. Bloky prenosu tepla vedením v časti 3 simulujú prenos tepla medzi krokovými motormi a konštrukciou CNC fréžky. Prostredie v kryte(vzduch) je v interakcii s CNC konštrukciou, krytom a okolím mimo krytu blokmi prenosu tepla zobrazenými v časti 4. Interakcia prostredia v kryte a mimo krytu nastáva v prípade otvorenia krytu pri výmene obroku, nástroja alebo nedokonalým tesnením krytu. Časti modelu 5,6,7 sú spojené s vytváraním tepla v procese frézovania. Bloky v časti 7 definujú vzniknuté teplo pri procese obrábania, ktoré sa následne rozdelí medzi nástroj, obrobok a triesku. Pomer rozdelenia tepla vzniknutého pri frézovaní je určený blokmi v oblasti 6 . Časť tepla vytvoreného v nástroji sa odovzdá kondukciou vretenu. Zvyšná časť tepla vytvoreného v priebehu frézovania sa skrz nástroj, obrobok a triesku uvoľní do prostredia. Dané zjednodušenie nemalo negatívny vplyv na výsledky simulácie.



Obr. 15 Tepelný model činnosti CNC frézky

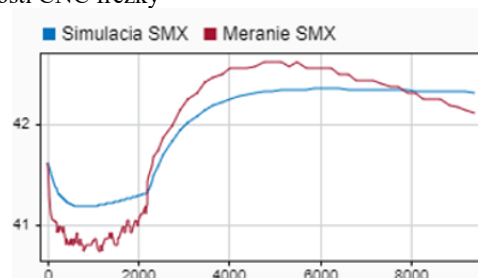
V. POROVNANIE VÝSLEDKOV SIMULAČNÉHO SYSTÉMU S REÁLNYM SYSTÉMOM

Chyba! Nenašiel sa žiaden zdroj odkazov. 1 zobrazuje hodnoty tolerancií pohonov v porovnaní s reálnym systémom. Zvážením výsledných parametrov modelu môžeme konštatovať, že model spĺňa predpoklady funkčného tepelného modelu CNC frézky. V nasledujúcej kapitole sa zameriame na využitie modelu v praktickom riešení predikcie teploty v monitorovacom systéme.

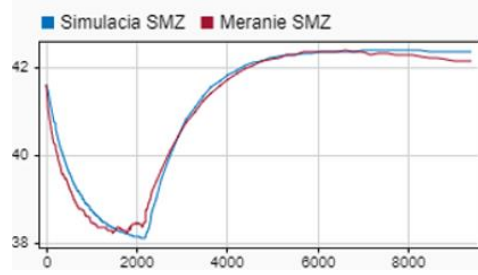
Tab. 1 Tolerancie modelu pohonov

Simulovaný pohon	Absolútna tolerancia[°C]	Relatívna tolerancia[%]
Vreteno	1,3	2,7
SMX	0,5	1,2
SMY	0,7	1,8
SMZ	0,6	1,5

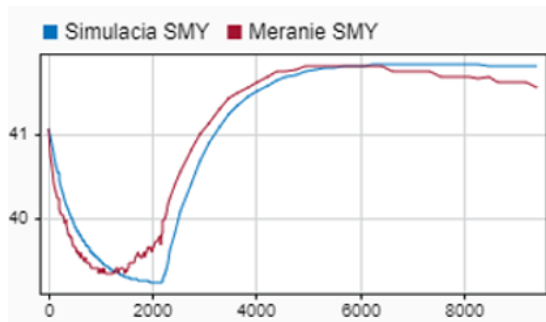
Všeobecné porovnanie príbehov teplôt tepelného modelu a reálneho CNC zariadenia je zobrazené na Obr. 16 až Obr. 19.



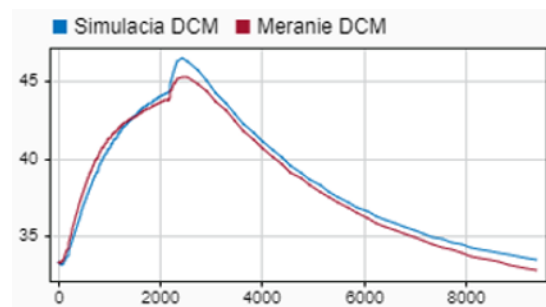
Obr. 16 Porovnanie teplôt krokového motora osi X



Obr. 17 Porovnanie teplôt krokového motora osi Z



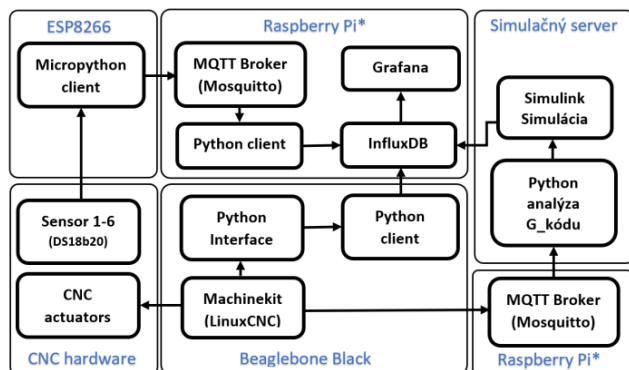
Obr. 18 Porovnanie teplôt krokového motora osi X



Obr. 19 Porovnanie priebehu teplôt jednosmerného motora

VI. PRAKTICKÁ IMPLEMENTÁCIA TEPELNÉHO MODELU ZA ÚČELOM PREDIKCIE

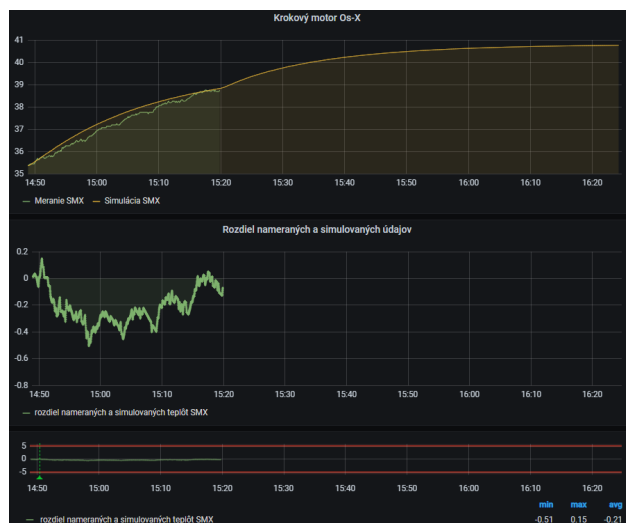
Na implementáciu tepelného modelu sa využil vytvorený monitorovací systém. Obr. 20 zobrazuje blokovú schému implementácie. Riadiaci program CNC frézky komunikuje so Simulačným serverom prostredníctvom MQTT protokolu. V celom systéme sa nachádza iba jeden MQTT broker, z dôvodu prehľadnosti sa v diagrame nachádza dva krát.



Obr. 20 Bloková schéma implementácie tepelného modelu

Po vykonaní simulácie tepelnej činnosti CNC frézky sa údaje zobrazia vo vizualizačnom nástroji. Využitie predikcie sa aplikovalo na ochranu pohonov pred prehriatím. Obr. 21 zobrazuje rozhranie pre krokový motor osi X. Vrchná časť obrázka zobrazuje graf priebehu teplôt simulácie a reálneho

motora. Graf v strede zobrazuje rozdiel meraného a simulovaného priebehu. Je to základný údaj v porovnávaní a vyhodnocovaní CNC systému. Grafana umožňuje robiť rôzne matematické operácie s dátami. Nevýhodou je, že po určitých typoch transformácii údajov napríklad ako je odčítanie dvoch priebehov, nie je možné na daný graf aplikovať výstrahy. Správna konfigurácia výstrah na rozdiel meranej a simulovanej veličiny je kľúčovým faktorom v hlásení vzniku možných porúch. Preto bolo potrebné doprogramovať do Python skriptu funkciu porovnávania simulovanej a aktuálnej veličiny. Po vykonaní rozdielu údajov je hodnota uložená do InfluxDB a následne vykreslená v grafe č.2 a č.3. V grafe č.3 je nastavená výstraha na prekorenie rozdielu teplôt v rozsahu $\pm 5^{\circ}\text{C}$. Po jej prekorení a splnení podmienky sa odošle upozornenie e-mailom. Na každú výstrahu je možné definovať samostatnú formu e-mailu.



Obr. 21 Vizualizácia predikcie teploty pohonu v monitorovacom systéme

Analogicky sme pristupovali k tvorbe grafického rozhrania pre vreteno CNC frézky.

ZÁVER

Podrobnou analýzou tepelných zdrojov CNC frézky sme si definovali všetky časti stroja, ktoré bude potrebné analyzovať pri fyzikálnom modelovaní systému.

Fyzikálna simulácia je tvorená v prostredí Matlab Simulink. Na vytvorenie správnej simulácie bolo potrebné získať všetky dostupné a merateľné parametre pohonov a iných častí stroja. Na získanie neznámych konštánt sa využil nástroj Parameter Estimator. Simulácie zahŕňajú výpočty parametrov pohonov, čiastočne simulácie jednotlivých častí a kompletný tepelný model CNC frézky.

Na analyzovanie CNC stroja sme v priebehu práce vyvinuli monitorovací systém, ktorý umožňuje zber údajov z meraného zariadenia. Medzi hlavné analyzované údaje patrí teplota, informácie o pohybe pohonov a údaje o aktuálnom obrábaní a konfigurácii zariadenia. Všetky namerané údaje sú ukladané v

databáze InfluxDB. Ďalej sa vytvorilo prepojenie databázy s Matlabom za účelom automatizácie práce pri vytváraní tepelných modelov.

V predposlednej časti sa vykonala verifikácia a posúdenie využitia celého systému. V prípade správnej činnosti modelu a CNC zariadenia, model nepresahuje maximálnu relatívnu odchýlku 3% a absolútnu odchýlku 1,5 °C. Model pracuje s postačujúcou presnosťou na využívanie modelu pre kontrolu činnosti CNC zariadenia a umožňuje správne detegovať možné ohrozenia monitorovaného systému.

REFERENCES

- [1] FECKO, Branislav: NÁVRH A RELIZÁCIA 3 – OSOVEJ CNC FRÉZKY (bakalárska práca). Technická univerzita v Košiciach, 2016.
- [2] POLIAK, František; FEDÁK, Viliam; ZBORAY, Ladislav; Elektrické pohony. Bratislava: Alfa – vydavateľstvo technickej a ekonomickej literatúry, 1987, 614 s, 063-568-87 EPO
- [3] VASILKO, Karol; Analytická teória trieskového obrábania. Fakulta výrobných technológií TU v Košiciach, 2007, ISBN 978-80-8073-759-7
- [4] ČEP, Robert; PETRŮ, Jana, Experimentální metody v obrábění: učební text. Ostrava: VŠB-TUO, Fakulta strojní, Katedra obrábění a montáže, 2011. SBN: 978-80-248-2533-5
- [5] SECO TOOLS: Managing thermal loads in milling processes [online] [cit: 21.7. 2021.]. Dostupné na internete : <<https://www.secotools.com/article/21483?language=en>>.
- [6] SECO TOOLS: Controlling mechanical loads in milling operations [online] [cit: 21.7. 2021.]. Dostupné na internete : <<https://www.secotools.com/article/21485?language=en>>.
- [7] Dallas Semiconductor, DS18B20, [online] [cit: 10.2. 2021.] Available on the Internet: <https://html.alldatasheet.com/html-pdf/58557/-DALLAS/DS18B20/181/1/DS18B20.html>
- [8] InfluxData, InfluxDB, [Online] [cit: 17.7. 2020.] Available on the Internet: <https://www.influxdata.com/products/influxdb/>
- [9] Grafana Labs, Grafana, [Online] [cit: 18.7. 2020.] Available on the Internet: <https://grafana.com/grafana/>
- [10] IoT Portal, MQTT, [Online] [cit: 18.7. 2020.] Available on the Internet: <https://www.iot-portal.cz/2016/05/24/mqtt/>

Optimalizácia výkonových strát v smart sieťach

Daniel Pál

Katedra elektroenergetiky
FEI TUKE
Košice, Slovenská republika
daniel.pal@tuke.sk

Lubomír Beňa

Katedra elektroenergetiky
FEI TUKE
Košice, Slovenská republika
lubomir.bena@tuke.sk

Abstrakt — Článok sa zaoberá problematikou zníženia výkonových strát v inteligentných sieťach. V úvode je vysvetlený súčasný stav výroby elektrickej energie. V druhej kapitole sú podrobne vysvetlené súčasné poznatky v oblasti inteligentnej siete. V ďalšej časti sú analyzované jednotlivé možnosti, ktorými je možné minimalizovať vzniknuté činné výkonové straty v sieti. V nasledujúcej časti sú popísané optimalizačné algoritmy, ktoré je možné používať v oblasti optimalizačných úloh. Potom nasleduje výpočet ideálneho jalového výkonu v regulačných uzloch pomocou enumeratívnej metódy a pomocou SOMA algoritmu. V záverečnej časti sú vyhodnotené získané výsledky riešenia optimalizačných úloh pomocou navrhnutých algoritmov.

Kľúčové slová — enumeratívna metóda; Inteligentné siete; minimalizácia; SOMA

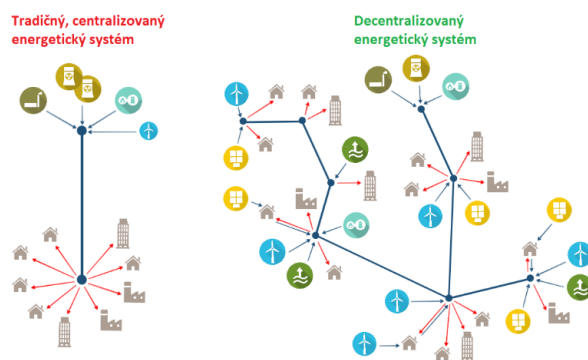
I. ÚVOD

V dnešnej dobe je používaná centralizovaná výroba elektrickej energie. Centralizovaná výroba predstavuje takú formu výroby elektrickej energie, v ktorej je elektrická energia vyrábaná centrálnou, vo veľkých, klasických elektrárňach. Medzi takéto elektrárne patria jadrové elektrárne, tepelné elektrárne a vodné elektrárne. Elektrárne sa nachádzajú vo veľkých vzdialenostiach, niekedy aj 100 - 200 km od koncových užívateľov. Aby dochádzalo k čo najmenším stratám pri prenose elektrickej energie prostredníctvom vedenia, následne po výrobe elektrickej energie je napätie zdroja transformované do vyššej napäťovej hladiny. V blízkosti spotrebiteľa je napätie transformované späť do nižšej napäťovej hladiny, ktoré je akceptovateľné pre spotrebiteľov [1].

V prípade decentralizovanej výroby je elektrická energia vyrábaná nielen centrálnou, ale jej výroba prebieha aj v blízkosti spotrebiteľov z dôvodu, aby sa výroba a spotreba nachádzala čo najbližšie k sebe. Aj v tomto prípade je väčšie množstvo elektriny vyrábané v klasických, centrálnych elektrárňach. Nevýhodou tohto formátu je, že akonáhle nastane porucha na hlavných vedeniach, mnoho spotrebiteľov ostáva bez elektriny. Z uvedeného dôvodu sú v decentralizovanej výrobe klasické elektrárne rozšírené malými elektrárňami, zabezpečujúcimi, že výpadok nejakého vedenia nebude mať vplyv na celú sústavu, ale iba na príslušnú časť, v ktorej došlo k poruche. V uvedenom prípade malé elektrárne sú schopné čiastočne nahradiť chýbajúce množstvo elektriny, nakoľko elektrická energia k zákazníkovi prichádza nie iba z jedného, ale z viacerých smerov. Malé elektrárne sú zvyčajne obnoviteľnými

zdrojmi energie, ako fotovoltika, veterné turbíny alebo biomasa.

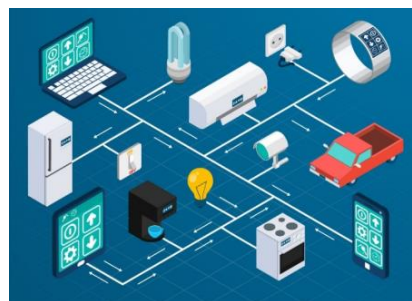
Rozdiely medzi centralizovanou a decentralizovanou výrobou sú zobrazené na nasledujúcom obrázku (Obr. 1).



Obr. 1 Porovnanie centralizovaného a decentralizovaného energetického systému [2]

II. INTELIGENTNÁ (SMART) SIEŤ

V 21. storočí sa s inteligentnými technológiami môže stretnúť vo všetkých oblastiach, stali sa súčasťou nášho života. Jednotlivé Smart technológie a zariadenia sú zobrazené na Obr. 2. Nachádzajú sa medzi nimi napr. Smartphones (mobilné telefóny), Smart Clock (inteligentné hodiny), Smart Lamps (inteligentné lampy), Smart Camery (inteligentné kamery), Smart Fridges (inteligentné chladničky) a mnohé iné. Technológie sa neustále rozvíjajú, vyvíjajú sa nové technologické zariadenia, ktorých cieľom je zlepšenie a uľahčenie života a úkonov v každodennom živote vo všetkých oblastiach.



Obr. 2 Smart zariadenia [3]

Všeobecne zadefinovať, že kedy považujeme sieť za inteligentnú, je dosť problematické, nakoľko každý štát, každý kontinent má k tejto problematike iný prístup. Z hľadiska posudzovania Európskej komisie inteligentnú sieť je možné opísať na základe nasledujúcich hľadísk:

- flexibilita – reaguje na požiadavky spotrebiteľov;
- dostupnosť – do siete je možné pripojiť všetky nové zdroje, vrátane obnoviteľných;
- spoľahlivosť – zaisťuje bezpečnosť a kvalitu dodávky elektrickej energie v každom okamihu;
- hospodárnosť – efektívne riadenie siete [4].

Podľa amerického úradu elektriny (The Office of Electricity) inteligentné siete charakterizujú nasledujúce body:

- vyššia efektívnosť pri prenose elektrickej energie
- nižšie straty
- rýchlejšia obnova siete v prípade poruchy
- zníženie prevádzkových a riadiacich nákladov – zníženie cien energie pre spotrebiteľov
- bezpečnosť – odolnosť siete voči fyzickému alebo kybernetickému zásahu
- integrácia obnoviteľných zdrojov energie [5]

TABUĽKA I slúži na prehľadné porovnanie súčasnej a inteligentnej siete podľa rôznych kategórií.

	Súčasná sieť	Inteligentná (SMART) sieť
Meranie	Elektromechanické	Digitálne
Komunikácia	Jednosmerná	Obojsmerná
Výroba	Centralizovaná	Centralizovaná a decentralizovaná
Monitorovanie	Niekoľko senzorov	Senzory v celej sieti
Prevádzka	Manuálne sledovanie zariadení	Sledovanie na diaľku
Údržba	Časový harmonogram údržby	Údržba v závislosti od sledovania
Regulácia toku energie	Obmedzená	Rozsiahla
Spoľahlivosť	Náchylné na poruchy. Môžu vzniknúť rozsiahle výpadky	Ochrana v reálnom čase
Obnova siete	Manuálna	Automatická, samo obnova siete

III. MOŽNOSTI ZNÍŽENIA STRÁT V SIEŤACH

Pri náraste podielu obnoviteľných zdrojov energie sa klasická centralizovaná schéma mení na decentralizovanú výrobu elektrickej energie. Vďaka obnoviteľným zdrojom energie je v súčasnosti ekonomicky výhodné vyrábať elektrickú energiu priamo na mieste spotreby. Pri prebytku energie je pomocou akumulátorov možné aj jej ukladanie na použitie v čase potreby. V elektrickej sieti sa preto čoraz častejšie objavujú nové zdroje a nové technológie s ohľadom

na rozvoj informačných technológií. Vzniká potreba novej komunikácie medzi týmito komponentami. Z elektrickej siete po týchto zmenách vzniká inteligentná sieť (SMART GRID).

Téma inteligentných sietí je v dnešnej dobe čoraz aktuálnejšia, keďže elektrickú energiu je potrebné využívať čo najefektívnejšie. Energiu je potrebné prenášať od výroby až do miesta spotreby s najmenšími možnými stratami. Momentálne vo výrobe elektrickej energie stále dominujú fosílna elektrárne, avšak zásoby fosílnych palív každoročne klesajú. Jednu z týchto alternatív predstavuje používanie obnoviteľných zdrojov energie. Podiel obnoviteľných zdrojov sa zvyšuje postupne na celom svete, a preto je viac ako nevyhnutné zaviesť elektrický systém, ktorý môže súčasne spĺňať podmienky bezpečnosti a spoľahlivosti. Jednou z možností je použitie technológií SMART, ktoré zabezpečia, aby boli sledované všetky parametre, ktoré sú dôležité z hľadiska prevádzky.

Pre zabezpečenie stálej dodávky elektrickej energie a kvalitatívnych ukazovateľov pri prijateľných vstupných nákladoch je potrebné minimalizovať straty, ktoré vznikajú pri prenose elektrickej energie. Predpokladá sa, že zásoby fosílnych zdrojov sa minú už v blízkej budúcnosti, čo si vyžaduje v blízkej budúcnosti hľadanie a nájdenie adekvátnej náhrady, aby výroba elektrickej energie bola bezproblémová aj po vyčerpaní jednotlivých zásob. Jednou z možností je používanie veľkých obnoviteľných zdrojov energie ako náhrada za klasické elektrárne, v blízkosti niekdajších elektrární by vznikli nové veľké elektrárne. Centralizovaná výroba by bola zachovaná aj naďalej. Druhou možnosťou je náhrada veľkých elektrární väčším počtom malých elektrární, ktoré by boli rozptýlené v sieti. Vznikla by takzvaná distribuovaná generácia, pomocou ktorej by vyrobená elektrická energia bola spotrebovaná hneď v blízkosti výroby, a preto by ju nebolo potrebné prenášať do vzdialenejších oblastí, čím by sa následne znížili aj straty vznikajúce pri prenose. Stopercentnú účinnosť nie je možné dosiahnuť, straty vzniknú vždy, ale ich možné minimalizovať. Práca je zameraná na dva možné spôsoby, ktoré by mohli straty znížiť, a to:

- využitie malých elektrární a distribúovanej generácie,
- rekonfigurácia existujúcich sietí,

A. Využitie distribúovanej generácie

Pri aplikovaní distribúovanej generácie sú zvyčajne používané obnoviteľné zdroje energie. Je oveľa jednoduchšie postaviť malé zdroje (napr. fotovoltaiky) ako veľké zdroje, napr. jadrovú elektráreň. Typické výkonové hodnoty distribúovanej generácie sú nasledovné:

- Vodné elektrárne
 - Malé → 0,025 MW – 1 MW
 - Veľké → 1 MW – 100 MW
- Veterné elektrárne → 200 W – 3 MW
- Solárne elektrárne → 20 W – 100 kW
- Geotermálne elektrárne → 5 MW – 100 MW
- Biomasa → 0,1 MW – 20 MW,

- Batéria → 0,5 MW – 5 MW [7][8].

Výhody používania distribuovanej generácie:

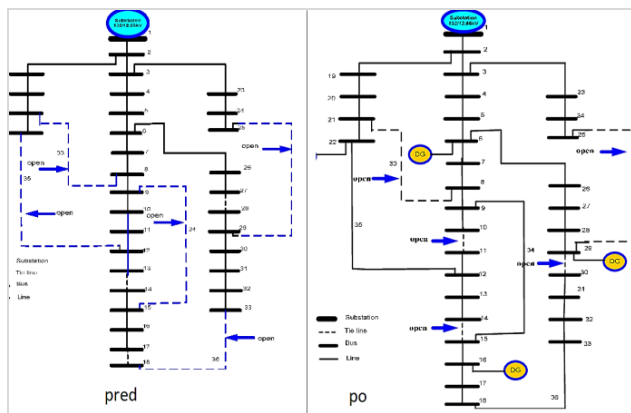
- zvyšovanie parametrov elektriny z hľadiska kvality a spoľahlivosti,
- znižovanie strát pri prenose energie z miesta výroby do miesta spotreby,
- znižovanie produkcie skleníkových plynov,
- výpadok jedného zdroja nespôsobuje problémy v sieti, lebo ostatné zdroje, ktoré sú rozptýlene umiestnené v sieti, sú schopné kompenzovať chýbajúci zdroj [7].

B. Rekonfigurácia siete

Tento pojem opisuje proces implementácie diaľkovo ovládaných úsekových vypínačov, ktoré sú vopred vybrané v sieti. Vypínače môžu byť zapnuté a vypnuté, a taká zmena topológie spôsobuje zmenu strát v sieti. Rekonfiguráciu znázorňuje Obr. 3. Na ľavej strane obrázku je znázornená sieť pred rekonfiguráciou a na pravej strane po rekonfigurácii. Niektoré úseky sú zmenené, sú vypnuté, čím je možné straty ovplyvniť. V tomto prípade boli zmenené vedenia medzi uzlami 9 – 15, 10 – 11, 12 – 22, 14 – 15, 18 – 33 a 29 – 30. Najčastejšie sa tieto metódy používajú spolu s distribuovanou generáciou [9][10].

Rekonfigurácia siete okrem znižovania strát má aj iné výhody. Medzi výhody je možné zaradiť:

- zníženie výpadkov siete,
- regulácia napätia,
- regulácia preťaženia siete [11].



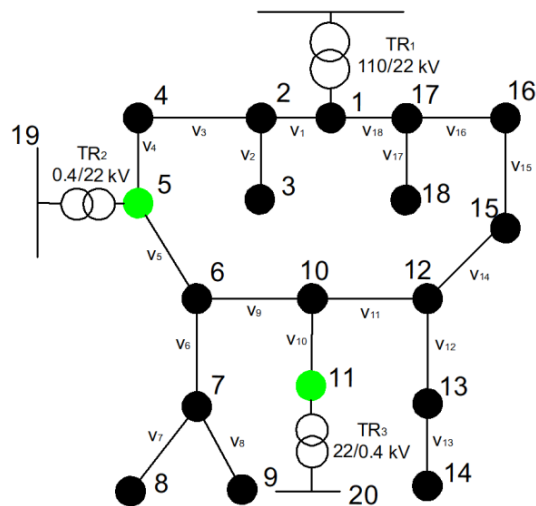
Obr. 3 Rekonfigurácia siete [12]

IV. SIMULOVANÁ SIETĚ

Schéma elektrickej siete sa nachádza na Obr. 4. Elektrická sieť sa skladá z 18 uzlov, ktoré sú spravidla odberovými uzlami. Do dvoch uzlov (5, 11) sú pripojené vn/nn transformátory. V týchto bodoch je možné napr. pomocou striedačov zmeniť jalový výkon, ktorý bude ovplyvňovať uzlové napätie a následne aj straty v sieti. Parametre vedení sa

nachádzajú v TABUĽKA II. Aby aj program vedel pracovať s transformátormi, bola schéma dodatočne doplnená ešte dvomi uzlami, a to uzlami 19, 20. Dané dodatočné uzly slúžia na umožnenie používania impedancie transformátorov. Používané impedancie transformátorov sa nachádzajú v TABUĽKA II.

Parametre uzlov sa nachádzajú v TABUĽKA III.



Obr. 4 Schéma siete

TABUĽKA II
Parametre vetiev

Názov ved.	R (Ohm)	X (Ohm)	G (Siemens)	B (Siemens)
v. 1	1,500	1,650	0	1,739E-05
v. 2	2,001	1,027	0	9,347E-06
v. 3	1,200	1,320	0	1,391E-05
v. 4	1,170	1,287	0	1,356E-05
v. 5	1,904	1,311	0	1,263E-05
v. 6	2,004	1,380	0	1,33E-05
v. 7	2,208	1,133	0	1,031E-05
v. 8	1,656	0,850	0	7,735E-06
v. 9	1,904	1,311	0	1,263E-05
v. 10	1,403	0,966	0	9,307E-06
v. 11	1,453	1,001	0	9,640E-06
v. 12	1,518	0,779	0	7,091E-06
v. 13	1,449	0,743	0	6,768E-06
v. 14	1,202	0,828	0	7,978E-06
v. 15	0,960	1,056	0	1,113E-05
v. 16	1,260	1,386	0	1,460E-05
v. 17	1,932	0,991	0	9,024E-06
v. 18	1,350	1,485	0	1,565E-05
v. 19 – TR ₁	6,292	28,350	0	0
v. 20 – TR ₂	6,292	28,350	0	0

TABUĽKA III
Parametre uzlov

Názov	Uzol č.	P (MW)	Q (MVar)	Ua (kV)	Ur (kV)
Uzol č.1	1	0	0	23	0
Uzol č.2	2	-2	-0,3		
Uzol č.3	3	-1,2	-0,15		
Uzol č.4	4	-1,6	-0,2		
Uzol č.5	5	0	0		
Uzol č.6	6	-1,2	-0,1		
Uzol č.7	7	-0,9	-0,1		
Uzol č.8	8	-1	-0,05		
Uzol č.9	9	-1,4	-0,25		
Uzol č.10	10	-0,1	-0,05		
Uzol č.11	11	0	0		
Uzol č.12	12	-1,3	-0,1		
Uzol č.13	13	-0,9	-0,15		
Uzol č.14	14	-1,5	-0,1		
Uzol č.15	15	0	0		
Uzol č.16	16	-2,5	-0,4		
Uzol č.17	17	-2,9	-0,4		
Uzol č.18	18	-3	-0,5		
Uzol č.19	19	1	0		
Uzol č.20	20	1	0		

TABUĽKA IV
Výsledky riešenia ustáleného chodu

Uzol č.	Napätie [kV]	Názov ved.	Prúd [kA]	Činné straty [MW]	Jalové straty [MVar]
1	23,000	v.1	0,209	0,196	0,215
2	22,362	v.2	0,031	0,006	0,003
3	22,248	v.3	0,125	0,056	0,062
4	22,056	v.4	0,083	0,024	0,027
5	21,855	v.5	0,108	0,067	0,046
6	21,468	v.6	0,091	0,050	0,034
7	21,127	v.7	0,028	0,005	0,003
8	21,020	v.8	0,039	0,008	0,004
9	21,007	v.9	0,015	0,001	0,001
10	21,519	v.10	0,026	0,003	0,002
11	21,580	v.11	0,009	0,000	0,000
12	21,503	v.12	0,066	0,020	0,010
13	21,322	v.13	0,041	0,007	0,004
14	21,216	v.14	0,092	0,031	0,021
15	21,712	v.15	0,067	0,013	0,014
16	21,847	v.16	0,134	0,067	0,074
17	22,195	v.17	0,080	0,037	0,019
18	21,908	v.18	0,290	0,340	0,374
19	22,102	v.19	0,026	0,013	0,058
20	21,830	v.20	0,026	0,013	0,059
Celkové straty				0,971	0,993

A. Výpočet ustáleného chodu siete

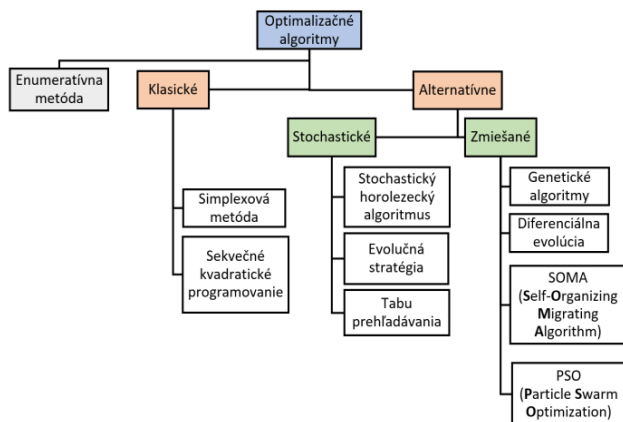
Na základe parametre uzlov a vetiev a na základe topológie siete (Obr. 4) bol vypočítaný ustálený stav siete. Ustálený stav siete je stav, v ktorom parametre v sieti, ktoré boli zadefinované, ostávajú nezmenené až do skončenia výpočtu ustáleného chodu siete. V reálnych podmienkach takýto stav - z dôvodu variabilnej záťaže, klimatických podmienok, porúch a podobne - neexistuje. Vo výpočtoch je však možné uvažovať so zjednodušením, kedy tieto stavy, tieto premenné, ostávajú nezmenené do skončenia výpočtu.

Výpočet ustáleného chodu pred aplikovaním optimalizačných algoritmov je dôležitý z dôvodu možnosti porovnávania výsledkov pred optimalizáciou a po optimalizácii. Výsledky riešenia ustáleného chodu v súčasnosti sa nachádzajú v TABUĽKA IV. Ako je možné vidieť, vo všetkých uzloch sú napätia v dovolených rozmedziach. Ani v jednom uzle nie je viac ako 25, 3 a nie menej ako 20, 7 kV. Najnižšie napätie je v 9. uzle - 21, 007 kV. Celkové činné straty v sieti majú hodnotu 971 kW a jalové straty dosahujú hodnotu 993 kVar.

V. OPTIMALIZÁCIA VÝKONOVÝCH STRÁT

Na optimalizáciu výkonových strát boli používané optimalizačné algoritmy. Úlohou optimalizačných metód/algoritmov je nájdenie minima resp. maxima v určitej populácii, kde pomocou vhodných kombinácií variantov dokáže nájsť najlepšiu a vyberá buď najmenšiu/najväčšiu hodnotu z nich. Tieto algoritmy sú zvyčajne používané tam, kde klasické výpočtové metódy nedosiahnu požadovanú presnosť alebo ich použitie pre daný typ úlohy nie je možné z hľadiska náročného zadefinovania potrebných výpočtových parametrov riešiť.

Pomocou optimalizačných metód je možné riešiť veľmi široké spektrum úloh, nie je obmedzené iba na jednu úlohu. Optimalizačné algoritmy je možné kategorizovať na základe niekoľkých kritérií. V článku optimalizačné algoritmy boli rozdelené na 4 podskupiny na základe toho, či sa jedná o klasické metódy alebo o nové algoritmy. Každá metóda má špecifické vlastnosti, na základe ktorých je treba zvoliť vhodnú metódu pri riešení konkrétnej úlohy tak, aby bola dosiahnutá požadovaná kvalita a presnosť. Spomínané rozdelenie algoritmov je možné vidieť na Obr. 5.



Obr. 5 Optimalizačné algoritmy [13]

1. *Enumeratívna metóda* – v tejto metóde dochádza k výpočtu všetkých kombinácií, ktoré sa vyskytujú pri riešení danej úlohy. Vzhľadom na to, že vypočíta všetky možné kombinácie variantov v problematike, výpočtový čas je veľmi dlhý, preto danú metódu odporúčajú iba pri riešení jednoduchých úloh v prípadoch, keď sa nepoužíva veľa parametrov. Pri zložitejších úlohách sa výpočtový čas na ukončenie metódy môže pohybovať od niekoľkých hodín až po niekoľko dní, komplikovanejšie príklady však vyžadujú ešte podstatne dlhší čas [13].

2. *Klasické optimalizačné metódy* – v týchto algoritmoch zvyčajne sú používané numerické, matematické modely, nakoľko v optimalizačných úlohách je potrebné zdefinovať rôzne vstupné obmedzenia a doplnujúce podmienky. Také predpoklady sú napr.:

- riešená úloha je lineárna alebo konvexná
- vyšetřovaný priestor je malý a súvislý
- účelová funkcia obsahuje iba jeden extrém, ...

Ak sú všetky tieto predpoklady jednoznačne definované, vtedy je možné nájsť iba jedno riešenie. Vo väčšine optimalizačných úloh takým riešením je nájdenie minima alebo maxima (v závislosti od zdefinovania úloh) funkcie [13].

3. *Alternatívne optimalizačné metódy* – nie všetky optimalizačné úlohy je možné riešiť iba pomocou predchádzajúcich metód. Tieto metódy nie sú vždy účinné a výpočtový čas je dlhý. Presne v takých prípadoch dochádza k zváženiu alternatívnych optimalizačných metód. Nové optimalizačné metódy sú schopné riešiť zložité optimalizačné úlohy relatívne rýchlo (výpočtový čas sa pohybuje len od niekoľkých minút do hodín). V prípade, ak sú parametre optimalizačných úloh veľmi dobre a presne zdefinované, je možné výpočet urobiť aj v reálnom čase. Aj v tomto prípade treba zvoliť čo najefektívnejšiu metódu na riešenie úlohy. Boli vyvinuté rôzne optimalizačné úlohy, zabezpečujúce riešiteľnosť všetkých vzniknutých problémov [13]. Alternatívne optimalizačné metódy je možné rozdeliť do dvoch skupín:

a) *Stochastické algoritmy* – tieto algoritmy vo svojich výpočtoch využívajú náhodné procesy. Pri výpočte sa náhodne

hľadajú hodnoty argumentov funkcie, ktorých výsledkom je vždy najlepšie riešenie, ktoré sa vyskytlo počas procesu. Najväčšou nevýhodou týchto algoritmov je, že vo svojich výpočtoch často používajú iba náhodné hľadanie, preto nie je možné zistiť, kedy bude daná úloha vyriešená a v ktorom kroku. Môže sa stať, že hneď v druhom kroku algoritmus náhodne vyberie najlepší variant, ale môže byť nájdená až v poslednom kroku. Preto najväčšia nevýhoda je neovplyvniteľný čas. Ďalšou nevýhodou je, že na základe odporúčani by sa stochastické algoritmy mali používať iba v prípadoch vyšetřovania malého priestoru. Súvisí to s používaním náhody v procesoch, a preto sú veľmi pomalé. Na základe odporúčani je vhodné stochastické algoritmy používať iba na hrubé odhady výsledkov. Pre nájdenia najlepších variantov bude používaná iná metóda [13].

b) *Zmiešané algoritmy* – tento algoritmus je kombináciou predošlých dvoch algoritmov, teda stochastického a klasického algoritmu. Každé predchádzajúce opísané metódy majú svoje výhody a nevýhody, kedy ich odporúčajú využívať. Optimálnou cestou je riešenie úloh prostredníctvom kombinácie algoritmov zabezpečujúcej dosiahnutie kvalitnejších výsledkov, akoby by to bolo pri využití len stochastickej alebo klasickej metódy.

Hlavnými výhodami používania zmiešaného algoritmu sú:

- robustnosť;
- rýchlosť;
- žiadne, resp. minimálne požiadavky na vstupné informácie;
- schopnosť vyriešiť viac úloh v priebehu jedného výpočtu;
- nie je potrebné analyticky popísať riešenú úlohu [13].

A. *Cieľová funkcia*

Pri riešení úloh je potrebné odvodiť cieľovú funkciu slúžiacu na opísanie cieľa riešenia danej úlohy. V tomto prípade je cieľom minimalizácia strát v sieti. Eliminácia strát znamená znižovanie finančných nákladov formou dodržiavania adekvátnych hodnôt.

Straty, ktoré vznikajú v sieti boli vyjadrované ako suma činných výkonov vo všetkých uzloch siete. Pri odvodení vzorcov boli používané predpoklady, že odberové uzly mali záporné znamienka a výrobné uzly majú kladné znamienka.

$$\Delta S = \sum_{i=1}^n \dot{S}_i = \sum_{i=1}^n (P_i - j * Q_i) \tag{1}$$

$$\Delta S = \sum_{i=1}^n [(U_{ai} - j * U_{ri}) * \sum_{j=1}^n (U_{aj} + j * U_{rj}) * (g_{ij} - j * b_{ij})] \tag{2}$$

Činné straty je možné vypočítať podľa vzorca (3) ako reálna časť strát zdanlivého výkonu. Druhou možnosťou výpočtu strát je použitie vzorca (4):

$$\Delta P = Re(\Delta S) \tag{3}$$

$$\Delta P = \sum_{i=1}^n P_i = \sum_{i=1}^n U_{ai} * \sum_{j=1}^n (U_{aj} * g_{ij} + U_{rj} * b_{ij}) \tag{4}$$

Obmedzenia použité pri výpočtoch:

- napätia v uzloch museli byť v rozmedzí $\pm 10\%$ oproti napätiu, ktoré bolo v bilančnom (v prvom) uzle. Napätie v bilančnom uzle malo hodnotu 23 kV, preto bolo potrebné dodržiavať napätie v rozmedzí 20,7 – 25,3 kV.
- hodnotu jalového výkonu v regulačných uzloch bolo potrebné dodržiavať v rozmedzí 0 – 1 MVar.

VI. RIEŠENIE OPTIMALIZAČNEJ ÚLOHY POMOCOU ENUMERATÍVNEJ METÓDY

Pri tejto metóde dochádza k výpočtu všetkých kombinácií, ktoré sa môžu vyskytnúť pri riešení danej úlohy. Enumeratívna metóda pri riešení problému má niekoľko výhod, ale aj nevýhod. Medzi výhody je možné zaradiť, že algoritmus pracuje s jednoduchými príkazmi, netreba definovať zložité funkcie, princíp výpočtu sa dá pomerne jednoducho pochopiť. Najväčšou nevýhodou tohto riešenia je výpočtový čas pohybujúci sa v časovom rozmedzí od niekoľkých minút až po niekoľko hodín.

V tomto prípade hodnotu jalového výkonu bolo možné ovplyvniť súčasne v dvoch miestach, konkrétne v 5. a 11. uzle. Výpočty ukázali, že hodnoty ideálneho jalového výkonu v prvom regulačnom uzle treba nastaviť na hodnotu 0,48 MVar a v druhom na 0,1 MVar. V prípade, že budú nastavené takéto hodnoty sa činné straty znížia na hodnotu 0,958 MW a zároveň budú vyššie napätia aj v uzloch. Výpočtový čas je 637,33 sekúnd.

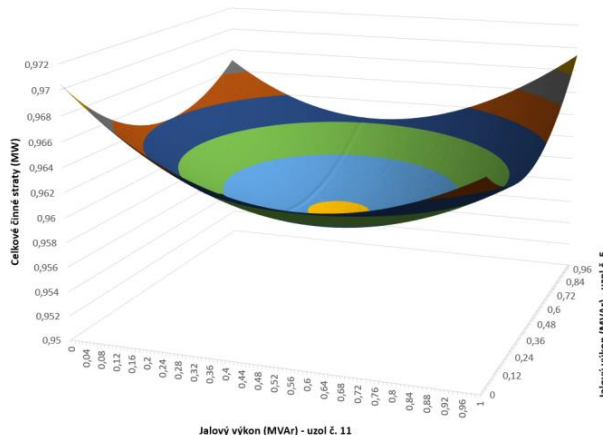
TABUĽKA V

Výsledky riešenia ustáleného chodu pomocou enumeratívnej metódy po optimalizačnom výpočte

Uzol č.	Napätie [kV]	Číslo ved.	Prúd [kA]	Činné straty [MW]	Jalové straty [MVar]
1	23,000	v.1	0,207	0,192	0,211
2	22,403	v.2	0,031	0,006	0,003
3	22,288	v.3	0,123	0,055	0,060
4	22,130	v.4	0,081	0,023	0,026
5	21,961	v.5	0,108	0,067	0,046
6	21,580	v.6	0,091	0,049	0,034
7	21,241	v.7	0,027	0,005	0,003
8	21,134	v.8	0,039	0,008	0,004
9	21,122	v.9	0,015	0,001	0,001
10	21,635	v.10	0,029	0,003	0,002
11	21,718	v.11	0,012	0,001	0,000
12	21,599	v.12	0,065	0,019	0,010
13	21,419	v.13	0,041	0,007	0,004
14	21,313	v.14	0,091	0,030	0,020
15	21,790	v.15	0,065	0,012	0,013
16	21,904	v.16	0,131	0,065	0,072
17	22,225	v.17	0,080	0,037	0,019

18	21,938	v.18	0,287	0,334	0,367
19	22,806	v.19	0,028	0,015	0,068
20	22,610	v.20	0,029	0,015	0,069
			Celkové straty		0,958 0,994
Výpočtový čas [s]:		637,33	Ideálny jalový výkon podľa výpočtu [MVar]		
Počet vypočítaných kombinácií		10 000	uzol č. 5		0,48
			uzol č. 11		0,51

Nižšie je uvedený funkčná závislosť funkcie činných výkonových strát od regulovaných premenných v 3-rozmernom priestore, ak sa uvažovalo s reguláciou jalového výkonu súčasne v dvoch rôznych miestach. V každom prípade je možné vidieť, že pomocou vhodných kombinácií jalového výkonu je možné dosiahnuť minimálne celkové činné straty v sieti. Minimum nastane vtedy, ak v 5. uzle bude nastavený 0,48 MVar výkon a v 11. uzle 0,51 MVar. Vtedy celkové činné straty budú mať hodnotu 0,958 MW.



VII. RIEŠENIE OPTIMALIZAČNEJ ÚLOHY POMOCOU SOMA ALGORITMU

V tomto prípade vo výpočtoch bol používaný SOMA (Samo-Organizujúci sa Migračný Algoritmus) algoritmus, ktorý patrí medzi optimalizačné algoritmy. SOMA algoritmus, tak ako všetky ostatné algoritmy, majú výhody aj nevýhody. Medzi výhody je možné zaradiť výpočtový čas. Program založený na báze SOMA algoritmu vypočíta dostatočne rýchlo výsledky a vhodným nastavením jeho parametrov je možné získavať kvalitné výsledky. Nastavené parametre, ktoré boli používané vo výpočtoch sú bližšie popísané v nasledujúcich podkapitolách.

Medzi nevýhody SOMA algoritmu je možné zaradiť pochopiteľnosť fungovania algoritmu a aj to, že nevhodným nastavením parametrov nemusí byť nájdené globálne, ale len lokálne minimum.

Nastavené parametre v algoritme SOMA

Pri používaní algoritmu SOMA v Matlab softvéri boli nastavené riadiace a ukončovacie parametre tak, aby SOMA pracoval na základe požiadaviek. Tieto parametre je možné vidieť v TABUĽKA VI.

TABUĽKA VI
SOMA parametre

Parameter	Typ	Nastavená hodnota
Dĺžka cesty	Riadiaci	1,20
Krok	Riadiaci	0,05 – 0,10
Veľkosť populácie	Riadiaci	10
Migrácia	Ukončovaci	10
Chyba	Ukončovaci	10^{-6}

V tomto prípade na ovplyvnenie strát v sieti bolo možné používať súčasne dva uzly. Optimálny jalový výkon v týchto dvoch miestach je potrebné nastaviť na hodnotu, ktorá je v rozpätí 0,49 – 0,496 MVar, čím sa dosiahne minimalizácia celkových činných strát v sieti. Pomocou hodnôt 0,49 MVar je možné straty znížiť až na hodnotu 0,958 MW. Výpočtový čas je v tomto prípade 60,98 sekúnd. Výsledky riešenia ustáleného chodu pomocou SOMA algoritmu sú zapísané do TABUĽKA VII.

TABUĽKA VII
Výsledky riešenia ustáleného chodu pomocou SOMA algoritmu

Uzol č.	Napätie [kV]	Názov ved.	Prúd [kA]	Činné straty [MW]	Jalové straty [MVar]
1	23,000	v.1	0,207	0,192	0,211
2	22,403	v.2	0,031	0,006	0,003
3	22,288	v.3	0,123	0,055	0,060
4	22,129	v.4	0,081	0,023	0,026
5	21,960	v.5	0,108	0,067	0,046
6	21,579	v.6	0,091	0,049	0,034
7	21,240	v.7	0,027	0,005	0,003
8	21,133	v.8	0,039	0,008	0,004
9	21,120	v.9	0,015	0,001	0,001
10	21,632	v.10	0,028	0,003	0,002
11	21,715	v.11	0,012	0,001	0,000
12	21,597	v.12	0,065	0,019	0,010
13	21,417	v.13	0,041	0,007	0,004
14	21,311	v.14	0,091	0,030	0,020
15	21,789	v.15	0,065	0,012	0,013
16	21,903	v.16	0,131	0,065	0,072
17	22,224	v.17	0,080	0,037	0,019
18	21,937	v.18	0,287	0,334	0,367
19	22,816	v.19	0,028	0,015	0,067
20	22,579	v.20	0,028	0,015	0,069
Celkové straty				0,958	0,994
Ideálny jalový výkon podľa výpočtu [MVar]:			0,49-0,496		
Výpočtový čas [s]:			60,98		

VIII. POROVNANIE DOSIAHNUTÝCH VÝSLEDKOV

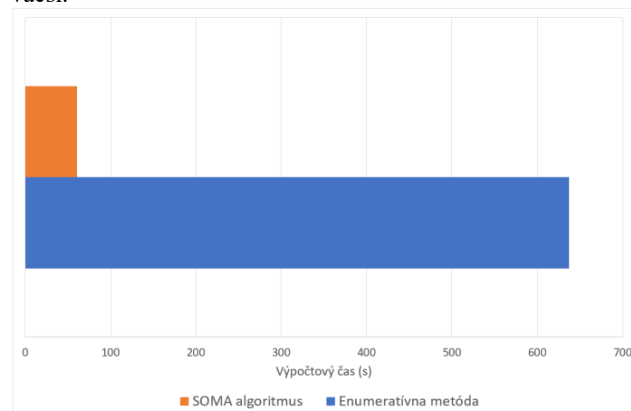
Porovnávanie a vyhodnotenie dosiahnutých výsledkov je podkladom pre záverečné určenie metódy, ktorá bola najvhodnejšou na získanie kvalitnejších výstupov. Porovnanie enumeratívnej metódy a SOMA algoritmu je možné posudzovať na základe rôznorodých kritérií a rozličných hľadísk. V článku na vyhodnotenie boli vybrané nasledujúce hľadiská:

- podľa výpočtového času,
- podľa hodnoty optimálneho jalového výkonu
- podľa uzlových napätí

A. Porovnanie podľa výpočtového času

Výpočtový čas je dôležitým parametrom, ktorý určuje, či navrhnuté algoritmy je alebo nie je možné používať aj v reálnom živote a v reálnych podmienkach. Ako príklad je možné uviesť situáciu, kedy dispečer siete nemôže čakať veľmi dlhú dobu na to, aby zistil, aký výkon má nastaviť. Musí rozhodovať veľmi rýchlo, preto treba používať veľmi kvalitné programy.

Algoritmus, ktorý bol navrhnutý pomocou enumeratívnej metódy má väčšie výpočtové časy v porovnaní s algoritmom SOMA. Ako je možné vidieť na Obr. 6, trvanie SOMA algoritmu bolo cca. 70 sekúnd, ale výpočtový čas enumeratívnej metódy bol 637 sekúnd, čo je skoro 10-násobne väčší.

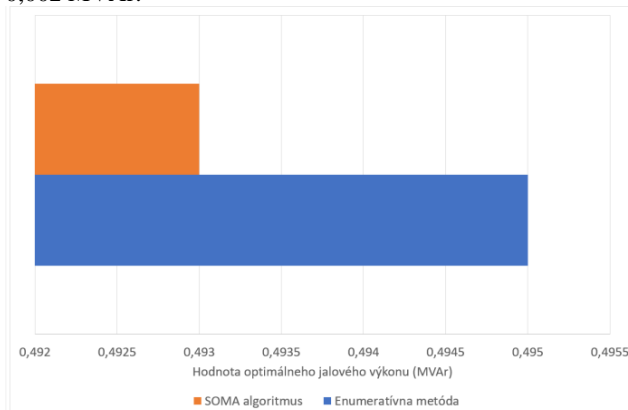


Obr. 6 Porovnanie podľa výpočtového času

B. Porovnanie podľa hodnoty optimálneho jalového výkonu

Porovnávanie na základe nájdenia ideálneho jalového výkonu bolo problematickejšie, nakoľko využitím SOMA algoritmu nie je daná iba jedna hodnota, ale je daný rozsah, v ktorom sa optimálny výkon nachádza medzi hodnotami. Ak je potrebné získať iba jednu hodnotu, vtedy sa vyžaduje parameter chyby nastaviť na 10^{-10} – 10^{-15} hodnotu, pričom však dochádza k nárastu výpočtového času. Na získanie presných výsledkov v dostatočne krátkom čase vo výpočtoch bola používaná hodnota 10^{-6} . Aby bolo možné urobiť porovnanie, z rozsahu SOMA algoritmu bola používaná stredná hodnota. V prípade používania enumeratívnej metódy program vypočítal 2 hodnoty v regulačných uzloch, preto na porovnanie bola používaná priemerná hodnota.

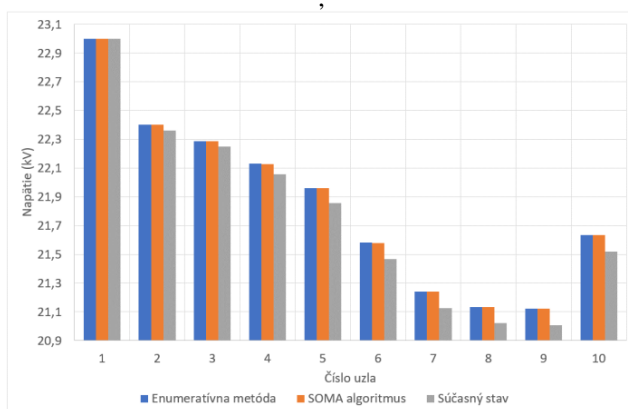
Získané porovnanie je možné vidieť na Obr. 7. Pruhový graf znázorňuje výsledky, ktoré sú pomerne totožné aj napriek tomu, že na nájdenie optimálneho výkonu boli používané rôzne optimalizačné algoritmy. V grafe je modrou farbou označený výsledok enumeratívnej metódy a oranžovou farbou výsledok získané pomocou SOMA algoritmu. Rozdiel je iba 0,002 MVA_r.



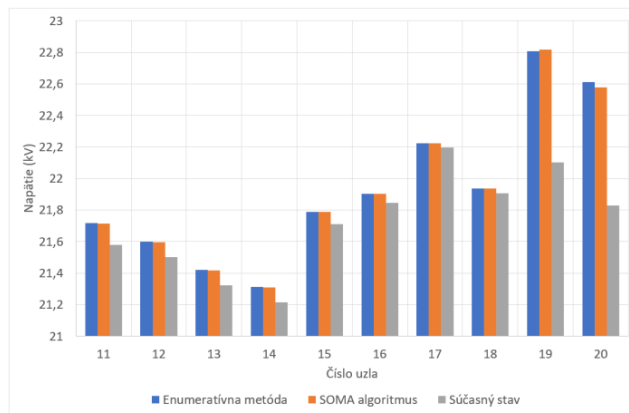
Obr. 7 Porovnanie podľa hodnoty optimálneho jalového výkonu

C. Porovnanie podľa uzlových napätí

Na Obr. 8 sú znázornené uzlové napätia od prvého po desiaty uzol a na Obr. 9 napätie od jedenásteho po dvadsiaty uzol. Zreteľne je možné vidieť – vzhľadom na skoro rovnaké hodnoty optimálneho jalového výkonu v regulačných uzloch – že aj výsledky uzlových napätí sú skoro rovnaké. Výber metódy na vyriešenie problému nemá vplyv na výsledok. Aktuálne napätie je v grafe vyznačené sivou farbou. Stĺpcový graf dostatočne vhodne prezentuje, že pomocou regulačných uzlov a pomocou jalového výkonu je možné ovplyvniť napätie v uzloch a takmer v každom uzle, okrem prvého, bilančného uzla, bolo dosiahnuté väčšie napätie.



Obr. 8 Porovnanie uzlových napätí (1.- 10. uzol) s navrhovanými algoritmi



Obr. 9 Porovnanie uzlových napätí (11.- 21. uzol) s navrhovanými algoritmi

ZÁVER

Hlavným cieľom tohto príspevku bolo predstaviť optimalizačné algoritmy, ktoré je možné používať pri riešení úloh vyskytujúcich sa pri prevádzke siete. Pomocou optimalizačných algoritmov je možné nájsť optimálnu hodnotu jalových výkonov v regulačných uzloch, s ktorými je možné ovplyvniť výkonové straty v sieti. Ako bol ukázaný aj v príspevku, aj enumeratívna metóda aj SOMA algoritmus má výhody a nevýhody. Vždy treba vybrať tú metódu, ktorá je najideálnejšia pre riešenie problematiky. Výsledky ukázali, že v tomto prípade najideálnejšie je používať program na báze SOMA algoritmu, lebo výsledky ideálneho jalového výkonu sú skoro totožné (rozdiel je iba 0,003 MVA_r), ale výpočtový čas SOMA algoritmu je 10-násobne nižší.

POĎAKOVANIE

Táto práca bola podporovaná Agentúrou na podporu výskumu a vývoja na základe zmluvy č. APVV-19-0576

LITERATÚRA

- [1] Electricity explained, How electricity is delivered to consumers. Electricity is delivered to consumers through a complex network. [online]. [cit. 2021-07-30]. Dostupné na internete: <https://www.eia.gov/energyexplained/electricity/delivery-to-consumers.php>
- [2] Paks II nélkül a világ – Mi kizámoltuk, és közzé is tesszük. [online]. [cit. 2021-07-30]. Dostupné na internete: <https://regiblogok.atlatszo.hu/2015/01/20/paks-ii-nelkul-a-vilag-kizsamoltuk-es-kozze-is-tesszuk/>
- [3] Internet of things, Mikko Hyppönen: Smart devices are "IT asbestos" [online]. [cit. 2021-07-30]. Dostupné na internete: <https://www.verdict.co.uk/mikko-hypponen-smart-devices-it-asbestos/>
- [4] European SmartGrids Technology Platform, Vision and Strategy for Europe's Electricity Networks of the future. Belgium, 2006, str. 38. ISBN: 92-79-01414-5.
- [5] What is the Smart Grid? [online]. [cit. 2021-07-30]. Dostupné na internete: https://www.smartgrid.gov/the_smart_grid/smart_grid.html
- [6] H. Farhangi, "The path of the smart grid," in IEEE Power and Energy Magazine, Ročník 8, číslo 1, str. 18 - 28, January-February 2010. doi: 10.1109/MPE.2009.934876. Dostupné na internete: <https://ieeexplore.ieee.org/document/5357331>
- [7] L. I. Dulá, M. Abrudean, D. Bicá, „Optimal Location of a Distributed Generator for Power Losses Improvement“ in Procedia Technology. Elsevier Ltd. 2016, Volume 22, pp. 734-739. Dostupné na internete:

- <<https://www.sciencedirect.com/science/article/pii/S2212017316000335>>
- [8] Dulău LI, Abrudean M and Bică D., „Automation of a Distributed Generation System“, 49th Universities' Power Engineering Conference (UPEC), IEEE Xplore, pp. 1-5, 2014. ISBN: 978-1-4799-6557-1. doi: 10.1109/UPEC.2014.6934734.
- [9] R. Rekowski, „Problematyka doboru współczynnika czułości algorytmu rekonfiguracji sieci dystrybucyjnej SN“ in XVIII Konferencja Aktualne Problemy w Elektroenergetyce APE'2017. Jastrzębia Góra, Polska, 2017, No. 53, pp. 103-106. Dostupné na internete: <<http://yadda.icm.edu.pl/baztech/element/bwmeta1.element/baztech-0dea7f44-a631->>>
- [10] J. E. Mendoza, M. E. Lopez, C. A. C. Coello and E. A. Lopez, "Microgenetic multiobjective reconfiguration algorithm considering power losses and reliability indices for medium voltage distribution network," in IET Generation, Transmission & Distribution, Vol. 3, no. 9, pp. 825-840, September 2009. ISSN 1751-8695. Dostupné na internete: <https://www.researchgate.net/publication/224588307_Microgenetic_multiobjective_reconfiguration_algorithm_considering_power_losses_and_reliability_indices_for_medium_voltage_distribution_network>
- [11] G. Vaskantiras, S. You: „Value Assessment of Distribution Network Reconfiguration: A Danish Case Study“ in Energy Procedia. Elsevier Ltd. 2016, Volume 100, pp. 336-341. Dostupné na internete: <<https://www.sciencedirect.com/science/article/abs/pii/S187661021631147X>>
- [12] Mohd Dahalan, Wardiah & Mokhlis, Hazlie & Halim, Abdul & Abu Bakar, Ab Halim & Jamian, J. J. „The Simultaneous Application of Optimum Network Reconfiguration and Distributed Generation Sizing Using PSO for Power Loss Reduction“. Dostupné na internete: https://www.researchgate.net/publication/260109311_The_Simultaneous_Application_of_Optimum_Network_Reconfiguration_and_Distributed_Generation_Sizing_Using_PSO_for_Power_Loss_Reduction
- [13] I. Zelinka, „Umělá inteligence v problémech globální optimalizace“. BEN – technická literatura, Praha 2002. ISBN 80-7300-069-5.

Review Of Feature Extraction Techniques From Still-3D Images For Facial Emotion Recognition

Fouzia Adjailia

Dep. Cybernetics and Artificial Intelligence,
Faculty of Electrical Engineering and Informatics.
Technical University of Košice
Košice, Slovak Republic
fouzia.adjailia@tuke.sk

Peter Sinčák

Dep. Cybernetics and Artificial Intelligence
Faculty of Electrical Engineering and Informatics.
Technical University of Košice
Košice, Slovak Republic
peter.sincak@tuke.sk

Abstract—Computer vision and image processing are receiving considerable attention among researchers and industry for the role it places as an effective variable to make machines adapt daily tasks and be adapted by humans. Facial emotion recognition is key point to make machines do that. It is important to make machines understand and perceive the emotions, and for that task, features must be extracted. Representing the human faces in three-dimensional representation showed a promising and better performance than two-dimensional representation. For a better operations on an image, features must be extracted. Features are an important parameter to enhance the results for the classification purposes. This paper serves as a review of feature extraction techniques for 3D-still images, we present the basic notions of facial emotion recognition, representation, modeling and feature extractions as well as an investigation on the most commonly used 3D image datasets for facial emotion recognition (BU-3DFE, D3DFACS, Bosphorus 3D and ND-2006, IMPA-FACE3D), alongside with that, we provide a comprehensive review of the current approaches for the feature extractions of 3D models like 2D texture descriptors and 3D Geometric descriptor. We identify the challenges concerning these techniques using well discussed and evaluated comparison based on the performance of those techniques.

I. INTRODUCTION

Vision is an important non-verbal communication channel, as a result, computer vision has attracted much attention from researchers. Research on the human face is very popular in image analysis and computer vision because the applications of automatic facial information analysis are numerous and span several fields, ranging from Human-Computer Interaction (emotion recognition) to law enforcement (face recognition).

To analyze a big amount of data that requires a huge amount of computing power and time, Feature extraction leads computers to better interpretations and to extract efficient, effective and non-redundant information from images. Feature extraction plays a critical step in computer vision.

Most research for face analysis utilized 2D static images or 2D dynamic videos. Recently, 3D static or dynamic data has received increasing attention due to its explicit representation of geometric structures and their inherent capacity to handle facial pose and illumination variations. Similar to the 2D images, 3D modality can also be represented in static space and dynamic space.

Facial expressions are those significant movements of the facial muscles that create a visual expression for other people which conveys emotion. The emotions displayed on the person's face are a result of components that control the intensity of the expressions. Recognizing an expression from an image can be done through three major steps from (1) Pre-processing (2) Feature extraction (3) Classification as shown in figure 1



Fig. 1. Typical facial emotion recognition system.

The first phase Facial Pre-processing consists of face detection, alignment, normalization or augmentation. Once the image is pre-processed some features are extracted from the image and for that task, many methods are dominated, global, local and temporal. Which will be well detailed later. The global feature extraction approach tries to extract features by encoding the whole face. However, the local feature extraction approach measures local human face areas such as eyes, eyebrows, nose, mouth...etc. The temporal approach also is a third approach that captures and extracts dynamically facial expression sequences. Last, the classification phase consists of producing a classification of the extracted feature using pre-trained classifiers like SVM, Adaboost and random forest.

To elaborate on the process of image processing and feature extraction, we must distinguish the types of features that can be distracted from images:

- Color features are pixel property commonly used the feature to emphasizes the descriptive aspects. It is robust, effective, requires simple implementation and computation as well as low storage. A number of important color features have been proposed like color histogram (1), color moments (2), color coherence vector (3) and color correlogram (4).

Color moments is defined as follows:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N \sum_{ij} R_{ji}^i \quad (1)$$

$$\sigma_i = \frac{1}{N} \sum_{j=1}^N \sum_{ij} R_{ji}^i - \mu_i^2 \quad (2)$$

$$\gamma_i = \frac{1}{N} \sum_{j=1}^N \sum_{ij} R_{ji}^i - \mu_i^3 \quad (3)$$

where R_{ji} is the color value of the i -th color component of the j -th image pixel and N is the total number of pixels in the image. $\mu_i, \sigma_i, \gamma_i (i = 1, 2, 3)$ denote the mean, standard deviation and skewness of each channel of an image respectively.

- Texture features, this feature is extracted by the computation of the pixel statistics or finding the local pixel structures in the original image domain. we can distinguish spatial texture feature extraction methods and spectral texture feature extraction methods. the most famous method used is the Gabor filter that works on obtaining the center frequency and orientation parameter (5), A two-dimensional Gabor function $g(x,y)$ can be defined as follows:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi jW_x \right] \quad (4)$$

where σ_x and σ_y are the scaling parameters of the filter (the standard deviations of the Gaussian envelopes), W is the center frequency, and θ determines the orientation of the filter.

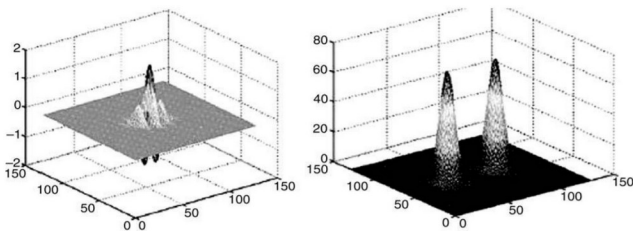


Fig. 2. Gabor function in the spatial domain

- Shapes features are some set of numbers that are produced to encode simple geometrical forms such as straight lines in different directions, and for that edge matching is used. Shape feature extraction techniques can be divided into contour based and region based methods (6), In (7) a well detailed A survey of shape feature extraction techniques was presented. according to (7) Shape descriptors should meet a specific requirements

such as a) be complete as possible b) should be stored compactly c) distance between descriptors should be simple.

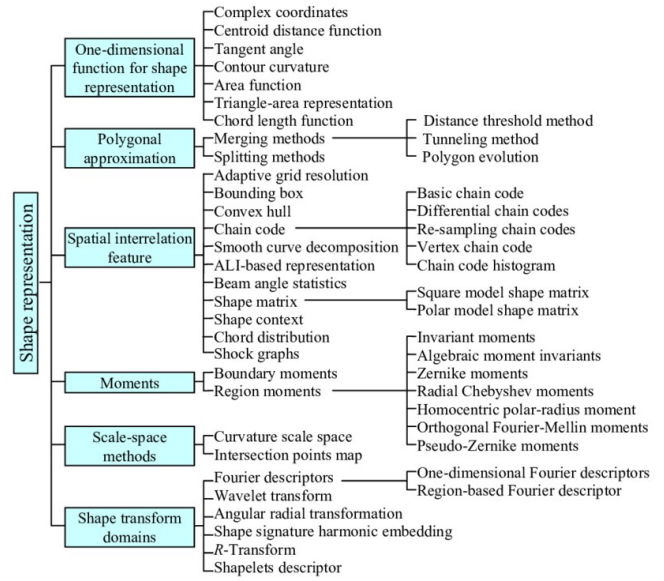


Fig. 3. An overview of shape description technique/2(7)

II. CONTRIBUTIONS OF THIS REVIEW

The main contributions of this review are as follows:

- Help new researchers understand basic notions of facial emotion recognition and its key components including feature extraction.
- Provide a new assimilation after compiling the literature and related work.
- Establish the boundaries of feature extraction techniques while also identifying core constructs and their relationships and key aspects compared between those approaches.
- Introduce the standard 3D datasets for facial emotion recognition databases along with their characteristics.

III. METHODOLOGY OF COMPARISON

We can distinguish two types of feature extraction techniques: 2D texture descriptors and 3D geometric descriptors:

Textured images are commonly used as an input source and are based on having visual entities. where they provide a representation in details about the entity, in which noise and other information unnecessary to the subject can be found. Hand-crafted texture descriptors can highlight certain characteristics of the image that is best suitable for its application.(8).

The Geometric descriptors are useful tool for consistency of information for facial emotion recognition where it is based on a 3D point cloud data.

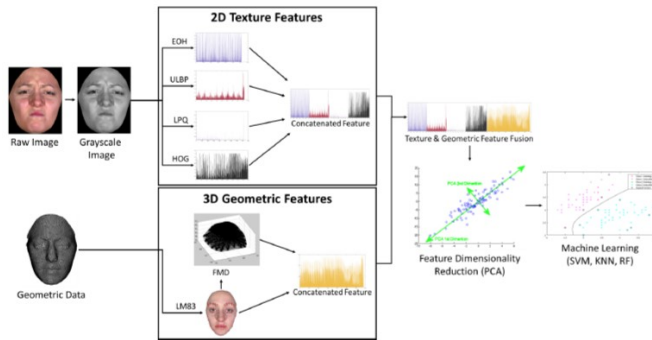


Fig. 4. Framework for Fusion based 3D/2D Facial Expression Recognition using static images. There are two initial branches that start with the texture data and the geometric data. Each of those branches has their respective features extracted, which are fused together using concatenation. Then the two branches are also fused and reduced in dimensionality before being used for machine learning(8)

IV. FACIAL EMOTION RECOGNITION

In order to make machine understand our emotions, researchers were able to propose emotional models, the figure demonstrates well used emotional models.

Categorical approach is a popular type of models that represents the emotions in a set of basic emotions [7].

However, the dimensional approach suggests an inter-connectivity between the emotions and it can be presented by two dimensions, The most widely used model is Russell with the valence and activation (or valence-arousal) dimensions represented by Russell's [8]

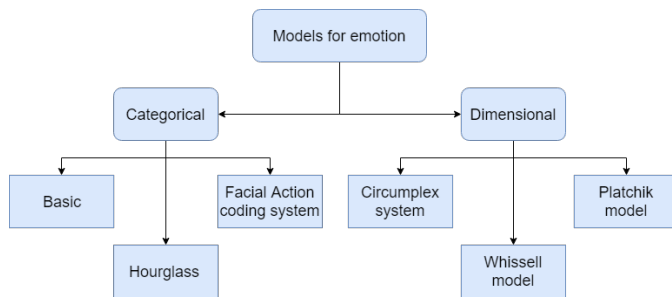


Fig. 5. Emotion Models

V. 3D DATASETS FOR FACIAL EMOTION RECOGNITION

A. Binghamton University 3D Facial Expression

The BU-3DFE (9) is a public database created by a research group at Binghamton university in 2006, it was designed for providing a rich 3D spectrum of emotional expressions available for the research community. the dataset includes both prototypical 3D facial expression shapes and 2D facial textures of 2500 3D scans from 100 subjects, 25 are females and 44 are males. with various each subject performing six expressions at 4 intensity levels.

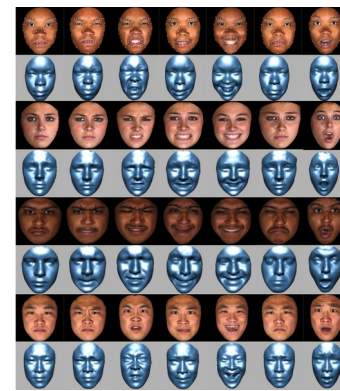


Fig. 6. Images from BU-3DFE dataset

The expressions included are Angry, Disgust, Fear, Happy, Sad and Surprise, along with a single sample of the subject's neutral face, and the intensity is in a range from mild to strong. The dataset is valuable resource for understanding the 3D behavior of the facial expression because it come from a variety of ethnic backgrounds.

B. Bosphorus 3D

The database (10) consists of 105 subjects consists of 60 men and 45 women aged between 25 and 35 in various poses, expressions and occlusion conditions. 18 men had beard/moustache and 15 other subjects had short facial hair, and most of the subjects are Caucasian. the total number of scans is 4652. Each scan has been manually labelled for 24 facial landmark points as shown in figure 7 such as nose tip, inner eye corners, etc.

we can distinguish two sets of emotion in this dataset: action units (AUs) that were grouped into 20 lower face AUs, five upper face AUs and three AU combinations. however the second set consists of a specific emotional expressions. These are: happiness, surprise, fear, sadness, anger and disgust 7.

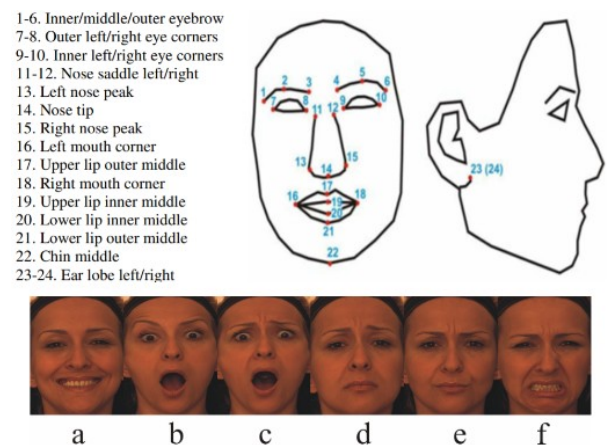


Fig. 7. Bosphorus 3D

C. Dynamic 3D Facial Action Coding System

In 2011 a dynamic 3D FACS data set for facial expression research was presented (11), the dataset contains 519 AU sequences where 10 subjects (6 females and 4 males) aged between 23 to 41 years, each performed between 19 and 97 different AUs both individually and in combination. The dataset required 4 expert FACS coders and 6 FACS-untrained participants figure 8

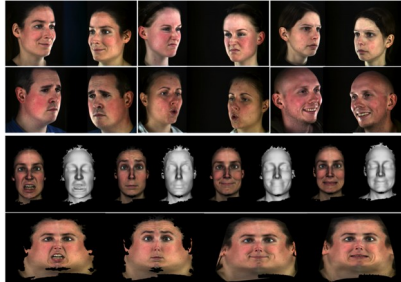


Fig. 8. Dynamic 3D Facial Action Coding System

D. Notre Dame 3D face dataset

The ND-2006 data set is the largest 3D face data set available that contains a 888 distinct subjects, where 63 images are provided for each subject (12). A total images is 13,450 with six different types of expressions (Neutral, Happiness, Sadness, Surprise, Disgust, and Other) are available in the ND-2006 data set as shown in FigV-D A total of. A table containing the number of images based on expression type is shown in ??.

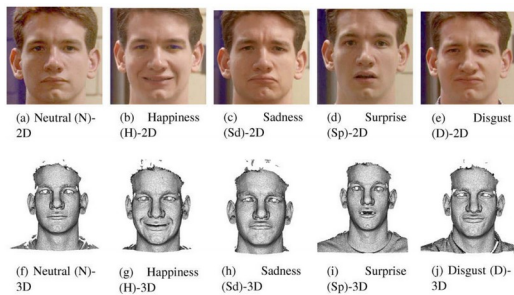


Fig. 9. Notre Dame 3D face dataset

E. IMPA-FACE3D

IMPA-FACE3D (13) dataset was created in 2008 for the purpose of analysis and synthesis of faces and expressions. This dataset contains both the geometric information with color and texture, where 38 subjects participated. The emotions in this dataset are neutral face (i.e. face with the front camera position -0 degrees and without facial expression), beside the six basic emotions: happiness, sadness, surprise, anger, disgust and fear 10

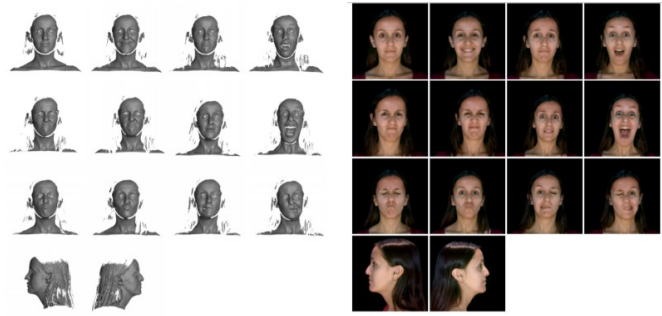


Fig. 10. IMPA-FACE3D

IMPA-FACE3D provides three-dimensional data from 3-D scanners. in a format for storing graphical objects that are described as a collection of polygons.

VI. FEATURE EXTRACTION TECHNIQUES

1) 2D Texture Descriptors Based Methods:

- **Local Binary Patterns:** Local Binary Pattern was introduced by (14). It is popular non-parametric descriptor technique that calculates each pixel which threshold the surrounding pixel in 3*3 blocks with the center value and considering the results as a binary number (15), figure, 11.

LBP is a powerful descriptor and showed a promising results using 2D datasets like JAFFE (16) Cohn-Kanade (17).

There are different kinds of patterns that can be detected using local binary patterns are Spot, Spot/Flat, Line end, Edge, Corner. While LBP features have been successfully used in 2D facial expression recognition, 3D feature extraction using LLBP was impossible but (18) made it possible to extract facial features using 2D methods proposes a novel facial feature, polytypic multi-block local binary patterns (P-MLBP). We can also distinguish types of LBP, figure 12, like Uniform LBP, Extended LBP (19)

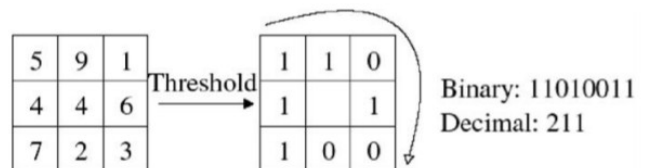


Fig. 11. The basic LBP operator (20)



Fig. 12. Visualisation of the ULBP operator applied to a Facial image demonstrating an intense Angry expression, where a histogram will be produced on each of the 64 windows to make the ULBP feature

- **Local Phase Quantisation:** Local Phase Quantisation is a robust technique proposed by Ojansivo and Heikkila (21) that works in the frequency domain to obtain the Fourier frequency and to image blurring due to the use of phase information. another technique like LBP but works in. and it is to analyse the differences between expressions that accrue in the frequently domain. in they used LPQ followed by two methods of feature extraction PCA and EFM before classification using SVM and they got results of recognition rate of 98.18% using CASIA 3D dataset (22).

2) *Histogram Based Methods:*

- **Histogram of Oriented Gradients:** Histogram of Oriented Gradients (23) is a common feature to describe image local texture in the field of computer vision by counting occurrences of gradient orientation in localized portions of an image. in (24) Shu Ana and Qiuqi Ruan used HOG for facial emotion recognition with SVM using BU-3DFE dataset (9). they obtained an accuracy of 80%. In (25) Histograms of Oriented Gradients was used along with Multiclass-SVM classification algorithm. Using the protocol proposed by Gong et al. [1] on the BU-3DFE dataset, and it showed a promising results while staying fully automatic.

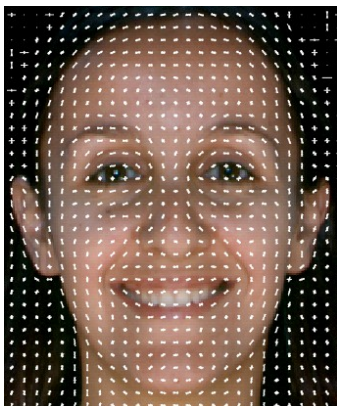


Fig. 13. Visualisation of the HOG descriptor applied on a facial image displaying a happy expression

3D Voxel HOG (3D VHOG) is based on the original Histogram of Oriented Gradients by (23). It extends the method by using voxels over pixels and expands the original histograms into 2 dimensions. The multiple steps for 3D gradient orientation descriptor calculation are shown in Figure 14.

The cell grid is split into the service area of the STIP. Likewise, each cell is divided into a grid of blocks. An integral video-based rapid measurement then calculates the 3D mean gradient in each row, the direction of which is quantified using a regular polyhedron to form a block histogram. After that a cell histogram is generated by summing up all the block histograms within that cell. Finally, within the support zone of the STIP, the HOG3D descriptor is a concatenation of all cell histograms.

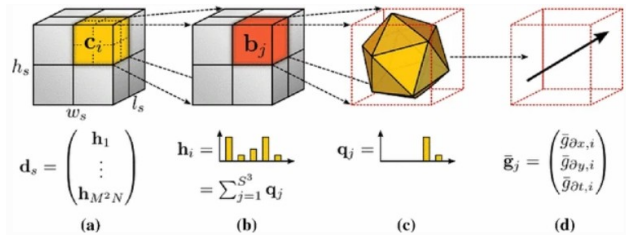


Fig. 14. Overview of HOG3D descriptor computation. (a) The support region of an STIP is divided into a grid of cells and the final descriptor is a concatenation of all cell histograms. (b) A cell is divided into a grid of blocks and a cell histogram is the sum of all block histograms within the cell. (c) A block histogram is obtained by regular polyhedron-based orientation quantization of mean block gradient. (d) Mean gradient in a block (26).

- **Edge Oriented Histogram:** Used mainly to counts occurrences of gradient orientations in localized (27), it is a robust way that focus mainly on edges to highlight the facial parts that have changed from the natural and initial position.

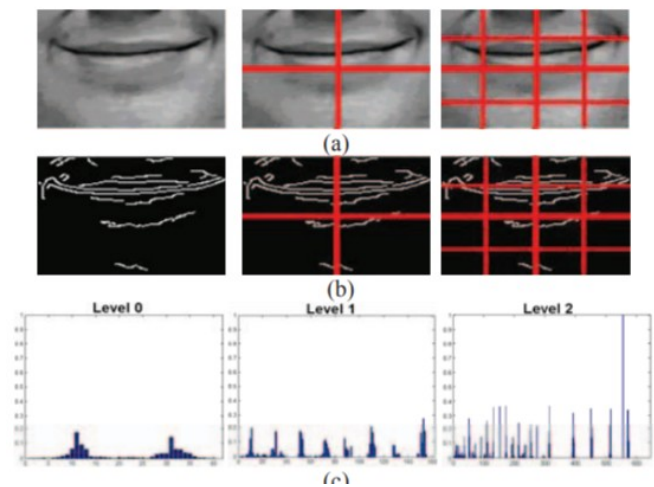


Fig. 15. Extraction of PHOG features from a smile image sample. (a) Grids at three pyramid resolution in the origin image, (b) edge contours are extracted using an edge detector, and (c) concatenation of all the HOG vectors in three pyramid resolutions to obtain the PHOG features of a sub-image (28).

3) 3D Geometric Descriptors Based Methods:

- **3D Facial Landmarks:** facial landmarks provide useful content information from images, for facial expression recognition in a way that points landmarks in the eyebrows, eyes, nose, mouth and around the face. Bosphorus database contains 22 landmarks.

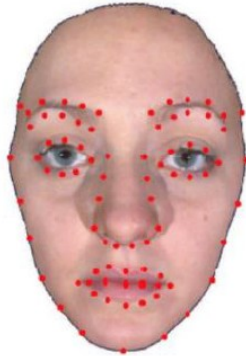


Fig. 16. 83 Facial Landmarks annotated on a BU-3DFE sample cropped face (8).

- **3D Facial Mesh Distances:** 3D Facial Mesh Distances from Geometric Data is a technique used to capture the variation of facial expressions from each other, a way to do this is to calculate the distance between the landmarks for each expression. A metric distance such as 3D Euclidean as shown in equation 5 can provide this information:

$$\delta_{i,j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2} \quad (5)$$

where $\delta_{i,j}$ the distance value between 3D points i and j , x, y, z the individual coordinates for each landmark i and j .

VII. CONCLUSION

This paper dedicated to outlining the most common 3D datasets for facial emotion recognition and a review of the techniques used for feature extraction from 3D images, we started by giving basic notions about facial emotion recognition and most commonly used 3D images datasets like BU-3DFE, D3DFACS, Bosphorus3D, ND-2006 and IMPA-FACE3D. We presented a comprehensive feature extraction techniques using the textured and geometric modalities in order to give a mathematical representation for the face.

REFERENCES

- [1] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern recognition*, vol. 29, no. 8, pp. 1233–1244, 1996.
- [2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic *et al.*, "Query by image and video content: The qbic system," *computer*, vol. 28, no. 9, pp. 23–32, 1995.
- [3] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," in *Proceedings Third IEEE Workshop on Applications of Computer Vision. WACV'96*. IEEE, 1996, pp. 96–102.
- [4] J. Huang, S. R. Kumar, M. Mitra, and W.-J. Zhu, "Image indexing using color correlograms," Jun. 12 2001, uS Patent 6,246,790.
- [5] B. S. Manjunath and W.-Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 8, pp. 837–842, 1996.
- [6] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, no. 1, pp. 1–19, 2004.
- [7] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A survey of shape feature extraction techniques," *Pattern recognition*, vol. 15, no. 7, pp. 43–90, 2008.
- [8] A. Jan, "Deep learning based facial expression recognition and its applications," Ph.D. dissertation, Brunel University London, 2017.
- [9] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3d facial expression database for facial behavior research," in *7th international conference on automatic face and gesture recognition (FGR06)*. IEEE, 2006, pp. 211–216.
- [10] A. Savran, N. Alyüz, H. Dibeklioglu, O. C. eliktutan, B. Gökberk, B. Sankur, and L. Akarun, "Bosphorus database for 3d face analysis," in *European workshop on biometrics and identity management*. Springer, 2008, pp. 47–56.
- [11] D. Cosker, E. Krumhuber, and A. Hilton, "A faces valid 3d dynamic action unit database with applications to 3d dynamic morphable facial modeling," in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 2296–2303.
- [12] T. C. Faltemier, K. W. Bowyer, and P. J. Flynn, "Using multi-instance enrollment to improve performance of 3d face recognition," *Computer Vision and Image Understanding*, vol. 112, no. 2, pp. 114–125, 2008.
- [13] "IMPA-FACE3D kernel description," <http://app.visgrafimpa.br/database/faces>, accessed: 2012-02-06.
- [14] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.
- [15] Y. Ma, "Number local binary pattern: An extended local binary pattern," in *Wavelet Analysis and Pattern Recognition (ICWAPR), 2011 International Conference on*. IEEE, 2011, pp. 272–275.
- [16] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," in *Proceedings Third IEEE international conference on automatic face and gesture recognition*. IEEE, 1998, pp. 200–205.

- [17] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 2010, pp. 94–101.
- [18] X. Li, Q. Ruan, Y. Jin, G. An, and R. Zhao, "Fully automatic 3d facial expression recognition using polytypic multi-block local binary patterns," *Signal Processing*, vol. 108, pp. 297–308, 2015.
- [19] H. Zhou, R. Wang, and C. Wang, "A novel extended local-binary-pattern operator for texture analysis," *Information Sciences*, vol. 178, no. 22, pp. 4314–4325, 2008.
- [20] A. Timo, "Face recognition with local binary patterns," in *Euro. Conf. on Computer Vision*, 2004.
- [21] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *International conference on image and signal processing*. Springer, 2008, pp. 236–243.
- [22] S. Zhang, A. Liu, J. Wan, Y. Liang, G. Guo, S. Escalera, H. J. Escalante, and S. Z. Li, "Casia-surf: A large-scale multi-modal benchmark for face anti-spoofing," *arXiv preprint arXiv:1908.10654*, 2019.
- [23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. IEEE, 2005, pp. 886–893.
- [24] S. An and Q. Ruan, "3d facial expression recognition algorithm using local threshold binary pattern and histogram of oriented gradient," in *2016 IEEE 13th International Conference on Signal Processing (ICSP)*. IEEE, 2016, pp. 265–270.
- [25] P. Lemaire, M. Ardabilian, L. Chen, and M. Daoudi, "Fully automatic 3d facial expression recognition using differential mean curvature maps and histograms of oriented gradients," in *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. IEEE, 2013, pp. 1–7.
- [26] A. Klaser, M. Marszałek, and C. Schmid, "A spatio-temporal descriptor based on 3d-gradients," in *BMVC 2008-19th British Machine Vision Conference*. British Machine Vision Association, 2008, pp. 275–1.
- [27] P. Carcagnì, M. Del Coco, M. Leo, and C. Distanto, "Facial expression recognition and histograms of oriented gradients: a comprehensive study," *SpringerPlus*, vol. 4, no. 1, p. 645, 2015.
- [28] Y. Bai, L. Guo, L. Jin, and Q. Huang, "A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 3305–3308.

Reinforcement Learning-Based Study Scheduling for Optimal Learning

Ján Magyar

Department of Cybernetics and
Artificial Intelligence
Technical University of Košice
Letná 9, 042 00 Košice
Email: jan.magyar@tuke.sk

Peter Sinčák

Department of Cybernetics and
Artificial Intelligence
Technical University of Košice
Letná 9, 042 00 Košice
Email: peter.sincak@tuke.sk

Abstract—Teaching often takes place in a one-size-fits-all way where all students are expected to follow to same schedule, even though cognitive research has shown that a personalized approach yields higher learning gains. Applying methods of machine learning promises highly scalable solutions that tailor the learning process to individual learners' needs. In this paper we propose a general adaptive scheduling method that can adjust the process of skill development by carefully selecting problems to be solved by students to help them better develop their skills. Our method uses reinforcement learning to recognize and address weaknesses in the student's understanding. In a simulated learning process we show that our proposed solution outperforms or performs comparably to a random scheduler.

I. INTRODUCTION

Learning is an inherent part of human life and development starting with infancy. It can take place in a classroom or in an informal setting, and it can be driven by internal motivation or external stimuli. As such, understanding the human learning process with the aim of optimizing it for the highest learning gain and retention rate has been the subject of cognitive research for decades. However, these results were hard to apply in real school settings, since teachers do not have the capacity to personalize their methods to each student. It was with the advent of artificial intelligence that this research gained new momentum since adaptive tutoring systems became feasible and scalable.

In this paper we present our results in designing a general framework for adaptive intelligent tutoring systems that can schedule a human's learning process for optimal learning in the context of skill acquisition. In our solution we use reinforcement learning, a machine learning approach that defines learning as an interaction between an agent and an environment [1]. By executing an action, the agent can bring about a change in the environment's state, which entails a numerical reward. This reward determines the suitability of the given action in the given state. The actor's goal is then to find a policy that maximizes the sum of rewards over the entire interaction.

In our framework, the agent is the adaptive tutoring system, and the environment is the student's knowledge. By carefully selecting study examples for the student to work on, the adaptive tutoring system can facilitate learning. The student's

knowledge is measured through frequent tests, and the change in his or her performance between study sessions is used to calculate the reward. This way we argue that the tutoring system can adapt its scheduling policy to the needs and preferences of individual learners.

The paper is structured as follows: in Section II we describe the theoretical background and major results of research into the cognitive processes behind human learning that our proposal is based on. We describe what an optimal learning process should look like, how it is possible to model human learning and retention, and how spaced repetition can support optimal learning.

In Section III we turn our attention to our proposal, present the general framework of an adaptive intelligent tutoring system, and outline a solution to be used in the context of skill acquisition. Section IV describes the experimental setup we used to evaluate our proposed framework to test its feasibility by comparing it to random study scheduling. We later present some of our results and discuss them in this section. Section V concludes the paper.

II. THEORETICAL BACKGROUND AND RELATED LITERATURE

There are two main types of human learning: knowledge acquisition and skill development. In most learning scenarios, both types are present, e.g. when learning to drive, a student must learn to recognize and interpret road signs, which would constitute knowledge acquisition, while operating the vehicle is an example of skill development. In this section we look at the theory of skill development, how we can model the human learning process and what systems aim to optimize it.

A. Skill development

The main steps of skill development were formally described by Broadwell in 1969 [2]. His four stages of competence describe the basic stages of the learning process; if we are able to correctly identify them, we can adjust the learning process in a way that leads to more optimal learning:

- **unconscious incompetence** – The learner is not aware of the existence and/or relevance of the skill.

- **conscious incompetence** – The learner understands the importance of the skill and perceives his or her own incompetence.
- **conscious competence** – The learner understands how to do something, but it requires concentration.
- **unconscious competence** – The learner has become proficient in performing the skill effortlessly.

Skill development in most settings is done through rote repetition whereby the learner keeps repeating the same actions with the hope that he or she will improve [3]. This approach might lead to some initial learning, but in later stages it results in stagnation or even in degradation of skills as shown in [4].

Further psychological and cognitive research led to the definition of more optimal types of practice. *Purposeful practice* has well-defined specific goals that are easily checked at the end of each practice session [3]. Such feedback is an important part of any practice and inherent belated feedback is present even in rote repetition, but with purposeful practice it is more present already during training and can engage the learner better. To maximize learning gain, the learner should perform right beyond the limits of his or her current comfort zone, challenging his or her abilities.

Optimal learning is highly dependant on the mental representations the learner uses to perform a skill, since having the wrong idea about the skill might limit the student's ability to perform it [5]. Therefore, Ericsson et al. described *deliberate practice* as a type of practice that directly targets and updates these mental representations [6]. It was first observed in domains with highly developed and formalized training methods, such as classical music training [7] and sports [8], as the ability to discern beginners from elite performers is crucial. Generally, deliberate practice involves a trainer or tutor who tailors the training process to the needs of individual learners to address their weaknesses, resulting in personalized training [3]. The introduction of machine learning methods and intelligence tutoring systems could partially replace the role of the teacher, opening up the possibility of deliberate practice to a larger number of people.

B. Memory models and spaced repetition

After the initial learning of a material, its retention and its transfer to long-term memory becomes desirable. It has been observed that for long-term retention, some degree of short-term forgetting is beneficial, as it strengthens the learner's ability to recall the information. This reality stems from the *spacing effect*, which implies that the amount of learned material is greater when learning takes place over a longer period of time and multiple study sessions instead of cramming.

This phenomenon was first described by Ebbinghaus [9]. He identified two variables determining the probability of a learner being able to recall an item: *reinforcement*, i.e. how many times the learner has come across that item, and *delay*, i.e. what time elapsed since the item was last reviewed. Based on these parameters, he defined an approximation of recall probability, which was later elaborated. In [10], recall

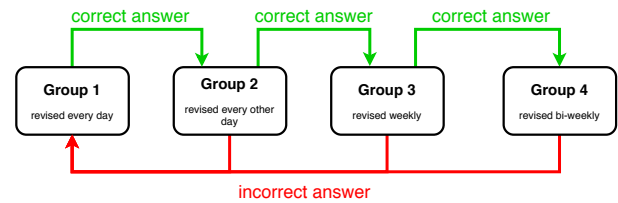


Fig. 1: The Leitner system

probability was calculated based on Ebbinghaus's model as follows:

$$P(\text{recall}) = \exp(-\theta \cdot \frac{D}{S}), \quad (1)$$

where $\theta \in \mathbb{R}_+$ is the item difficulty, $D \in \mathbb{R}_+$ is the delay since the item was last reviewed, and $S \in \mathbb{R}_+$ is the student's memory strength for the item, which can be calculated in various ways.

Half-life regression [11] builds on Ebbinghaus's exponential forgetting curve, but it approximates the item's difficulty from the student's past success in recalling it. This means that the forgetting curve can be personalized to a student's past study history and we arrive at more precise predictions on his or her ability to recall an item correctly.

Even though the spacing effect was described early, the first learning technique that explicitly utilised it – spaced repetition – was introduced only in the 1930s by C. A. Mace [12]. The idea behind spaced repetition is that bits of information that are harder to recall should be revised more frequently than the ones the student can recall easily. Its effectiveness in improving recall was shown in the 1970s [13], and the first general spaced repetition system for flashcards was proposed in 1972 by Sebastian Leitner [14].

The Leitner system (shown in Fig. 1) sorts flashcards into groups based on how hard the learner finds recalling the information described on them. During each study session, the student attempts to recall the information on a given flashcard. Upon successful recall, the flashcard is moved to the next group which is revised at longer intervals. In the case of an incorrect recall, the flashcard is moved to the first group of cards, or into the previous group in some variations. Modern spaced repetition services such as Pimsleur [15], Anki [16], Memrise [17] and Duolingo [18] build on the Leitner system, introducing some level of adaptability to the spacing intervals.

C. Adaptive spaced repetition

Most commercial spaced repetition systems still provide a one-size-fits-all approach to retention with scheduling reflecting a learner's study history of a given item, but following a pre-defined algorithm for calculating review frequency changes. Adaptive review scheduling using machine learning to infer the student's real grasp of a concept would be more precise as it would reflect both an item's strength in the student's memory and his or her ability to retain this information.

Mozer et al. proposed a content-independent approach to modeling student learning for domains where knowledge can be decomposed into distinct components or items [19]. They pointed out that while memory models can describe general forgetting, they do not have predictive power for a single student's ability to recall an item. They therefore introduced the DASH model (difficulty, ability and study history) to predict the likelihood of a student making a correct response at a certain trial. The proposed model was tested in a system offering personalized recommendations to students about materials to study in language learning. The DASH model was continually trained on newly acquired data, and experimental results showed an increase in the amount of material remembered with the scheduler using DASH (over 10% better retention rates compared to a generic scheduler).

Another research project used deep reinforcement learning to provide optimized review schedules [20]. The proposed scheduler balanced conflicting priorities of introducing new items and reviewing old ones while still maintaining a reasonable amount of time spent on studying. The authors showed that an item's difficulty is best expressed with its corresponding position in the Leitner system, which also enables the scheduler to consider some items mastered, i.e. once they are put in the n -th group in the Leitner system. The scheduler's goal is then to maximize the number of items in the n -th group.

From the point of view of reinforcement learning, the study process was described as a partially-observable Markov decision process (POMDP), in which the agent has no complete understanding of the environment's state, similarly to [21]. This stems from the fact that the environment (the student's ability to recall an item) is described by parameters of the memory model, but the scheduling agent can only observe the correctness of the student's answers. The agent's action space is defined as the items it can show to the student, spacing and cramming can be balanced through the careful use of the discount factor γ – a higher value promotes long-term learning. The POMDP problem was then solved through trust region policy optimization with a gated recurrent unit neural network. The input was a triplet with the previously reviewed item, the outcome of the recall attempt, and the time elapsed since the last recall attempt; the network's output was the item to be shown next to the student. The authors compared the proposed scheduler to the Leitner system, the SuperMemo algorithm, and a random scheduler with three student memory models. In each case, their solution outperformed the random and Leitner-based schedulers, with SuperMemo performing better only for one student model.

III. METHODOLOGY

In the previous section, we described intelligent tutoring systems and provided some examples that used reinforcement learning to adapt to individual learners. While these solutions have been implemented and used successfully in various domains, they are usually domain-specific, require intensive domain modelling, eliciting expert knowledge from experienced teachers in the given domain, fine-tuning, and are

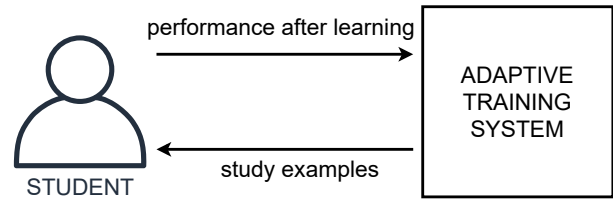


Fig. 2: Interaction between a student and a training system

tested under controlled conditions, which does not relay any information on their usability in a real-life study scenario. If these solutions are to be used in a different domain, extensive reconfiguration must be done.

In this section we propose an adaptive component that could be used in any tutoring system independent of domain. Our framework is proposed in the context of skill acquisition to support adaptive training of students, but the same methodology could be applied in knowledge acquisition. The framework was proposed with two characteristics in mind:

- It is to be used for adaptive training where the student must learn a skill that consists of executing multiple possible tasks. The learner must recognize when to apply what action and then must execute it correctly.
- Our solution should use as little domain-specific information as possible, inferring characteristics of the study material directly from learner performance instead of relying on past expert knowledge.

A. System Architecture and Workflow

For our purposes, we view learning and training in particular as an iterative process, where the student is given a set of problems to solve with the expectation that solving these problems will lead to learning. By carefully selecting these problems, learning can take place quicker and can lead to higher retention. Therefore, we can consider the use of an adaptive training system as an interaction between student and the training system as depicted in Fig. 2. The training system is adaptive when it adjusts study examples to the student's current level of knowledge. It is important to note that the end goal is for each student to reach the same level of understanding, but their paths may be different and should reflect differences in their current knowledge (i.e. one student will already have a working understanding of some tasks related to the skill that is being learned). For optimal learning, training should focus on areas where the student's knowledge is lacking the most. While this will result in a more challenging experience, it will also lead to the most learning.

By exploring the interaction between student and training system, we can easily view it in the context of reinforcement learning. This way, the adaptive training system is the agent whose actions define the set of study examples to be solved by the student. The student's knowledge is then the agent's environment, and by solving the provided problems, the student will learn and his or her performance will change.

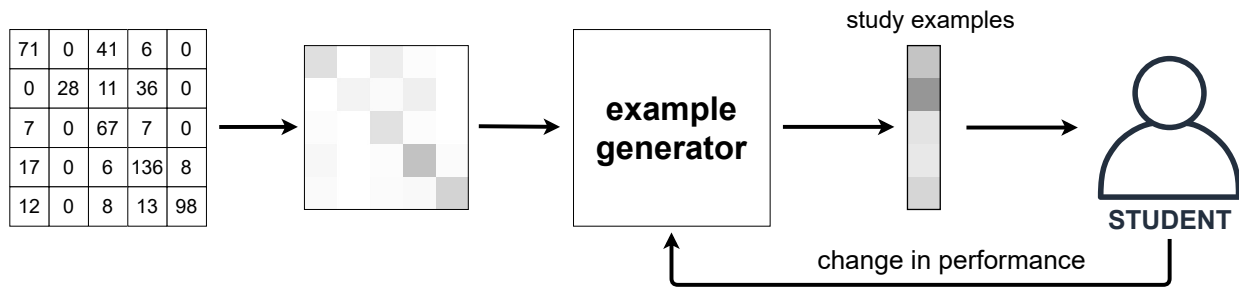


Fig. 3: Architecture and workflow of the adaptive tutor

The student's performance is measured frequently, and based on this, the adaptive training system gets vital feedback on the suitability of the examples it gave the student: the higher the increase in student performance, the better the selected study examples. By using this feedback as a base for reward, we can then train the adaptive training system on-line while it interacts with the student. This enables the system to adjust its behavior to the student.

We can also describe this interaction more concretely in the context of skill acquisition, see Fig. 3. The input to the adaptive trainer is a description of the student's current performance in the given study domain. Considering the way our system is proposed, this domain can be arbitrary, as long as the student's understanding can be represented as a confusion matrix. This means that to solve a problem, the student must recognize what type of problem he or she faces, what tasks comprising the skill are relevant to solving the problem, and then must apply these tasks. If the student chooses the inappropriate concept or skill, this is considered to be an error and we assume that the student fails to solve the problem. This way, we represent training and skill acquisition as a classification-like problem, where students must learn to recognize the skills they must apply when solving a problem. Naturally, for effective learning the student must get better at these skills through repetition, but that is not in the scope of our proposed solution; the adaptive tutor simply selects the problems the student will benefit from the most.

The generated confusion matrix serves as the state representation and is converted into an image. The image is then fed as input to the adaptive tutor, which consists of an artificial convolutional neural network (shown as example generator in Fig. 3). The tutor generates the next batch of problems the student must solve. On the lowest level, the tutor generates only the distributions of the problems for each category of problems (related to individual skills the student must learn), an independent component must then generate the problems themselves based on this distribution. The problems are then showed to the student who reviews and solves them, learning in the process. We then observe the change in the student's performance, which serves as reward signal for the adaptive tutor. As in most related research, the normalized learning gain (NLG) is used for this purpose.

Student performance can be measured in two ways: in

explicit test rounds or by considering the last n problems the student solved. In either case, the system must determine how often he or she was able to apply the right skill or sub-skill to solve the problem. However, if we consider only the last examples solved by the student without an explicit test round, the adaptive tutor might bypass the learning goal by giving the student only examples the student can already solve easily. While this will lead to temporary learning gains from the system's point of view, in reality, there will be no significant learning in the long run. This problem can be solved by penalising unbalanced problem distributions of study examples. However, the system must be given leeway so it can favor types of examples causing students problems, while neglecting skills that are already mastered.

B. Learning Process Modeling

To model the learning process and the student's knowledge, we used the exponential forgetting curve to calculate the probability that a student gives a correct answer for a problem of a given type:

$$P(\text{correct}) = e^{-\theta \cdot \frac{c \cdot D}{S}}, \quad (2)$$

where θ is the skill's difficulty, D is the delay since the item was last reviewed, and S is the student's memory strength for the item. For our experiments, θ was sampled from the interval $< 0, 1 >$. We further adjusted the elapsed time by introducing a time delay constant c . Instead of simply calculating the number of days passed since the last recall attempt (as in spaced repetition systems), we calculated the number of training sessions since the skill was used last and then multiplied it by a constant c to arrive at more realistic recall rates. The effect of different values of c on recall probability with respect to elapsed time D is shown in Fig. 4.

As in [10], we considered three variations of the exponential forgetting curve. First, we used the number of problems within a session for the given skill the student faced. Second, we considered the number of correct recognition of the skill, but only considering at most one correct attempt per past session. Finally, we used a counter analogous to the Leitner system, where the strength of an item increased after every correct recognition (at most one per session), and decreased after an incorrect recognition (possibly multiple times within

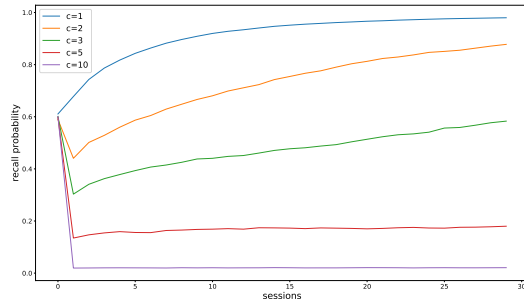


Fig. 4: Effect of time delay constant c on recall probability $P(\text{recall})$ with random scheduling

a session). The smallest value for the Leitner counter was set to 1 with no upper bound.

We simulated a test round after each session, and calculated the normalized learning gain between performances in two test rounds as follows:

$$NLG = \frac{\text{posttest} - \text{pretest}}{1 - \text{pretest}}. \quad (3)$$

IV. EXPERIMENTS AND RESULTS

We tested our proposed solution on simulated training of three types of students with three different study materials. The items in the study material were generated randomly at the beginning of each test run with their difficulty sampled from a set interval with uniform distribution. In the first study material, item difficulty ranged between 0.8 and 1.0; in the second material, it ranged between 0.6 and 0.8; in the third material, it ranged between 0.35 and 0.45.

In each study material, there were 10 items (10 skills). To ensure that the system had to adapt to the student, the set of study examples comprised six examples at each iteration. This forces the adaptive scheduling agent to select only some of the skills for each training session. Otherwise, it would give the student an example for each skill, which would result in high learning, but requires no adaptation. During a simulation, we considered 30 study sessions, and determined the student's knowledge level from the probability of correct recognition for each item. To this end, we simulated student answers for 5 questions per item (50 questions in total), remembering correct and incorrect answers. For incorrect answers we simulated with what other skill the student confused the correct one; this incorrect answer was generated based on a randomly set similarity between skills at the beginning of each test run. The generated confusion matrix served as input to the adaptive trainer.

For the implementation of the trainer, we used a deep deterministic policy gradient (DDPG) agent [22] with a continuous action space. The action of the agent was described by values ranging from 0 to 1 for each item, the values representing the portion of questions for the given item in the next batch of training examples. The student's knowledge was then trained

on these examples by simulating student answers and updating memory strength. At the end of each session, the adaptive trainer remembered the experience (student performance before training, action as the distribution of study examples, reward, student performance after training), and was trained on a sample of 64 such experiences. Learning rate was 0.001 for the actor network and 0.002 for the critic. The discount rate was set to 0.99.

As described earlier, we considered three types of students who differed in the way memory strength was defined. In total, we ran 9 test scenarios (3 materials \times 3 student types) with different time delay constants c . We compared learning gain with random scheduling (the distribution of study items in the next set of study examples was sampled randomly) to that with our proposed adaptive component (the distribution was defined by the output of our DDPG agent).

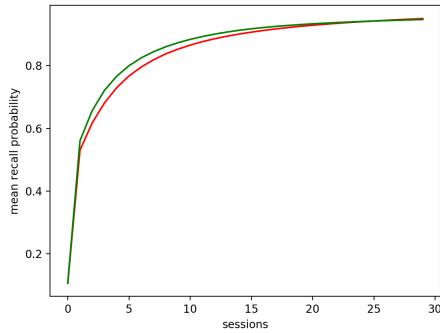
The learning gain was calculated by considering the average recall likelihood of the 10 items in the given study material based on the exponential forgetting curve. We ran 1000 runs with 30 sessions for each setup, and considered the average over these runs. Recall likelihood at the beginning of training was defined as $1 - \theta$. We expected that our model would lead to faster and more learning, showing as a steeper growth in average recall probability across items and higher average recall likelihood in the final session.

We ran these experiments with time delay constant values 1, 2, 3 and 5. Here we show and discuss results for $c = 1$, but similar trends were observed with other values. Graphs with results from experiments with other c values are available in the Appendix.

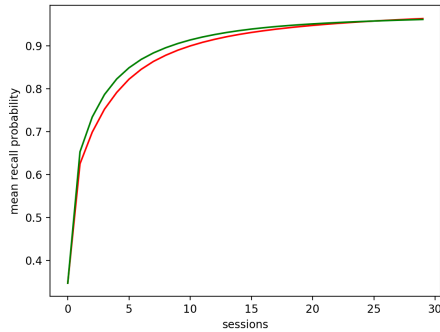
The smallest difference was observed when memory strength was defined by the total number of recall attempts for each item (see Fig. 5). Here, while learning was quicker with adaptive scheduling in early phases, learning gains with random scheduling caught up with adaptive scheduling by session 25, and slightly outperformed adaptive scheduling by the end of training at session 30. The same results were observed across different study materials, the items' difficulty influenced only the final recall probability. For study material 1 (highest average difficulty), final average recall likelihood was 0.94 with the adaptive trainer, 0.95 with random scheduling; for study material 2, the final value was 0.96 with both schedulers; and for study material 3, the final value was 0.97 with both schedulers. We can see that there was no significant difference between the effect of either scheduling on learning.

When using the number of correct recalls to define memory strength (shown in Fig. 6), the differences between scheduling were bigger. Training based on an adaptive schedule resulted once again in faster learning in early phases, and the random scheduler failed to catch up with its adaptive counterpart in thirty sessions, although it got closer. For study material 1, final average recall likelihood was 0.92 with adaptive training and 0.91 with random scheduling; for study material 2, the values were 0.95 and 0.94 respectively; for study material 3 were 0.97 and 0.96.

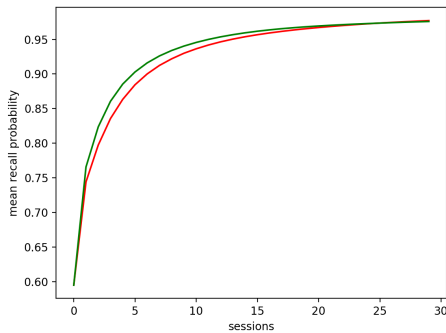
The largest difference was observed when using an analogy



(a) Study material 1

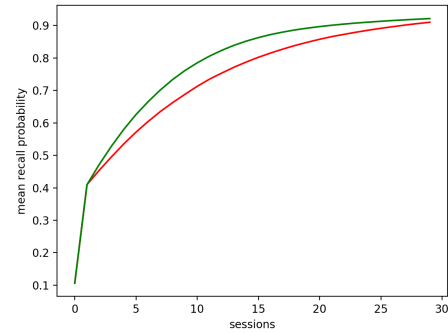


(b) Study material 2

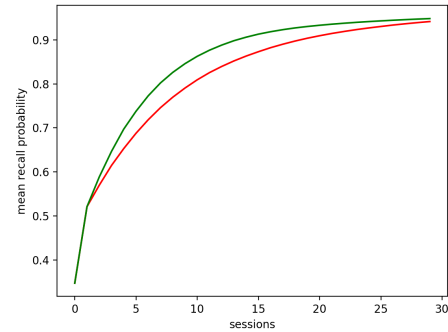


(c) Study material 3

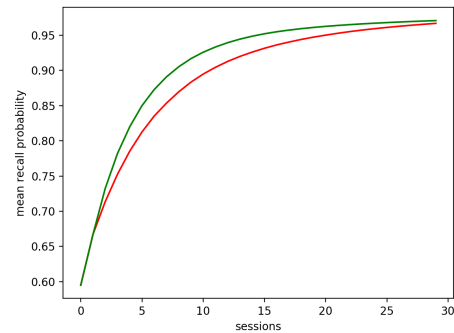
Fig. 5: Change in the mean recall likelihood across items for study materials with memory strength defined by the total number of recall attempts with random scheduling (red) and adaptive scheduling (green)



(a) Study material 1



(b) Study material 2



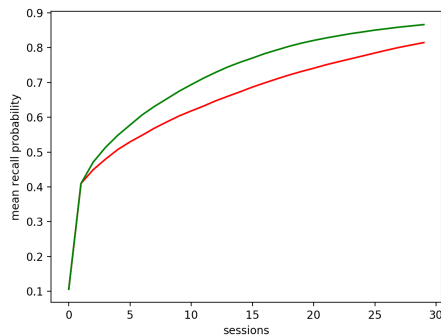
(c) Study material 3

Fig. 6: Change in the mean recall likelihood across items for study materials with memory strength defined by the number of correct recalls with random scheduling (red) and adaptive scheduling (green)

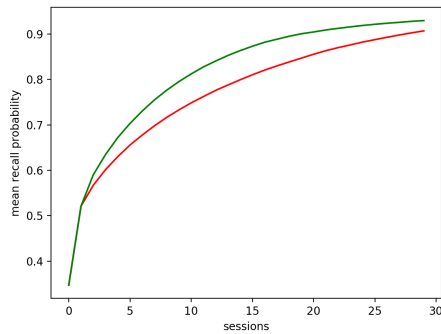
of the Leitner system for memory strength (see Fig. 7). This modeling is the one that gets closest to real learning on average. We observed a clear dominance of the adaptive training component, although the effect was mitigated by smaller study item difficulty. For study material 1, the final average recall likelihood was 0.87 with the adaptive trainer and 0.81 with random scheduling; for study material 2 the values were 0.93 and 0.91 respectively; for study material 3, the values were 0.97 and 0.96. Even though the average recall likelihoods were close after session 30, the adaptive scheduling resulted in steeper learning in early phases.

V. CONCLUSION

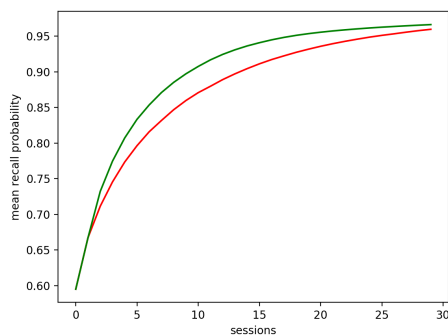
In this paper we described a general architecture of an adaptive scheduling component for optimizing human learning for skill development. To this end, we considered the learning process as an iterative interaction between the tutoring system and a student where the student is given a set of problems to solve by the system. By careful selection of these problems, the learning process can be optimized for the highest learning gain. The proposed system uses reinforcement learning to model the student’s understanding of the material being taught and to adjust the question selection for each individual study



(a) Study material 1



(b) Study material 2



(c) Study material 3

Fig. 7: Change in the mean recall likelihood across items for study materials with memory strength defined defined by analogous position in the Leitner system with random scheduling (red) and adaptive scheduling (green)

session. We then simulated human learners using the exponential forgetting curve model of human memory, and showed that our system leads to steeper and higher learning than random scheduling in most test scenarios. Even when the random scheduler outperformed our solution, their performance was comparable.

The main contribution of our proposed solution is that our scheduler can be used in the context of skill development instead of knowledge acquisition, as long as the skill can be represented as a classification problem where the student must learn to correctly recognize the task or sub-skill relevant to solving a given problem. In this way, our method could be

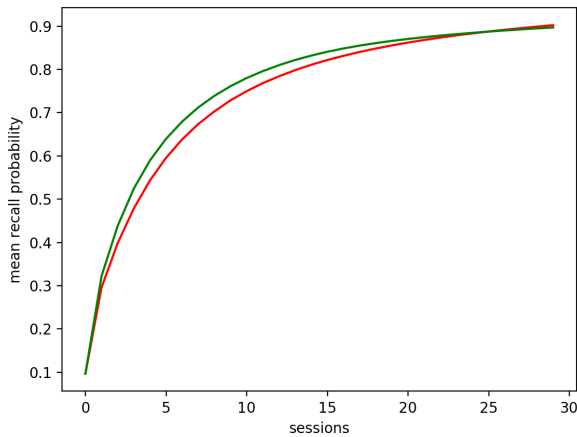
adapted to knowledge acquisition if the goal is to teach the student to correctly classify or recognize items (e.g. assigning a periodic element to its correct group).

The results presented in this paper support the feasibility of the proposed solution, but further extensive testing is needed. Future research might include testing with other student models, and also developing an application incorporating the adaptive scheduler to support skill development. This application can then be used to measure the effects of the use of the proposed adaptive training scheduler in a real-life learning scenario.

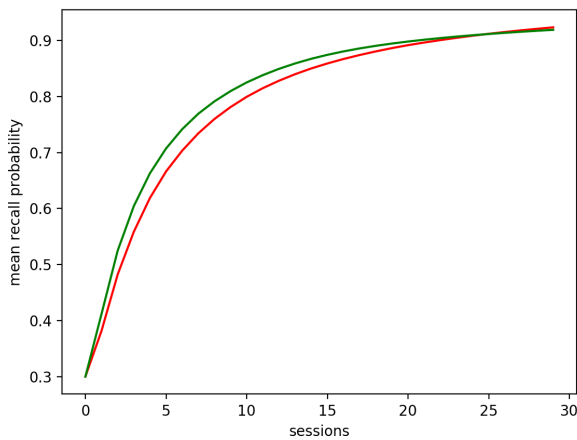
REFERENCES

- [1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [2] M. M. Broadwell, "Teaching for Learning (XVI)," *The Gospel Guardian*, vol. 20, no. 41, pp. 1–3, 1969.
- [3] A. Ericsson and R. Pool, *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.
- [4] K. A. Ericsson, "Acquisition and Maintenance of Medical Expertise: A Perspective from the Expert-Performance Approach with Deliberate Practice," *Academic Medicine*, vol. 90, no. 11, pp. 1471–1486, 2015.
- [5] W. G. Chase and K. A. Ericsson, "Skilled Memory," *Cognitive Skills and Their Acquisition*, pp. 141–189, 1981.
- [6] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review*, vol. 100, no. 3, p. 363, 1993.
- [7] K. A. Ericsson, C. Tesch-Römer, and R. T. Krampe, "The Role of Practice and Motivation in the Acquisition of Expert-Level Performance in Real Life: An Empirical Evaluation of a Theoretical Framework," in *Encouraging the Development of Exceptional Skills and Talents*. British Psychological Society, 1990, pp. 109–130.
- [8] K. A. Ericsson, "Development of Elite Performance and Deliberate Practice: An Update from the Perspective of the Expert Performance Approach," *Expert Performance in Sports: Advances in Research on Sport Expertise*, pp. 49–84, 2003.
- [9] H. Ebbinghaus, "Memory: A Contribution to Experimental Psychology," *Annals of Neurosciences*, vol. 20, no. 4, pp. 155–156, 2013.
- [10] S. Reddy, I. Labutov, S. Banerjee, and T. Joachims, "Unbounded human learning: Optimal scheduling for spaced repetition," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 2016, pp. 1815–1824.
- [11] B. Settles and B. Meeder, "A Trainable Spaced Repetition Model for Language Learning," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2016, pp. 1848–1858.
- [12] C. A. Mace, "The Psychology of Study," 1932.
- [13] A. W. Melton, "The Situation with Respect to the Spacing of Repetitions and Memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 9, no. 5, pp. 596–606, 1970.
- [14] S. Leitner, *So lernt man lernen: Der Weg zum Erfolg [How to Learn to Learn: The Way to Success]*. Freiburg, Germany: Herder, 1972.
- [15] P. Pimsleur, "A Memory Schedule," *The Modern Language Journal*, vol. 51, no. 2, pp. 73–75, 1967.
- [16] Anki, "Anki main web page," <https://apps.ankiweb.net/>, 2021.
- [17] Memrise, "Memrise main web page," <https://www.memrise.com/>, 2021.
- [18] Duolingo, "Duolingo main web page," <https://www.duolingo.com/>, 2021.
- [19] M. C. Mozer and R. V. Lindsey, "Predicting and Improving Memory Retention: Psychological Theory Matters in the Big Data Era," *Big Data in Cognitive Science*, p. 34, 2016.
- [20] S. Reddy, S. Levine, and A. Dragan, "Accelerating Human Learning with Deep Reinforcement Learning," in *NIPS Workshop: Teaching Machines, Robots, and Humans*, 2017.
- [21] A. N. Rafferty, E. Brunskill, T. L. Griffiths, and P. Shafto, "Faster Teaching via POMDP Planning," *Cognitive Science*, vol. 40, no. 6, pp. 1290–1332, 2016.
- [22] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

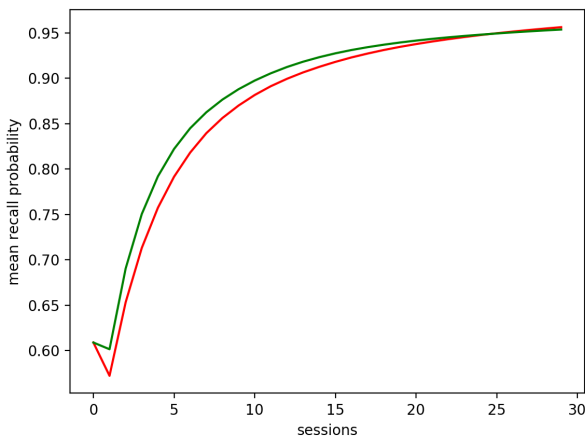
APPENDIX



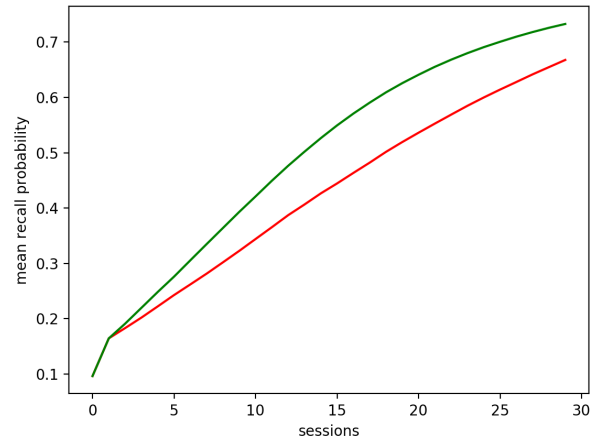
(a) Study material 1



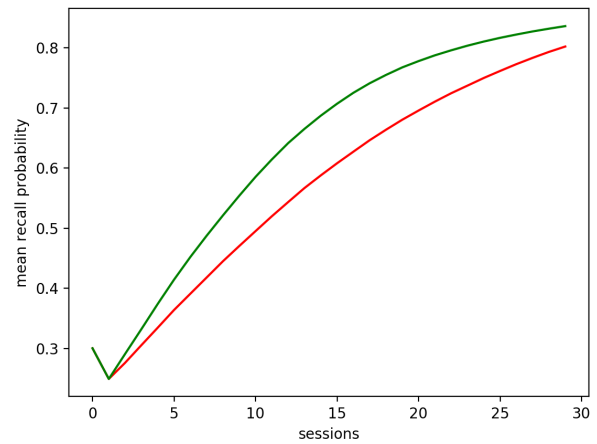
(b) Study material 2



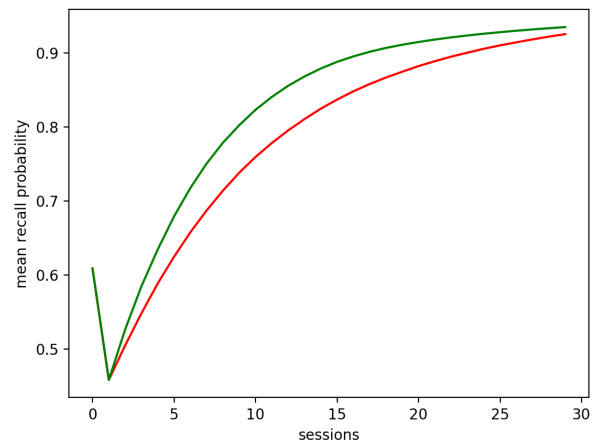
(c) Study material 3



(a) Study material 1



(b) Study material 2



(c) Study material 3

Fig. 8: Change in the mean recall likelihood across items for study materials with memory strength defined by the total number of recall attempts with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 2$

Fig. 9: Change in the mean recall likelihood across items for study materials with memory strength defined by the number of correct recalls with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 2$

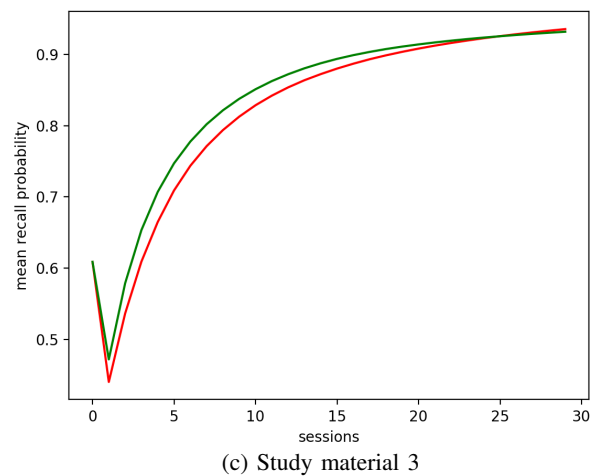
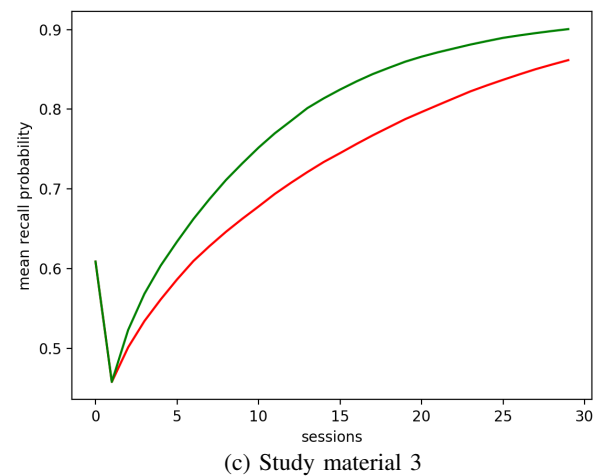
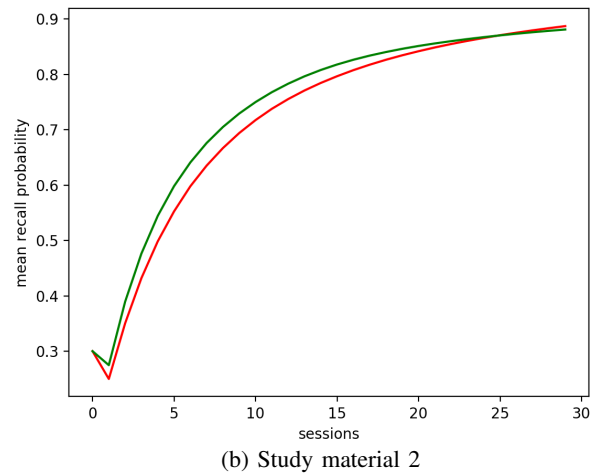
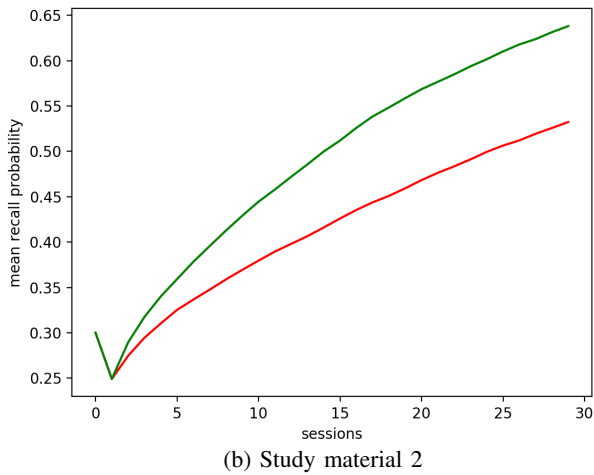
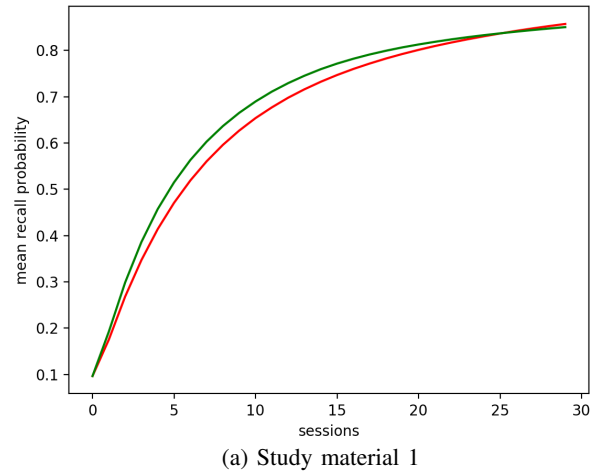
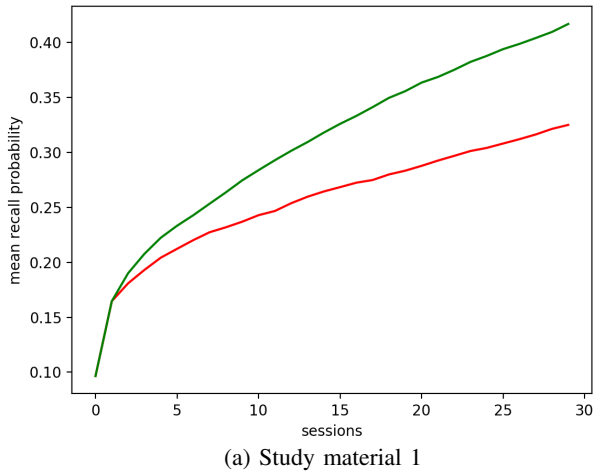


Fig. 10: Change in the mean recall likelihood across items for study materials with memory strength defined by analogous position in the Leitner system with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 2$

Fig. 11: Change in the mean recall likelihood across items for study materials with memory strength defined by the total number of recall attempts with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 3$

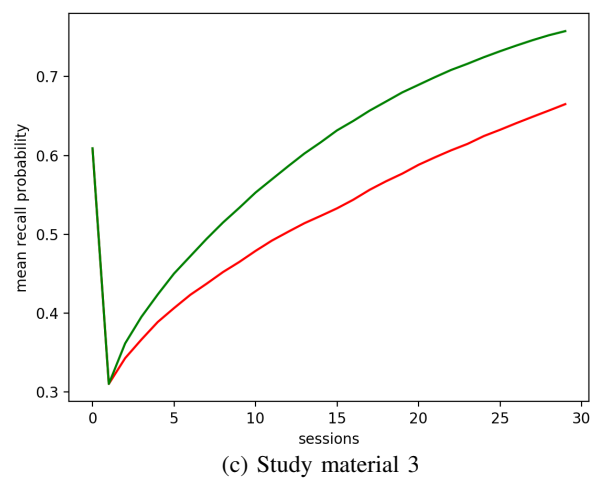
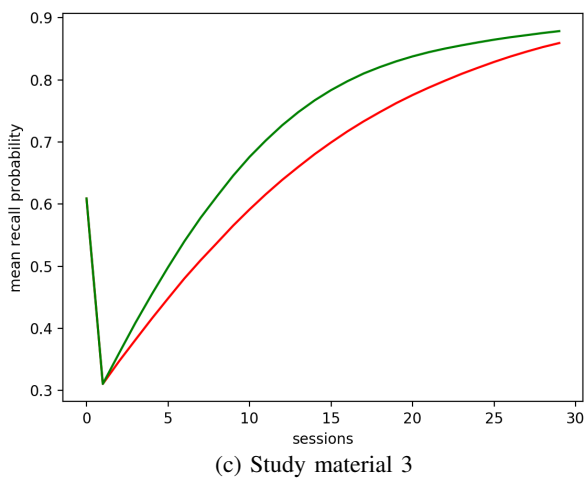
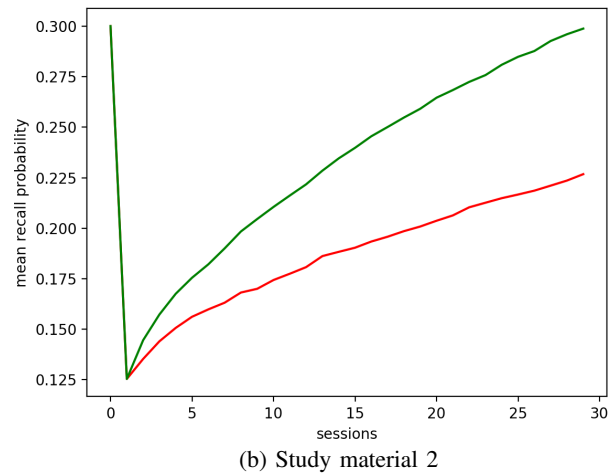
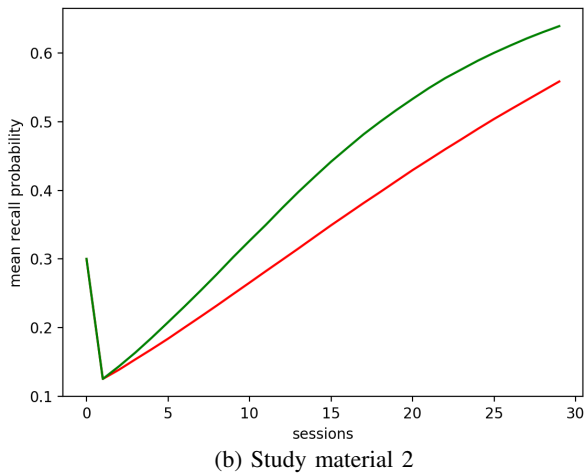
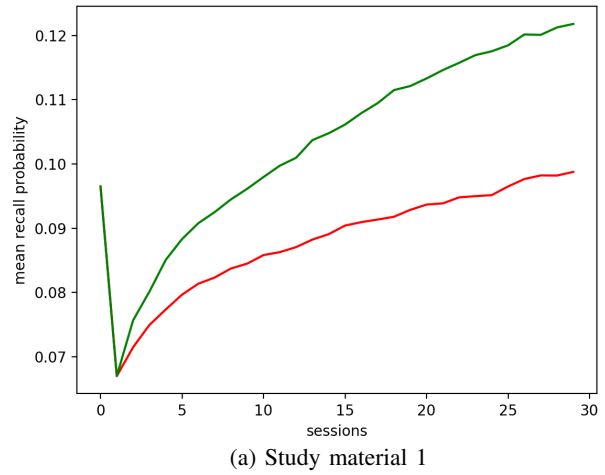
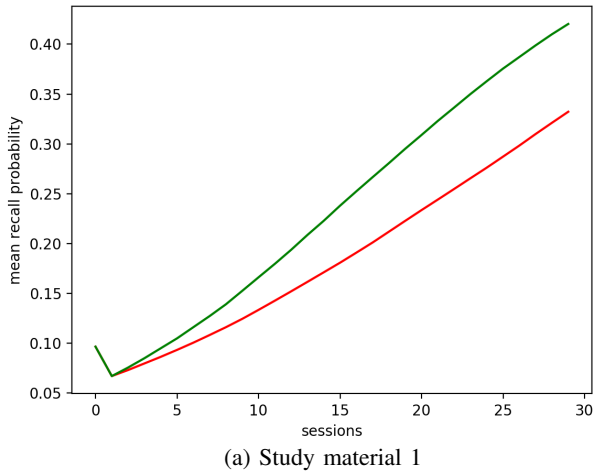


Fig. 12: Change in the mean recall likelihood across items for study materials with memory strength defined by the number of correct recalls with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 3$

Fig. 13: Change in the mean recall likelihood across items for study materials with memory strength defined by analogous position in the Leitner system with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 3$

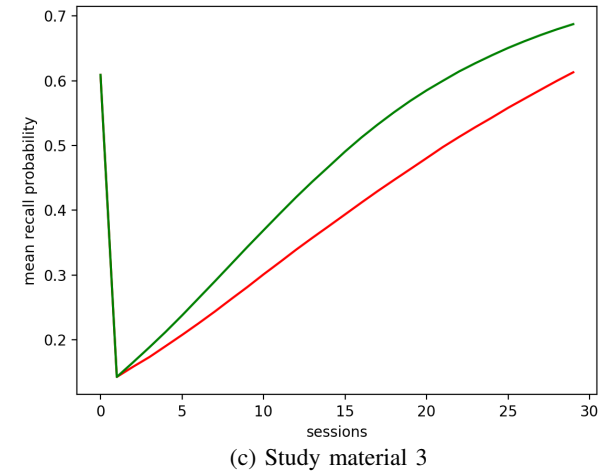
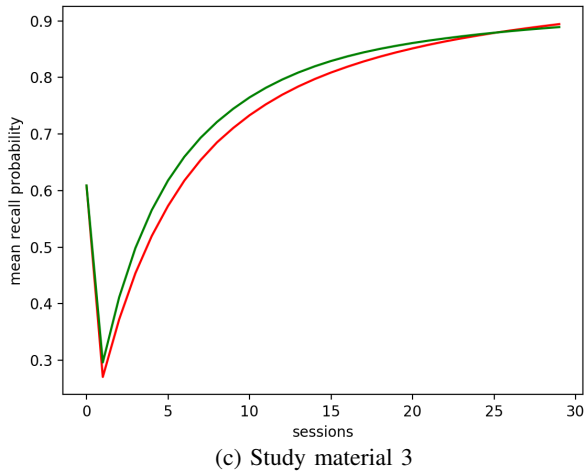
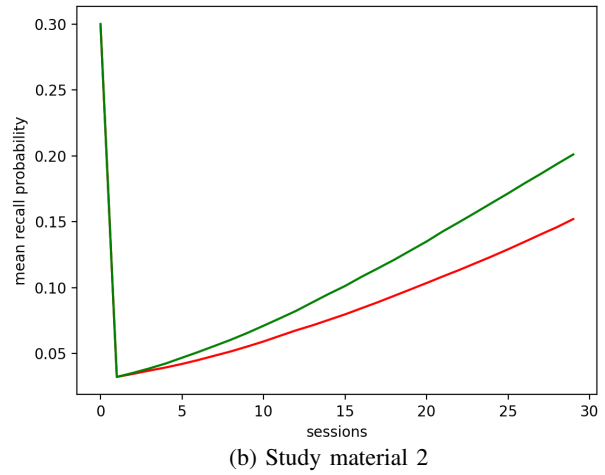
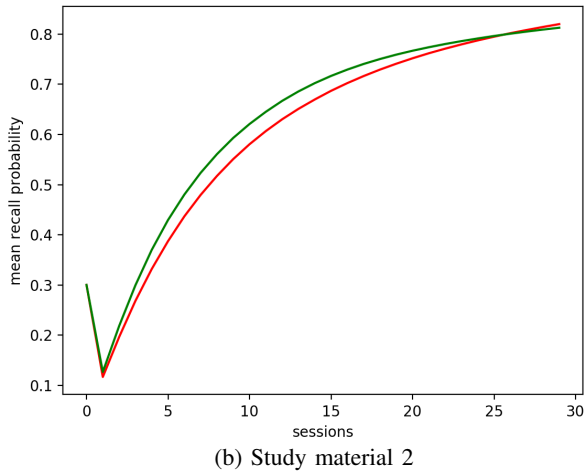
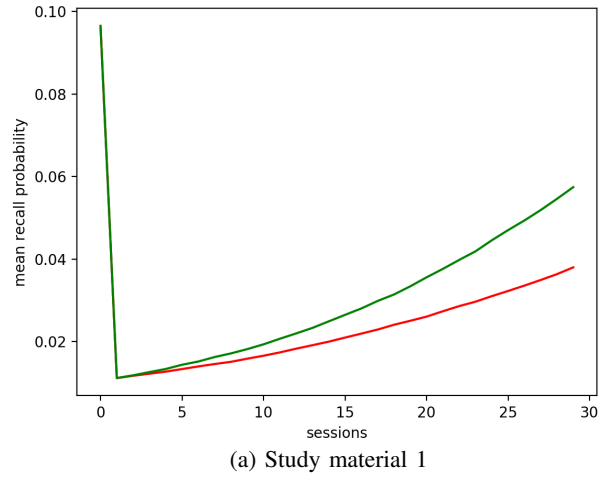
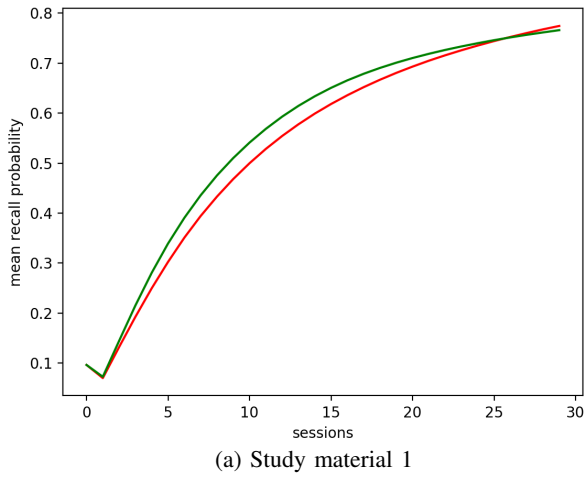
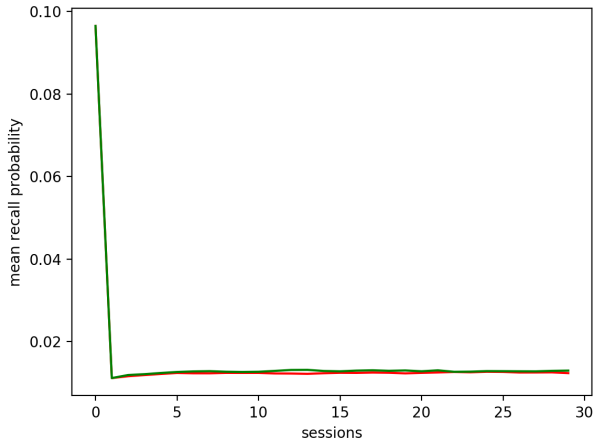
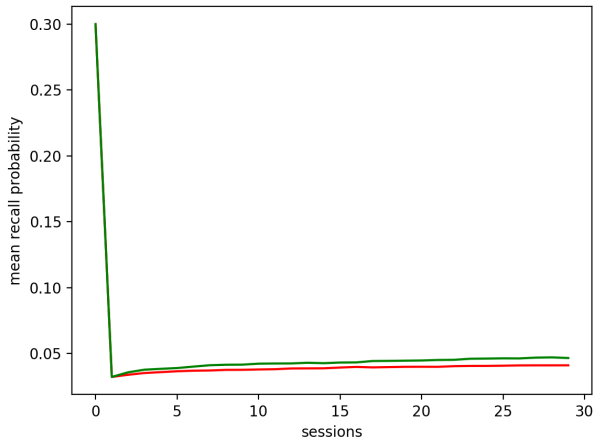


Fig. 14: Change in the mean recall likelihood across items for study materials with memory strength defined by the total number of recall attempts with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 5$

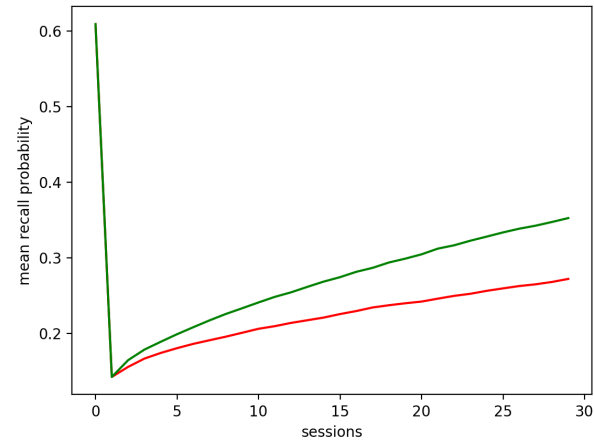
Fig. 15: Change in the mean recall likelihood across items for study materials with memory strength defined by the number of correct recalls with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 5$



(a) Study material 1



(b) Study material 2



(c) Study material 3

Fig. 16: Change in the mean recall likelihood across items for study materials with memory strength defined by analogous position in the Leitner system with random scheduling (red) and adaptive scheduling (green) – time delay constant $c = 5$

Návrh metodiky pre modelovanie, analýzu a simuláciu efektívne kráčajúcich robotických systémov

Lukáš Koska

Katedra kybernetiky
a umelej inteligencie

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
lukas.koska@tuke.sk

Anna Jadlovská

Katedra kybernetiky
a umelej inteligencie

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
anna.jadlovska@tuke.sk

Slávka Jadlovská

Katedra kybernetiky
a umelej inteligencie

Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
slavka.jadlovska@tuke.sk

Abstrakt—Tento príspevok prezentuje vytvorenú metodiku pre modelovanie, analýzu a riadenie dvojnôhých kráčajúcich robotov v kontexte hybridných systémov. Cieľom navrhutej metodiky modelovania je umožniť generovanie matematických modelov podaktuovaných kráčajúcich robotov s využitím klasickej mechaniky a teórie hybridných systémov. Na následnú analýzu sú využité limitné (medzné) cykly a Poincarého mapy. V rámci metodiky riadenia je riešená úloha sledovania generovaných referenčných trajektórií v riadiacej štruktúre s dopredným riadením, čomu predchádza úprava modelov s využitím metódy parciálnej spätnoväzobnej linearizácie spolu so zapojením príslušného algoritmu riadenia. Predstavené sú vybrané prípadové štúdie v oblasti modelovania a riadenia na báze rôznych konfigurácií kráčajúcich robotov. Softvérová aplikácia, umožňujúca simulačne overiť správnosť navrhnutých postupov, je implementovaná v programovom prostredí MATLAB/Simulink/Stateflow.

Kľúčové slová—kráčajúce roboty, hybridné systémy, modelovanie nelineárnych systémov, riadenie kráčajúcich robotov, kružidlová chôdza, plánovanie trajektórie, sledovanie trajektórie, dvojbodová okrajová úloha, optimalizácia evolučným algoritmom

Abstract—This paper presents the design and development of a methodology for modeling, analysis and control of bipedal walking robots in the context of hybrid systems. The aim of proposed modeling methodology is to enable the generation of mathematical models of underactuated walking robots using classical mechanics and hybrid systems theory. Limit cycles and Poincaré maps are used for subsequent analysis. Control methodology focuses on solving the task of tracking generated reference trajectories in the feedforward control structure, which is preceded by the modification of models using the partial feedback linearization method with the involvement of an appropriate control algorithm. Selected case studies in the field of modeling and control, based on different configurations of underactuated walking robots, are presented. The software application, enabling simulation to verify the accuracy of proposed procedures, is implemented in the MATLAB/Simulink/Stateflow environment.

Index Terms—walking robots, hybrid systems, modeling of nonlinear systems, control of walking robots, compass gait, trajectory planning, trajectory tracking, boundary value problem, evolutionary algorithm

I. ÚVOD

Kráčajúce roboty (roboty vykonávajúce pohyb pomocou mechanizmov podobných nohám) využívajú jednu z najsofistikovanejších a najefektívnejších foriem pohybu - chôdzu. Dvojnôhá chôdza je charakterizovaná striedavým dvíhaním dvojice chodidiel na konci nôh, pričom v ktoromkoľvek momente sa aspoň jedno z nich nachádza na zemi. Je zrejmé, že význam dvojnohej chôdze vychádza z inšpirácie ľudskou chôdzou. Z hľadiska výskumu v oblasti robotiky ľudská chôdza vyčnieva medzi inými formami dvojnôhého pohybu, kvôli skutočnosti, že počas veľkej časti kroku pohybujúce sa telo nie je v statickej rovnovážnej polohe [1]. Keďže ľudské telo má mnoho stupňov voľnosti, je prijateľnejšie pre výskum v oblasti kráčajúcich robotov a ich chôdze vychádzať zo zjednodušeného modelu ľudského tela a využívať jednoduchšie typové kráčajúce roboty, ktoré majú menej stupňov voľnosti.

Riadenie kráčajúcich robotov, tak ako aj riadenie iných mechanických systémov, vyžaduje presný matematický model uvažovaného riadeného systému [2]. Z pohľadu modelovania sú kráčajúce roboty výnimočné tým, že okrem spojitej dynamiky existujú pri kráčaní aj diskkrétne udalosti, napríklad pri dopade švihovej nohy na podložku. Z tohto dôvodu je vhodné využívať pre modelovanie a analýzu kráčajúcich robotov koncept hybridných systémov.

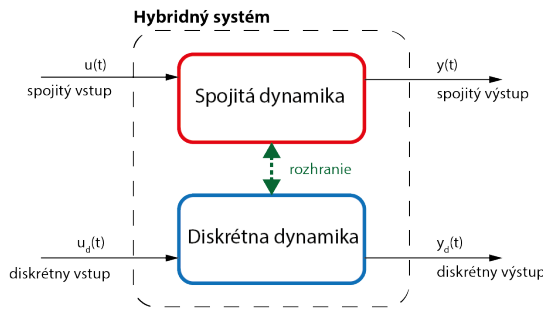
Pri návrhu riadenia kráčajúceho robota je možné uvažovať s plne aktuovaným alebo podaktuovaným robotom [3][4]. Na rozdiel od implementácie dvojnôhého pohybu vo forme statickej chôdze, kde pri riadení všetkých stupňov voľnosti (plne aktuovaný robot) zotrúva priemet ťažiska neustále v oblasti chodidiel za cenu zvýšenej spotreby energie a istej neprirodzenosti pohybu, dynamická chôdza vychádza z prirodzenej dynamiky systému a teda je výrazne energeticky úspornejšia. Takéto systémy je vhodné modelovať ako podaktuované (t.j. majúce menej riadiacich vstupov ako stupňov voľnosti) a cieľom riadenia je dosiahnuť biologicky inšpirovanú, energeticky úspornú chôdzu. Z pohľadu modelovania

vania nedochádza k žiadnym zmenám oproti plneaktuovaným systémom. Podaktuovaný systém je možné vytvoriť z rôznych plneaktuovaných systémov, ktoré majú vhodné vnútorné väzby v rámci dynamiky systému. Zložité podaktuované systémy sa dajú zjednodušiť na takzvané typové (benchmarkové) systémy. Medzi takéto jednoduchšie typové podaktuované systémy patrí napríklad vozík s kyvadlom, z ktorého sa postupne rôznymi úpravami vyvinuli aj systémy schopné robotického chôdze [5].

II. MODELOVANIE A ANALÝZA KRÁČAJÚCICH ROBOTOV AKO HYBRIDNÝCH SYSTÉMOV

Hybridné systémy sú dynamické systémy, ktoré pozostávajú z komponentov so spojitým a diskretným správaním a preto predstavujú syntézu spojitých a diskretných dynamík. Hybridný systém musí tiež obsahovať interakciu medzi spojitou a diskretnou dynamikou. Každý diskretný stav vyniká svojou vlastnou spojitou dynamikou. Interakcia je obzvlášť dôležitá v mnohých technologických procesoch, v ktorých logické rozhodovanie a integrované riadiace činnosti sa spájajú so spojitými fyzikálnymi procesmi. Na zachytenie správania týchto systémov sú potrebné matematické formalizmy, ktoré kombinujú dynamiku spojitých častí systému s dynamikou diskretných častí [6].

Koncept hybridných systémov je možné použiť aj na kráčajúce roboty, ktorých dynamika vykazuje skokové zmeny alebo sa s časom mení. Pre modelovanie kráčajúcich robotov s hybridnou dynamikou stačí využiť len časť z celej rozsiahlej teórie hybridných systémov a aj z tohto dôvodu táto prezentácia výsledkov dizertačnej práce neobsahuje prehľad celej teórie matematických zápisov hybridných systémov.



Obr. 1. Všeobecná štruktúra hybridného systému

Hybridné systémy v zmysle Obr. 1 zahŕňajú spojitú dynamiku so spojitým vstupom $u(t)$ a spojitým výstupom $y(t)$ a diskretnú dynamiku s diskretným vstupom $u_d(t)$ a diskretným výstupom $y_d(t)$, ako aj rozhranie medzi spojitou a diskretnou dynamikou [7].

A. Modelovanie spojitej dynamiky

Použitie Lagrangeových rovníc II. druhu pre modelovanie kráčajúcich robotov umožňuje algoritimizáciu odvodenia diferenciálnych pohybových rovníc [8] s využitím *Symbolic Toolbox-u* programového prostredia MATLAB [9]. S využitím Lagrangeových rovníc II. druhu je pre m stupňov voľnosti kráčajúceho robota nutné odvodiť m diferenciálnych

pohybových rovníc druhého rádu popisujúcich dynamiku jednotlivých stupňov voľnosti.

Základom odvodenia spojitých pohybových rovníc je definovanie zovšeobecnených súradníc hmotných bodov $\theta(t)$ [10], pričom v tejto práci sú uvažované nasledovne:

$$\theta(t) = [\theta_0(t) \quad \theta_1(t) \quad \dots \quad \theta_n(t)] \quad (1)$$

Príklad, ako môžu byť definované zovšeobecnené súradnice $\theta(t)$ hybridný model kružidlovej chôdze je na Obr. 2. Pomocou vektora zovšeobecnených súradníc $\theta(t)$ je možné definovať pre každý i -ty hmotný bod kinetickú E_{k_i} a potenciálnu E_{p_i} energiu [10], pričom celková kinetická E_k a potenciálna E_p energia pre n hmotných bodov je definovaná nasledovne:

$$E_k = \sum_{i=1}^n E_{k_i} = \sum_{i=1}^n \frac{1}{2} m_i v_i^2 \quad (2)$$

$$E_p = \sum_{i=1}^n E_{p_i} = \sum_{i=1}^n m_i g h_i$$

Rozdiel kinetickej a potenciálnej energie je vyjadrený ako Lagrangeova funkcia [11] zovšeobecnených súradníc $\theta(t)$ a ich rýchlostí $\dot{\theta}(t)$:

$$L(\theta(t), \dot{\theta}(t)) = E_k(\theta(t), \dot{\theta}(t)) - E_p(\theta(t)) \quad (3)$$

S uvažovaním $Q^*(t)$ ako vektora vonkajších síl pôsobiacich na systém, Lagrangeove rovnice II. druhu majú tvar:

$$\frac{d}{dt} \left(\frac{\partial L(t)}{\partial \dot{\theta}(t)} \right) - \frac{\partial L(t)}{\partial \theta(t)} = Q^*(t) \quad (4)$$

Derivácia a parciálna derivácia Lagrangeovej funkcie (3) vedie k systému pohybových diferenciálnych rovníc $\dot{\theta}(t) = f(\theta(t), t)$ (4) popisujúcich spojitú dynamiku jednotlivých stupňov voľnosti kráčajúceho robota [12]. Ďalšou možnosťou zápisu výslednej pohybovej rovnice kráčajúcich robotov je zápis pomocou tzv. štandardnej minimálnej formy (ODE) [13], ktorá má tvar:

$$M(\theta(t))\ddot{\theta}(t) + N(\dot{\theta}(t), \theta(t))\dot{\theta}(t) + P(\theta(t)) = V(t)u(t) \quad (5)$$

kde $M(t)$ je matica zotrvačnosti, $N(t)$ je matica Coriolisových síl, $P(t)$ je matica potenciálov a matica $V(t)$ mapuje vstupy $u(t)$ do všeobecných súradníc [14]. Ak je hodnota matice $V(t)$ rovná rozmeru vektora zovšeobecnených súradníc, systém je plneaktuovaný $rank(V) = size(\theta)$, ak hodnota matice $V(t)$ je menšia ako je rozmer vektora zovšeobecnených súradníc, hovoríme, že systém je podaktuovaný $rank(V) < size(\theta)$ [15].

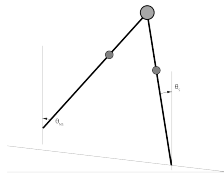
Spojité stavový vektor $x(t)$ pre konkrétny i -ty diskretný stav x_d môže byť definovaný ako:

$$x(t) = [\theta(t) \quad \dot{\theta}(t)] \quad (6)$$

Pre analýzu alebo návrh algoritmu riadenia hybridného modelu kráčajúceho robota je ďalej možné prepísať ODE formu (5) s uvažovaním stavového vektora (6) na tvar:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ &= \begin{pmatrix} \dot{\boldsymbol{\theta}}(t) \\ \mathbf{M}(\boldsymbol{\theta}(t))^{-1}[\mathbf{N}(\dot{\boldsymbol{\theta}}(t), \boldsymbol{\theta}(t))\dot{\boldsymbol{\theta}}(t) + \mathbf{P}(\boldsymbol{\theta}(t)) - \mathbf{V}(t)\mathbf{u}(t)] \end{pmatrix} \end{aligned} \quad (7)$$

kde $\mathbf{f}(\mathbf{x}, \mathbf{u})$ sa skladá z jednotlivých nelineárnych diferenciálnych rovníc popisujúcich dynamiku stupňov voľnosti a popisuje celkovú spojitú dynamiku kráčajúceho robota pre nemeniaci sa diskretný stav x_d .



Obr. 2. Schematický náčrt kružidlovej chôdze. θ_S označuje uhol stojnej nohy, θ_{NS} označuje uhol švihovej nohy

B. Modelovanie diskretnéj dynamiky

Diskretná dynamika tvoriaca jednu z častí celkovej dynamiky hybridného systému je definovaná existenciou udalostí špecifikovaných veličinami charakterizujúcimi spojitú dynamiku $\mathbf{x}(t), \mathbf{u}(t)$ a diskretnú dynamiku $x_d(t), u_d(t)$

Pre popis diskretnéj dynamiky hybridného modelu kráčajúceho robota je nutné najskôr zadefinovať hybridný stavový vektor $\zeta(t)$:

$$\zeta(t) = \begin{pmatrix} \mathbf{x}(t) \\ x_d(t) \end{pmatrix} \quad (8)$$

Ďalej je nutné definovať prepínací priestor S_i .

$$S_i : s_i(\mathbf{x}, \mathbf{u}, x_d, u_d) = 0 \quad (9)$$

V modelovaní kráčajúcich robotov je nutné uvažovať prepínacie priestory v jednom konkrétnom smere z diskretnéj stavovej premennej x_d^- do diskretnéj stavovej premennej x_d^+ . Tiež je nutné uvažovať, že stavový vektor \mathbf{x} , môže vstúpiť do prepínacieho priestoru len z jedného konkrétneho smeru.

Tiež je nutné pre modelovanie diskretnéj dynamiky zadefinovať prepínaciu mapu Φ_i pre i -tu diskretnú udalosť. Prepínacia mapa Φ_i je funkcia, ktorá určuje nové hodnoty hybridného stavového vektora $\zeta(t)$ v momente po výskyte diskretnéj udalosti.

V okamihu, keď sa vyskytne nespojitost', čiže nastane diskretná udalosť a dôjde k skokovej zmene hybridného stavového vektora $\zeta(t)$, je tento vektor definovaný pomocou vhodnej prepínacej mapy:

$$\zeta^+ = \Phi_i(\mathbf{x}, \mathbf{u}, x_d, u_d, t^-) \quad (10)$$

Prepínacia mapa ako súčasť diskretnéj dynamiky kráčajúceho robota

Kráčajúci robot sa posúva dopredu so zmenou bodu, kde sa dotýka podporná (stojná) noha podložky v každom kroku, ako je zrejme z Obr. 2. Keď sa švihová noha dotkne zeme, stáva sa novou podpornou nohou so zmenou podporného bodu a z prednázarovej podpornej nohy sa stáva nová švihová noha [4].

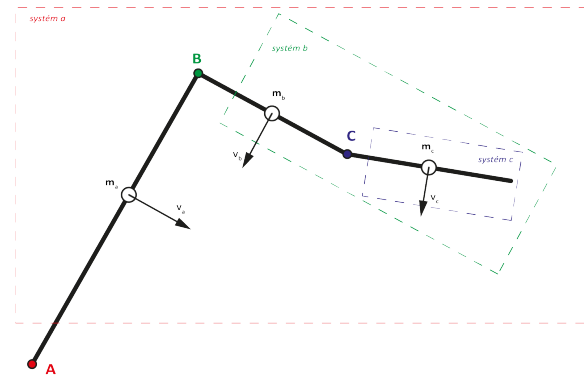
Predpokladá sa, že kolízia švihovej nohy nie je pružná a nedochádza ku sklzu čo znamená, že moment hybnosti robota \mathbf{h} je počas prechodu zachovaný $\mathbf{h}^- = \mathbf{h}^+$. Tento fakt bude využitý pri výpočte kolíznych matíc, určujúcich stavový vektor $\mathbf{x}^+(t)$ po dopade [16].

$$\mathbf{x}^+(t) = \Phi_i \mathbf{x}^-(t) \quad (11)$$

Predpokladá sa, že kolízia švihovej nohy nie je pružná a nedochádza ku sklzu čo znamená, že moment hybnosti robota \mathbf{h} je počas prechodu zachovaný $\mathbf{h}^- = \mathbf{h}^+$. Tento fakt bude využitý pri výpočte kolíznych matíc, určujúcich stavový vektor $\mathbf{x}^+(t)$ po dopade [16].

$$\mathbf{x}^+(t) = \Phi_i \mathbf{x}^-(t) \quad (12)$$

Pre modelovanie prechodovej fázy pri kráčajúcich robotov bude využitý fakt, že každý planárny robot môže byť zjednodušený do kinematického reťazca, ako na Obr. 3. Pri skúmaní, čo sa deje pri dopade švihovej nohy stačí použiť zjednodušenie čo sa deje pri kolízii v kinematickom reťazci [16].



Obr. 3. Jednoduchý kinematický reťazec pozostávajúci z 3 tuhých tyčí spojených kĺbmi [16]

Kolízie v tomto reťazci môžu byť buď interné alebo externé a sú modelované ako diskretné udalosti, pretože sa predpokladajú okamžité zmeny rýchlosti v v dôsledku kolízie tuhého telesa. Pokiaľ nárazové sily F_i nevytvárajú moment, moment hybnosti \mathbf{h} sa v tomto bode zachováva. Práve tento fakt bude využitý pri tvorbe a odvodení diskretného modelu kráčajúcich robotov [12]. Ak nastane kolízia v bode A, pôsobiaca sila F_i nevytvára žiadny krútiaci moment v tomto bode, pretože to je jediná vonkajšia sila F_i v systéme a (ako je to znázornené na Obr.3). Moment hybnosti \mathbf{h}_a pre systém a je zachovaný okolo bodu nárazu, A. Táto kolízia je ekvivalentná udalosti, keď švihová noha, ktorá je bodom nárazu, dopadne na pevnú

podložku. Okrem toho účinky kolízie sa prenášajú cez tuhé telo. Pre systém b bude izolované všetko za bodom B , jediná sila, ktorá môže byť vytvorená z impulzu pri bode A je sila v bode B . Pretože to je jediná vonkajšia sila F_i od kolízie, moment hybnosti h_b systému b sa zachováva okolo B . Podobne je zachovaný moment hybnosti h_c okolo bodu C pre systém c . Týmto spôsobom je možné získať n rovníc pre n kĺbov [16].

Aby mohli byť odvodené rovnice prechodu pre konkrétny príklad, sú jednotlivé segmenty pomenované označením obidvoch koncových bodov v dolnom indexe (napríklad A_a zodpovedá priamke vedenej z bodu A a pozície hmotnosti m_a). Lineárne rýchlosti z určitého bodu sú označené v a zodpovedajúceho indexu. Indexy, $-$ a $+$ označujú rýchlosti pred a po kolízii.

Pomocou tejto notácie je možné napísať rovnice zachovania momentu hybnosti takto:

Pre systém a :

$$\begin{aligned} m_a \vec{A}_a \times v_a^+ + m_b \vec{A}_b \times v_b^+ + m_c \vec{A}_c \times v_c^+ &= \\ = m_a \vec{A}_a \times v_a^- + m_b \vec{A}_b \times v_b^- + m_c \vec{A}_c \times v_c^- \end{aligned} \quad (13)$$

Pre systém b :

$$m_b \vec{B}_b \times v_b^+ + m_c \vec{B}_c \times v_c^+ = m_b \vec{B}_b \times v_b^- + m_c \vec{B}_c \times v_c^- \quad (14)$$

Pre systém c :

$$m_c \vec{C}_c \times v_c^+ = m_c \vec{C}_c \times v_c^- \quad (15)$$

Naopak, vnútorná kolízia, ako napríklad impulzný krútiaci moment \vec{M}_M v bode C (z blokovacieho spojenia c , napríklad, keď sa zároveň c s ramenom b), ako je to v prípade udalosti vystretí kolena pri chôdzi chodiaceho robota s kolenami. V tomto prípade, vnútorná sila F_v je nárazová sila F_i na koleno, keď sa dolná výkyvná noha otáča dopredu a zaistí sa rovná poloha nohy, čiže kĺb C bude blokovaný.

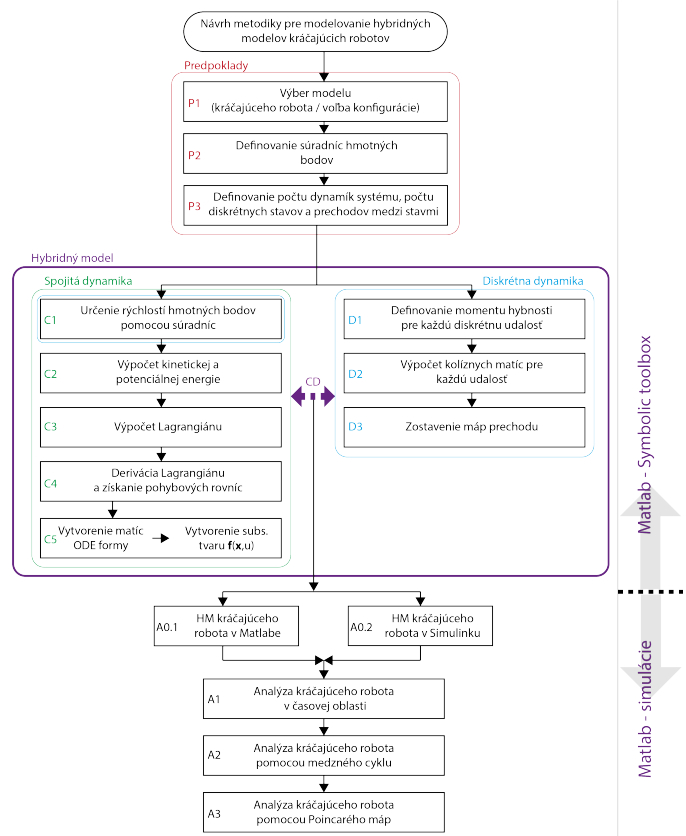
Tento krútiaci moment \vec{M}_M opäť bude jedinou silou pôsobiacou na celý systém. Pretože je to interný krútiaci moment \vec{M}_M pre systém a a systém b , je možné povedať, že pre ne je zachovaný tak lineárny, ako aj uhlový moment systému. Avšak moment hybnosti h_c systému c sa mení pôsobením krútiaceho momentu pri C .

Vytvorenie kolíznych rovníc je potrebné pre modelovanie diskretnej dynamiky kráčajúcich robotov, aby bolo možné vypočítať uhlové rýchlosti $\dot{\theta}^+$ po dopade švihovej nohy na podložku, alebo v prípade zložitejšej robotickej chôdze s kolenami, aby bolo možné modelovať zmenu uhlových rýchlostí $\dot{\theta}^+$ po vyrovnaní a zaistení kolena.

C. Návrh metodiky pre modelovanie kráčajúcich robotov

Navrhaná metodika v dizertačnej práci autora znázorňuje, že modelovania kráčajúcich robotov v rôznych konfiguráciách využíva analogické kroky. Metodika navrhovaná v rámci riešenia úloh dizertačnej práce (Obr. 4) je rozdelená na tri hlavné moduly. Prvým modulom je definovanie predpokladov. Druhý modul obsahuje tri podmoduly a to vytvorenie spojitej dynamiky/dynamík, diskretnej dynamiky a prechodov medzi

nimi. Tretí modul obsahuje vytvorenie samotných simulačných modelov a analýzu kráčajúcich robotov, čo predstavuje nutné kroky pre návrh správnych metód riadenia.

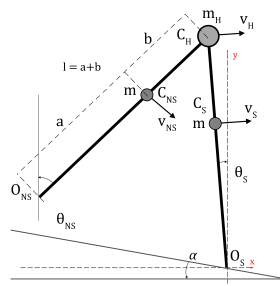


Obr. 4. Metodika pre modelovanie a riadenie kráčajúcich robotov

Prvým modulom $P1$ v navrhovanej metodike je výber modelu kráčajúceho robota. V $P1$ je potrebné definovať nie len základný náčrt kráčajúceho robota, ale aj všetky jeho rozšírenia, s ktorými je uvažovaný, napríklad: obsahuje model kolená, alebo len pevné nohy, uvažujeme chodidlá, ak áno, tak rovné alebo zakrivené, prípadne budeme uvažovať aj trup, alebo bude robot bez trupu. Na prvý krok nadväzuje druhý modul $P2$ v ktorom je nutné definovať súradnice hmotných bodov, z ktorých sa uvažovaný kráčajúci robot skladá.

Tretí modul $P3$ v module predpokladov je špecifický pre kráčajúce roboty pretože obsahuje definovanie počtu dynamík hybridného modelu kráčajúceho robota, počtu diskretných stavov a prechodov medzi nimi. Samotné odvodenie spojitej dynamiky a diskretných prechodov bude algoritmicke spracované v metodike v príslušných moduloch pre odvodenie spojitej a diskretnej dynamiky. Napríklad kráčajúci robot kružidlovej chôdze obsahuje len jednu spojitú dynamiku systému popisujúcu kráčajúceho robota a jeden prechod, kým kráčajúci robot s kolenami obsahuje dve spojité dynamiky a dva prechody medzi nimi.

Druhá hlavná časť navrhovanej metodiky je vytvorenie hybridného modelu kráčajúceho robota. Na Obr. 4 je táto časť vyznačená fialovou farbou. Najskôr popíšeme vytvorenie



Obr. 5. Model kružidlovej chôdze

Fyzikálne parametre modelu kružidlovej chôdze, ktorý bude uvažovaný sú uvedené v tabuľke I. Z tohto modelu je možné odvodiť zložitejšie kráčajúce roboty napríklad rozšírené o chodidlá, trup alebo kolena [18].

m	hmotnosť v ťažisku nohy
m_H	hmotnosť v bedrovom kĺbe
a	dĺžka od konca nohy po ťažisko
b	dĺžka od ťažiska nohy po bedrový kĺb
l	celková dĺžka nohy
α	uhol naklonenej roviny

Tab. I

PARAMETRE MODELU KRUŽIDLOVEJ CHÔDZE

spojitej dynamiky, následne vytvorenie diskkrétnej dynamiky a ako posledné popíšeme návrh prepojenia a medzi spojitou a diskrétnou dynamikou.

Prvým modulom *C1* v modelovaní spojitej dynamiky predstavuje určenie rýchlostí hmotných bodov, ktoré boli definované v module *P2*. Získanie predpisu pre rýchlosti hmotných bodov je dôležité pre modul *C2*, v ktorom je potrebné určiť kinetickú a potenciálnu energiu kráčajúceho robota. Pomocou kinetickej a potenciálnej energie je v module *C3* úlohou vypočítať tzv. Lagrangian pomocou vzťahu (3). Posledným, resp. predposledným modulom *C4* je derivácia Lagrangianu podľa súradníc, podľa rýchlostí a podľa času. Modul *C5* obsahuje algoritmické vytvorenie matíc ODE formy z pohybových rovníc a substitučného kanonického tvaru pre uvažovanú dynamiku systému.

Pre získanie diskkrétnej dynamiky je najskôr potrebné určiť momenty hybností, ktoré musia byť definované pre každú uvažovanú diskrétnu udalosť. Tento modul je na Obr. 4 označený ako *D1*. Modul *D2* predstavuje výpočet kolíznych matíc, ktoré určujú rýchlosti hmotných bodov v okamihu po kolízii. Posledným modulom *D3* je zostavenie máp prechodu, ktoré budú určovať nie len polohy ale aj rýchlosti hmotných bodov v okamihu tesne po diskkrétnej udalosti.

Modul *CD* implementuje postup vytvorenia ODE formy z vygenerovaných matíc spojitej dynamiky podľa vzťahu (7) a prepojenie spojitých dynamík s vygenerovanými diskrétnymi dynamikami kráčajúceho robota. Takto vygenerovaný hybridný model kráčajúceho robota je už vhodný pre simuláciu a následnú analýzu jeho dynamiky.

V treťom module, ktorý obsahuje navrhovaná metodika je popísaná analýza kráčajúceho robota a to analýza v časovej oblasti, analýza pomocou medzných cyklov a tiež analýza Poincarého mapou (tiež nazývanou mapa prechodu). Analýza pomocou Poincarého máp je vhodná na rýchlu analýzu stability robotického chôdze.

1) Prípadová štúdia: kráčajúci robot kružidlovej chôdze:

Kráčajúci robot kružidlovej chôdze (**Compass Gait - CG**), ktorý je znázornený na Obr. 5 je základný pasívny (neaktívny) robot, ktorý je prirodzene schopný chôdze po naklonenej rovine s dvoma stupňami voľnosti. Model kružidlovej chôdze pozostáva z dvoch tuhých nôh bez kolien a chodidiel spojených kĺbom v bedrách, ktorý v prípade aktívnej verzie chôdze obsahuje aktuátor [17].

Kráčajúci robot kružidlovej chôdze je založený na nasledovných predpokladoch:

- prechod medzi fázami chôdze je okamžitý,
- náraz švihovej nohy na zem je úplne neelastický,
- v bodoch kontaktu stojnej nohy so zemou nie je pruženie ani sklz.

Prípadová štúdia 1 overuje použitie metodiky pre modelovanie a analýzu kráčajúcich robotov na najjednoduchšom type kráčajúceho robota kružidlovej chôdze.

Prvým krokom metodiky je modul *P1*, ktorý predstavuje výber modelu. Splnenie tohto kroku predstavuje už samotné zameranie prípadovej štúdie 1. Druhým modulom *P2* je definovanie súradníc hmotných bodov a stupňov voľnosti. Model CG pozostáva z 2 stupňov voľnosti [1], ktoré podľa vzťahu (1) je možné popisovať nasledovným vektorom zovšeobecnených súradníc:

$$\theta(t) = (\theta_{NS}(t), \theta_S(t)) \tag{16}$$

kde $\theta_{NS}(t)$ predstavuje uhol natočenia švihovej nohy a $\theta_S(t)$ uhol natočenia stojnej nohy. Hmotné body C_{SS} , C_{tS} , C_H , C_{SNS} a C_{tNS} spolu s koncami nôh O_S a O_{NS} majú nasledovné súradnice:

$$\begin{aligned} C_S &= [-a \sin \theta_S \quad a \cos \theta_S] \\ C_H &= [-l \sin \theta_S \quad l \cos \theta_S] \\ C_{NS} &= [-l \sin \theta_S + b \theta_{NS} \quad l \cos \theta_S - b \cos \theta_{NS}] \\ O_S &= [0 \quad 0] \\ O_{NS} &= [-l \sin \theta_S + l \sin \theta_{NS} \quad l \cos \theta_S - l \cos \theta_{NS}] \end{aligned} \tag{17}$$

Spojité dynamiky kráčajúceho robota kružidlovej chôdze $f_{x_d}(x, u, t)$ bude pozostávať len z jednej spojitej dynamiky popísanej diskrétnou premennou $x_d = 1$. Výstupom modulu *P3* je teda popis jednej spojitej dynamiky kráčajúceho robota ako celku:

$$\dot{x}(t) = f_1(x, u, x_d, t) \tag{18}$$

Vykonaním modulu *C1* sú automaticky vygenerované rýchlosti hmotných bodov, ktoré sú získané časovou deriváciou súradníc hmotných bodov (17). Realizáciou modulu *C2* je generovaná celková kinetická energia $E_k(t)$ a potenciálna energia $E_p(t)$ kráčajúceho robota:

$$E_k(t) = \frac{1}{2}(m_H l^2 + m l^2 + m a^2) \dot{\theta}_S^2(t) + \frac{1}{2} m b^2 \dot{\theta}_{NS}^2(t) - m b l \dot{\theta}_S^2(t) \dot{\theta}_{NS}^2(t) \cos(\theta_{NS}(t) - \theta_S(t)) \quad (19)$$

$$E_p(t) = (m_H g l + m g l + m g a) \cos(\theta_S(t)) + m g b \cos(\theta_{NS}(t))$$

Z kinetickej a potenciálnej energie (19) je modulom *C3* vypočítaný Lagrangian podľa vzťahu (3) a následne sú v module *C4* generované pohybové rovnice pre jednotlivé stupne voľnosti. Nelineárna diferenciálna rovnica popisujúca dynamiku švihovej nohy má nasledovný tvar:

$$m b^2 \ddot{\theta}_{NS}(t) - m l b \cos(\theta_S(t) - \theta_{NS}(t)) \ddot{\theta}_S(t) + m l b \sin(\theta_S(t) - \theta_{NS}(t)) \dot{\theta}_S^2(t) + m g b \sin(\theta_{NS}(t)) = 0 \quad (20)$$

Nelineárna diferenciálna rovnica popisujúca správanie stojnej nohy má nasledovný tvar:

$$(m_H l^2 + m(l^2 + a^2)) \ddot{\theta}_S(t) - m l b \cos(\theta_S(t) - \theta_{NS}(t)) \ddot{\theta}_{NS}(t) - m l b \sin(\theta_S(t) - \theta_{NS}(t)) \dot{\theta}_{NS}^2(t) - (m_H l + m(l + a)) g \sin(\theta_S(t)) = 0 \quad (21)$$

Sústavu dvoch nelineárnych diferenciálnych rovníc (20) a (21) je možné upraviť do ODE formy (5), ktorej matice majú tvar:

$$\mathbf{M}(\boldsymbol{\theta}(t)) = \begin{bmatrix} m b^2 & -m l b \cos(\theta_S - \theta_{NS}) \\ -m l b \cos(\theta_S - \theta_{NS}) & (m_H l^2 + m(l^2 + a^2)) \end{bmatrix} \quad (22)$$

$$\mathbf{N}(\dot{\boldsymbol{\theta}}(t), \boldsymbol{\theta}(t)) = \begin{bmatrix} 0 & m l b \dot{\theta}_S \sin(\theta_S - \theta_{NS}) \\ -m l b \dot{\theta}_{NS} \cos(\theta_S - \theta_{NS}) & 0 \end{bmatrix} \quad (23)$$

$$\mathbf{P}(\boldsymbol{\theta}(t)) = \begin{bmatrix} m g b \sin(\theta_{NS}) \\ -(m_H l + m(l + a)) g \sin(\theta_S) \end{bmatrix} \quad (24)$$

Ak je uvažovaná aktívna podaktuovaná chôdza, matica vstupov $\mathbf{V}(t)$ ODE formy má tvar:

$$\mathbf{V}(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (25)$$

Odvodenie nelineárnych diferenciálnych pohybových rovníc a matíc ODE formy je naprogramované v simulačnom prostredí MATLAB vo funkcii `underactuated_modeling_compass()`. Spojitý stavový vektor pre model kráčajúceho robota kružidlovej chôdze je definovaný nasledovne:

$$\mathbf{x}(t) = [x_1(t) \ x_2(t) \ x_3(t) \ x_4(t)]^T = [\theta_{NS}(t) \ \theta_S(t) \ \dot{\theta}_{NS}(t) \ \dot{\theta}_S(t)]^T \quad (26)$$

Hybridný stavový vektor $\zeta(t)$ vznikne zo spojitého stavového vektora (26) pridaním diskkrétnej stavovej premennej:

$$\zeta(t) = [\theta_{NS}(t) \ \theta_S(t) \ \dot{\theta}_{NS}(t) \ \dot{\theta}_S(t) \ x_d(t)]^T \quad (27)$$

Spojité dynamiky kráčajúceho robota zahŕňa stojnú a švihovú fázu kráčajúceho robota a je možné ju zapísať nasledovne:

$$\dot{\mathbf{x}}(t) = \mathbf{f}_1(\mathbf{x}, \mathbf{u}) = \begin{pmatrix} \dot{\boldsymbol{\theta}}(t) \\ \mathbf{M}(\boldsymbol{\theta}(t))^{-1} [\mathbf{N}(\dot{\boldsymbol{\theta}}(t), \boldsymbol{\theta}(t)) \dot{\boldsymbol{\theta}}(t) + \mathbf{P}(\boldsymbol{\theta}(t)) - \mathbf{V}(t) \mathbf{u}(t)] \end{pmatrix} \quad (28)$$

Pre vygenerovanie diskkrétnej dynamiky kráčajúceho robota je podľa modulu *P3* potrebné určiť jednotlivé spojité dynamiky systému kráčajúceho robota ako celku a prechody medzi nimi. Diskrétna dynamiku modelu kráčajúceho robota obsahuje len jednu diskrétnu situáciu, keď švihová noha dopadne na podložku. Diskrétny stavový vektor $\mathbf{x}_d(t) \in 1 \subset \mathbb{Z}$ popisuje práve túto jednu udalosť. Diskrétna dynamika je popísaná podľa modulu *P3* len jedným prechodom a preto obsahuje len jednu podmienku a jednu prepínaciu mapu Φ_1 ako je ukázané na Obr. 6. Horným indexom "+" je označovaný stav v okamžiku hneď po kolízii a horným indexom "-" je označovaný stav tesne pred kolíziou.

Podmienka, kedy dochádza k výskytu diskkrétnej udalosti vyplýva z geometrie modelu a má nasledovný tvar:

$$s_1(\boldsymbol{\theta}) : \theta_{NS}(t) + \theta_S(t) + 2\alpha = 0 \quad (29)$$

Ak je podmienka (29) splnená, nastáva výmena nôh, zo švihovej nohy sa stáva stojná a zo stojná sa stáva švihová.

Podľa modulov *D1* a *D2* je nutné definovať momenty hybnosti pre každú diskrétnu udalosť a pomocou nich vypočítať vzťah uhlových rýchlostí $\dot{\boldsymbol{\theta}}(t)$ pred a po kolízii. Tieto vzťahy pred a po kolízii popisujú rovnice $\mathbf{Q}^-(\gamma)$ a $\mathbf{Q}^+(\gamma)$, ktoré sú získané riešením kinematického reťazca s uvažovaním zákona zachovania momentu hybnosti \mathbf{h} a majú nasledovný tvar:

$$\mathbf{Q}^-(\gamma) = \begin{bmatrix} -mab & -mab + (m_H l^2 + 2mal) \cos(2\gamma) \\ 0 & -mab \end{bmatrix} \quad (30)$$

$$\mathbf{Q}^+(\gamma) = \begin{bmatrix} mb(b - l \cos(2\gamma)) \\ mb^2 \\ ml(l - b \cos(2\gamma)) + ma^2 + m_H l^2 \\ -m l b \cos(2\gamma) \end{bmatrix} \quad (31)$$

kde $\gamma = \frac{\theta_{NS}(t) - \theta_S(t)}{2}$.

Diskrétna udalosť kráčajúceho robota kružidlovej chôdze je popísaná prepínacou mapou $\zeta^+ = \Phi_1(\zeta^-)$. Definovanie prepínacích máp $\Phi(\zeta^-)$ je realizované vykonaním modulu *D3*. Keďže z predpokladov pre model kráčajúceho robota

kružidlovej chôdze vyplýva, že pri dopade nenastáva sklz a je neelastický, uhly nôh pred $\{\theta_{NS}^-, \theta_S^-\}$ a po $\{\theta_{NS}^+, \theta_S^+\}$ dopade sú určené nasledovným vzťahom:

$$\theta^+ = J\theta^- \quad (32)$$

kde

$$J = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (33)$$

Matica J popisuje výmenu hodnôt uhlov jednotlivých nôh vyplývajúcu z konfigurácie robota. Uholvé rýchlosti po kolízii $\dot{\theta}_{NS}^+(t)$ a $\dot{\theta}_S^+(t)$ sú vygenerované použitím vzťahov (30) a (31) nasledovne:

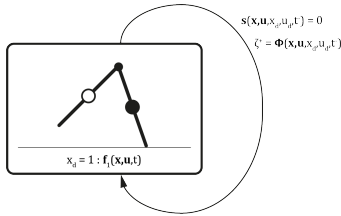
$$\dot{\theta}^+ = Q^+(\gamma)^{-1}Q^-(\gamma)\dot{\theta}^- \quad (34)$$

Spojením vzťahu (32) a (34) je možné získať prepínicu mapu Φ_1 pre spojité stavový priestor:

$$x^+(t) = \Phi_1(x^-(t)) = \begin{bmatrix} J & 0 \\ 0 & H \end{bmatrix} x^-(t) \quad (35)$$

kde

$$H^+ = Q^+(\gamma)^{-1}Q^-(\gamma) \quad (36)$$

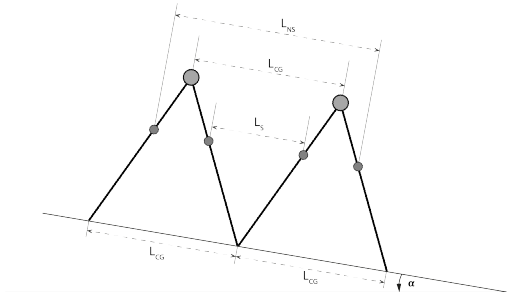


Obr. 6. Graf prechodov pre model kružidlovej chôdze

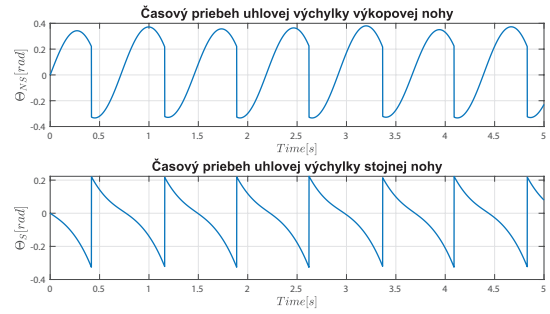
V tomto momente sú vygenerované všetky potrebné vzťahy a rovnice pre určenie výsledného hybridného modelu kráčajúceho robota. Hybridný model kráčajúceho robota kružidlovej chôdze bude mať tvar:

$$\begin{aligned} \dot{x}(t) &= f_1(x, u, t) \quad \text{ak } \zeta(t) \notin s_1 \\ \zeta^+ &= \Phi_1(x, u, x_d, u_d, t^-) \quad \text{ak } \zeta(t) \in s_1 \end{aligned} \quad (37)$$

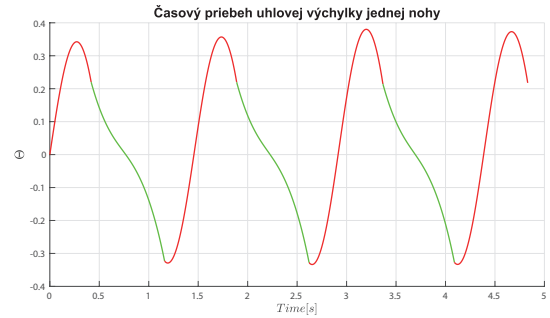
Spojité dynamiku kráčajúceho robota $f_1(x, u, x_d, t)$ vznikne úpravou matíc (22), (23) a (24) podľa vzťahu (7).



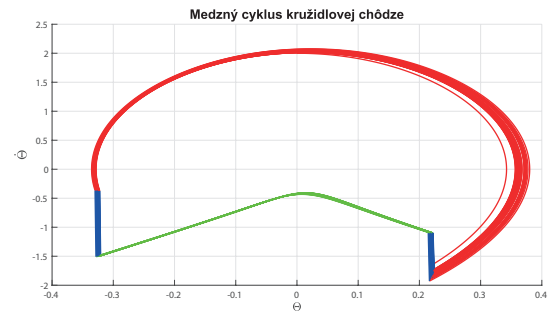
Obr. 7. Znázornenie kompletného kroku, pri ktorom ťažisko v bedrovom kĺbe klesne o $L_{CG} \sin(\alpha)$ pričom α je sklon naklonenej roviny a L_{CG} je dĺžka kroku



Obr. 8. Časové priebehy švihová nohy $\theta_{NS}(t)$ a stojnej nohy $\theta_S(t)$ modelu kružidlovej chôdze

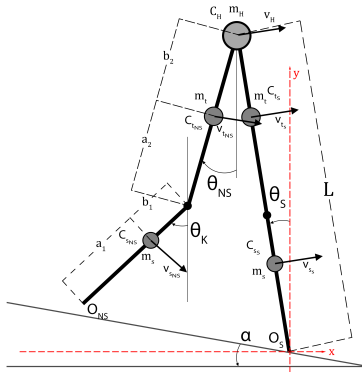


Obr. 9. Časové priebehy uholvej výchylky $\theta(t)$ z pohľadu na jednu nohu, ktorá sa mení z švihová nohy $\theta_{NS}(t)$ (červená farba) na stojnú nohu $\theta_S(t)$ (zelená farba)



Obr. 10. Zjednodušený fázový portrét typu medzný cyklus kružidlovej chôdze

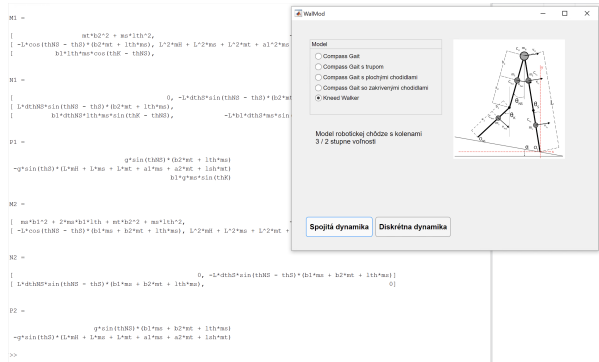
V rámci modulu $A1$ analyzujeme správanie modelu v časovej oblasti. Časové priebehy uholvých výchyliek $\theta_{NS}(t)$ a $\theta_S(t)$ sú na Obr. 8. Časový priebeh uholvých výchyliek striedajúcich sa nôh kráčajúceho robota je na Obr. 9. Podľa predpokladov, modelu kráčajúceho robota kružidlovej chôdze tvorí uzavretý medzný cyklus (Obr. 10), ktorého analýza je náplňou modulu $A2$. Medzný cyklus je získaný použitím nelineárnych diferenciálnych rovníc (20) a (21) pre spojité dynamiku a rovnice (68) pre diskretnú dynamiku. Červená časť grafu na Obr. 10 predstavuje švihová časť jednej nohy $\theta_{NS}(t)$, zelená časť zobrazuje stojnú časť $\theta_S(t)$ a modrá časť znázorňuje fázu prechodu, v ktorej sa skokovo mení stavový vektor $x(t)$ modelu kružidlovej chôdze. Pre simuláciu boli použité nasledovné fyzikálne parametre $m = 5kg, m_H = 10kg, a =$



Obr. 11. Kráčajúci robot s kolenami

$0.5m, b = 0.5m, l = a + b, g = 9.81m/s^2, \alpha = -0.1048rad$.

2) *Prípadová štúdia - modelovanie kráčajúceho robota s kolenami:* Pokrok vo výskume kráčajúcich robotov kružidlovej chôdze priniesol prirodzenú potrebu rozšírenia základného kráčajúceho robota kružidlovej chôdze. Existujú rôzne modifikácie kráčajúceho robota kružidlovej chôdze; jednou zo základných modifikácií je pridanie kolien k modelu kráčajúceho robota kružidlovej chôdze, čím vznikne jedno zo základných pripodobnení ľudskej chôdzi.



Obr. 12. Ukážka vygenerovaných pohybových rovníc v ODE forme pre kráčajúceho robota Kneed Walkera s využitím navrhutej metodiky modelovania

Kračajúci robot s kolenami je modelovaný ako pasívny, rovnako ako základný kráčajúci robot kružidlovej chôdze. Ako je zobrazené na Obr. 11, z pohľadu modelovania chôdze s kolenami je švihová noha modelovaná na začiatku kroku ako systém dvoch kyvadiel a keď dôjde k vyrovnaniu kolena, kolenný kĺb sa zablokuje a počas ďalšej fázy kroku sa švihová noha modeluje ako jednoduché kyvadlo [16]. Stojná noha je stále modelovaná ako jednoduché kyvadlo, keďže stojná noha má blokováný kolenný kĺb [19]. Kráčajúci robot s kolenami obsahuje tri, resp. dva stupne voľnosti. Uvedené predpoklady sú zohľadnené v moduloch P1 až P3 metodiky a prípadová štúdia 2 overí správnosť týchto predpokladov.

Modelovanie kráčajúceho robota s kolenami prináša zvýšenie zložitosti v podobe pridania ďalšieho stupňa voľnosti

a tiež pridania ďalšej spojitej dynamiky, ktorú je potrebné modelovať.

Prvý predpokladový modul P1 je realizovaný zvolením kráčajúceho robota s kolenami. Ďalším modulom predpokladov P2 pre modelovanie, je definovanie stupňov voľnosti a súradníc hmotných bodov. Model kráčajúceho robota s kolenami pozostáva z 3 stupňov voľnosti [1], ktoré podľa vzťahu (1) je možné popísať nasledovným vektorom zovšeobecnených súradníc:

$$\theta(t) = (\theta_{NS}(t), \theta_S(t), \theta_K(t)) \quad (38)$$

kde $\theta_{NS}(t)$ predstavuje uhol natočenia švihovej nohy v bedrovom kĺbe, $\theta_S(t)$ je uhol natočenia stojnej nohy a $\theta_K(t)$ je uhol natočenia kolena švihovej nohy. Poradie bolo zvolené s ohľadom na zachovanie čo najväčšej podobnosti s modelom základného kráčajúceho robota kružidlovej chôdze.

Hmotné body C_S, C_H a C_{NS} spolu s koncami nôh O_S a O_{NS} majú nasledovné súradnice:

$$\begin{aligned} C_{SS} &= [-a_1 \sin \theta_S \quad a_1 \cos \theta_S] \\ Ct_S &= [(l_s + a_2) \sin \theta_S \quad (l_s + a_2) \cos \theta_S] \\ C_H &= [-L \sin \theta_S \quad L \cos \theta_S] \\ C_{SNS} &= [L \sin \theta_S - l_t \sin \theta_{NS} - b_1 \sin \theta_K \\ &\quad L \cos \theta_S - l_t \cos \theta_{NS} - b_1 \cos \theta_K] \\ Ct_{NS} &= [L \sin \theta_S - b_2 \sin \theta_{NS} \quad L \cos \theta_S - b_2 \cos \theta_{NS}] \\ O_S &= [0 \quad 0] \\ O_{NS} &= [L \sin \theta_S - l_t \sin \theta_{NS} - l_s \sin \theta_K \\ &\quad L \cos \theta_S - l_t \cos \theta_{NS} - l_s \cos \theta_K] \end{aligned} \quad (39)$$

Nasleduje realizácia modulu P3, ktorý definuje počet jednotlivých spojitých dynamík a diskkrétne prechody medzi nimi. Keďže kráčajúci robot s kolenami obsahuje dve rozdielne spojité dynamiky systému, je nutné určiť nelineárne diferenciálne pohybové rovnice pre obidva z nich. V tejto prípadovej štúdiu bude najskôr uvedené vygenerovanie dynamiky trojitého kyvadla a následne dynamiky dvojitého kyvadla. Metodika samotného odvodenia nelineárnych diferenciálnych pohybových rovníc je rovnaká ako v prípade kráčajúceho robota kružidlovej chôdze. Odvodenie vyúsťujúce do automatického vygenerovania hybridného modelu kráčajúceho robota je uvedené z dôvodu overenia metodiky aj pre zložitejšie kráčajúce roboty. Odvodenie diskkrétnej dynamiky a zostavenie máp prechodov bude uvedené v časti o diskkrétnej dynamike po odvodení spojitých dynamík kráčajúceho robota.

Spojité dynamiky kráčajúceho robota s kolenami $f_{x_d}(x, u, t)$ pozostáva z dynamiky s uvoľneným kolonom $x_d = 1$ a z dynamiky so zaisteným kolonom $x_d = 2$:

$$\dot{x}(t) = f(x, u, x_d, t) = \begin{cases} f_1(x, u, t) & \text{pre } x_d = 1 \\ f_2(x, u, t) & \text{pre } x_d = 2 \end{cases} \quad (40)$$

Spojité dynamiky $f_1(x, u, t)$ a $f_2(x, u, t)$ sa skladajú z dynamík odpovedajúcich jednotlivým stupňom voľnosti.

Dynamika trojitého kyvadla kráčajúceho robota s kolenami

Podľa metodiky je ďalším krokom modul *C1*, teda určenie rýchlosti hmotných bodov. Následne realizáciou modulu *C2* je automatické odvodenie kinetickej a potenciálnej energie celého robota.

Z kinetickej a potenciálnej energie kráčajúceho robota a je nutné vypočítať Lagrangián podľa vzťahu (3), čo je realizované modulom *C3*. Z Lagrangiánu sú určené nelineárne diferenciálne pohybové rovnice pre jednotlivé stupne voľnosti. Matice ODE formy pre trojlinkovú konfiguráciu kráčajúceho robota sú generované funkciou *KneedWalkerModeling()*. Výstup modulu *C4* a modulu *C5* je nasledovný:

$$\mathbf{M}(\boldsymbol{\theta}(t)) = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \quad (41)$$

$$\begin{aligned} M_{11} &= m_t b_2^2 + m_s l_t^2 m_s b_1 l_t \cos(\theta_K(t) - \theta_{NS}(t)) \\ M_{12} &= -(m_t b_2 + m_s l_t) L \cos(\theta_{NS}(t) - \theta_S(t)) \\ M_{13} &= m_s b_1 l_t \cos(\theta_K(t) - \theta_{NS}(t)) \\ M_{21} &= -(m_t b_2 + m_s l_t) L \cos(\theta_{NS}(t) - \theta_S(t)) \\ M_{22} &= m_s a_1^2 + m_t (l_s + a_2)^2 + (m_H + m_t + m_s) L^2 \\ M_{23} &= -m_s b_1 L \cos(\theta_K(t) - \theta_S(t)) \\ M_{31} &= m_s b_1 l_t \cos(\theta_K(t) - \theta_{NS}(t)) \\ M_{32} &= -m_s b_1 L \cos(\theta_K(t) - \theta_S(t)) \\ M_{33} &= m_s b_1^2 \end{aligned}$$

$$\mathbf{N}(\dot{\boldsymbol{\theta}}(t), \boldsymbol{\theta}(t)) = \begin{bmatrix} 0 & N_{12} \dot{\theta}_S(t) & N_{13} \dot{\theta}_K(t) \\ N_{21} \dot{\theta}_{NS}(t) & 0 & N_{23} \dot{\theta}_S(t) \\ N_{31} \dot{\theta}_{NS}(t) & N_{32} \dot{\theta}_S(t) & 0 \end{bmatrix} \quad (42)$$

$$\begin{aligned} N_{12} &= -(m_t b_2 + m_s l_t) L \sin(\theta_{NS}(t) - \theta_S(t)) \\ N_{13} &= -m_s b_1 l_t \sin(\theta_K(t) - \theta_{NS}(t)) \\ N_{21} &= -N_{12} \\ N_{23} &= m_s b_1 L \sin(\theta_K(t) - \theta_S(t)) \\ N_{31} &= -N_{13} \\ N_{32} &= -N_{23} \end{aligned}$$

$$\mathbf{P}(\boldsymbol{\theta}(t)) = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} \quad (43)$$

$$\begin{aligned} P_1 &= (m_t b_2 + m_s l_t) g \sin(\theta_{NS}(t)) \\ P_2 &= -[m_s a_1 + m_t (l_s + a_2) + \\ &\quad + (m_H + m_t + m_s) L] g \sin(\theta_S(t)) \\ P_3 &= m_s b_1 g \sin(\theta_K(t)) \end{aligned}$$

Dynamika dvojitého kyvadla kráčajúceho robota s kolenami

Keďže kráčajúci robot s kolenami obsahuje dve dynamiky, celý proces odvodenia spojitých dynamik je nutné zopakovať

aj pre druhú dynamiku $f_2(\mathbf{x}, u, t)$. Dynamiku švihovej nohy v druhej polovici švihovej fázy je nutné modelovať ako jednoduché kyvadlo. Rovnako aj stojná noha je modelovaná so zaisteným kolenom.

Z kinetickej energie a potenciálnej energie je v module automaticky vypočítaný *C3* Lagrangián pre dynamiku f_2 . Z Lagrangiánu sú odvodené pohybové rovnice kráčajúceho robota (modul *C4*), ktoré sú upravené do matíc ODE formy v module *C5*. Tieto matice ODE formy majú tvar:

$$\mathbf{M}(\boldsymbol{\theta}(t)) = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \quad (44)$$

$$\begin{aligned} M_{11} &= m_t b_2^2 + m_s (l_t + b_1)^2 \\ M_{12} &= -(m_t b_2 + m_s (l_t + b_1)) L \cos(\theta_{NS}(t) - \theta_S(t)) \\ M_{21} &= -(m_t b_2 + m_s (l_t + b_1)) L \cos(\theta_{NS}(t) - \theta_S(t)) \\ M_{22} &= m_s a_1^2 + m_t (l_s + a_2)^2 + (m_H + m_t + m_s) L^2 \end{aligned}$$

$$\mathbf{N}(\dot{\boldsymbol{\theta}}(t), \boldsymbol{\theta}(t)) = \begin{bmatrix} 0 & N_{12} \dot{\theta}_S(t) \\ N_{21} \dot{\theta}_{NS}(t) & 0 \end{bmatrix} \quad (45)$$

$$\begin{aligned} N_{12} &= -(m_t b_2 + m_s (l_t + b_1)) L \sin(\theta_{NS}(t) - \theta_S(t)) \\ N_{21} &= (m_t b_2 + m_s (l_t + b_1)) L \sin(\theta_{NS}(t) - \theta_S(t)) \end{aligned}$$

$$\mathbf{P}(\boldsymbol{\theta}(t)) = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \quad (46)$$

$$\begin{aligned} P_1 &= (m_t b_2 + m_s (l_t + b_1)) g \sin(\theta_{NS}(t)) \\ P_2 &= -[m_s a_1 + m_t (l_s + a_2) + \\ &\quad + (m_H + m_t + m_s) L] g \sin(\theta_S(t)) \end{aligned}$$

Spojitéy stavový vektor $\mathbf{x}(t)$ pre kráčajúceho robota s kolenami pre dynamiku f_2 je zadaný nasledovne:

$$\mathbf{x}(t) = [\theta_{NS}(t) \quad \theta_S(t) \quad \theta_{NS}(t) \quad \dot{\theta}_{NS}(t) \quad \dot{\theta}_S(t) \quad \dot{\theta}_{NS}(t)]^T \quad (47)$$

Hybridný stavový vektor $\boldsymbol{\zeta}(t)$ vznikne zo spojitých stavového vektora pridaním diskretných stavov premennej:

$$\boldsymbol{\zeta}(t) = \begin{bmatrix} \theta_{NS}(t) & \theta_S(t) & \theta_{NS}(t) \\ \dot{\theta}_{NS}(t) & \dot{\theta}_S(t) & \dot{\theta}_{NS}(t) \\ x_d(t) \end{bmatrix}^T \quad (48)$$

Modelovanie diskretných dynamik kráčajúceho robota s kolenami

Po odvodení spojitých dynamik kráčajúceho robota je nutné určiť vzťahy, ktorými sa riadi diskretná dynamika, teda podmienky prechodov medzi spojitými dynamikami.

Kračajúci robot s kolenami obsahuje okrem kolízie, ktorá nastane pri dopade švihovej nohy na podložku aj druhú kolíziu, ktorá nastáva v prípade vyrovnania a teda okamžitého uzamknutia kolenného kĺbu. Z tohto dôvodu prípadová štúdia 2 overuje automatické vygenerovanie diskretných dynamik pre zložitejší hybridný model.

Predpoklady z modulu $P3$, ktoré sa uplatňujú pri modelovaní kolízií sú rovnaké, ako v prípade kráčajúceho robota jednoduchej kružidlovej chôdze. Ako prvú uvedieme kolíziu a jej predpoklady pre vyrovnanie kolena.

Vyrovnanie a blokovanie kolena

Vyrovnanie kolena nastáva ak je splnená podmienka, že uhol švihovej nohy θ_{NS} a uhol kolena θ_K sa rovnajú, čiže:

$$s_{1,2} : \theta_{NS}(t) - \theta_K(t) = 0 \quad (49)$$

Ak sú splnené podmienky, teda kráčajúci robot sa správa podľa dynamiky f_1 ($x_d = 1$) a stavový vektor vstúpi do prepínacieho priestoru, nastane diskretná udalosť a kráčajúci robot sa prepne z dynamiky $x_d = 1$ do dynamiky $x_d = 2$.

Okrem definovania zákona zachovania momentu hybnosti, v module $D1$ určíme aj zmenu natočení jednotlivých uhlov. Zmenu natočenia jednotlivých uhlov kráčajúceho robota s kolenami definujeme nasledovne:

$$\begin{bmatrix} \theta_{NS}^+ \\ \theta_S^+ \\ \theta_K^+ \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{NS}^- \\ \theta_S^- \\ \theta_K^- \end{bmatrix} \quad (50)$$

čo môžeme v maticovom tvare interpretovať aj ako:

$$\theta^+ = J_{1,2} \theta^- \quad (51)$$

Vzťahy popisujúce uhlové rýchlosti jednotlivých uhlov v momente pred a po nastaní diskretné udalosti je možné získať riešením kinematického reťazca, ktorý uvažuje platnosť zákona zachovania momentu hybnosti. V rámci modelovania diskretné dynamiky je potrebné v module $D1$ definovať zákon zachovania momentov hybnosti.

V ďalšom module $D2$ modelovania diskretné dynamiky sú z rovníc zachovania momentu hybnosti určené matice $Q^-(\theta(t))$ a $Q^+(\theta(t))$, ktoré určujú vzťah uhlových rýchlostí jednotlivých kĺbov v momente diskretné udalosti.

$$Q^-(\gamma) = \begin{bmatrix} Q_{11}^- & Q_{12}^- & Q_{13}^- \\ Q_{21}^- & Q_{22}^- & Q_{23}^- \end{bmatrix} \quad (52)$$

$$\begin{aligned} Q_{11}^- &= -m_s b_1 l_t \cos(\theta_{NS} - \theta_K) - b_1^2 m_s \\ Q_{12}^- &= (m_s l_t + m_t b_2) L \cos(\theta_S - \theta_{NS}) + \\ &\quad + m_s b_1 L \cos(\theta_S - \theta_K) \\ Q_{21}^- &= -m_s b_1 l_t \cos(\theta_{NS} - \theta_K) - m_t b_2^2 - m_s l_t^2 \\ Q_{21}^- &= (m_s l_t + m_t b_2) L \cos(\theta_S - \theta_{NS}) - \\ &\quad - m_s b_1 l_t \cos(\theta_{NS} - \theta_K) - m_t b_2^2 - m_s l_t^2 \\ Q_{22}^- &= (m_s l_t + m_t b_2) L \cos(\theta_S - \theta_{NS}) + \\ &\quad + m_s b_1 L \cos(\theta_S - \theta_K) - (m_t + m_s + m_H) L^2 \\ &\quad - m_s a_1^2 - m_t (l_s + a_2)^2 \\ Q_{23}^- &= -m_s (b_1 \cos(\theta_K) (b_1 \cos(\theta_K) - L \cos(\theta_S) + l_t \cos(\theta_{NS})) \\ &\quad + b_1 \sin(\theta_K) (b_1 \sin(\theta_K) - L \sin(\theta_S) + l_t \sin(\theta_{NS}))) \end{aligned}$$

$$Q^-(\gamma) = \begin{bmatrix} Q_{11}^+ & Q_{12}^+ \\ Q_{21}^+ & Q_{22}^+ \end{bmatrix} \quad (53)$$

$$\begin{aligned} Q_{11}^+ &= -m_s (l_t + b_1)^2 - m_t b_2^2 \\ Q_{12}^+ &= (m_s (b_1 + l_t) + m_t b_2) L \cos(\theta_S - \theta_{NS}) \\ Q_{21}^+ &= (m_s (b_1 + l_t) + m_t b_2) L \cos(\theta_S - \theta_{NS}) - \\ &\quad - m_s (l_t + b_1)^2 - m_t b_2^2 \\ Q_{22}^+ &= (m_s (b_1 + l_t) + m_t b_2) L \cos(\theta_S - \theta_{NS}) - \\ &\quad - (m_t + m_s + m_H) L^2 - m_s a_1^2 - m_t (l_s + a_2)^2 \end{aligned}$$

Vykonaním modulu $D3$ - výsledná prepínacia mapa Φ_1 pre túto diskretnú udalosť bude mať tvar:

$$x^+(t) = \Phi_1(x^-(t)) = \begin{bmatrix} J & 0 \\ 0 & H \end{bmatrix} x^-(t) \quad (54)$$

kde

$$H^+ = Q^{+1} Q^- \quad (55)$$

Dopad švihovej nohy na podložku

Pri dopade švihovej nohy na podložku dochádza ku druhej diskretné udalosti a mení sa dynamika z f_2 na dynamiku f_1 . Pri dopade švihovej nohy na podložku dochádza k zmene dynamiky z dvojitého kyvadla na trojité kyvadlo. Táto diskretná udalosť nastáva, keď je splnená podmienka:

$$s_{2,1} : \theta_{NS}(t) + \theta_S(t) + 2 * \alpha = 0 \quad (56)$$

Zmena uhlových výchyliek pri tejto diskretné udalosti je nasledovná:

$$\begin{bmatrix} \theta_{NS}^+ \\ \theta_S^+ \\ \theta_K^+ \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \theta_{NS}^- \\ \theta_S^- \\ \theta_K^- \end{bmatrix} \quad (57)$$

čo je možné v maticovom tvare interpretovať aj ako:

$$\theta^+ = J_{2,1} \theta^- \quad (58)$$

Znova je nutné vykonať modul $D1$ a definovať zákon zachovania momentu hybnosti, v momente dopadu švihovej nohy, vzhľadom na koniec švihovej nohy P pred kolíziou a bod O po kolízii.

Riešením rovníc zachovania momentu hybnosti sa získajú matice kolízií $Q^-(\theta(t))$ a $Q^+(\theta(t))$, ktoré určujú pomer uhlových rýchlostí jednotlivých kĺbov kráčajúceho robota v momente tesne pred dopadom švihovej nohy na podložku a tesne po dopade. Vykonaním modulu $D2$ sú získané vzťahy $Q^-(\theta(t))$ a $Q^+(\theta(t))$:

$$Q^- = \begin{bmatrix} Q_{11}^- & Q_{12}^- \\ Q_{21}^- & Q_{22}^- \end{bmatrix} \quad (59)$$

$$\begin{aligned} Q_{11}^- &= 0 \\ Q_{12}^- &= m_s a_1 (a_2 + b_1 + b_2) + m_t b_2 (a_1 + a_2 + b_1) \\ Q_{21}^- &= m_s a_1 (a_2 + b_1 + b_2) + m_t b_2 (a_1 + a_2 + b_1) \\ Q_{22}^- &= -m_s a_1^2 + L (m_s a_1 + m_t a_2 + m_t l_s) - \\ &\quad - m_t (a_2 + l_s)^2 + (-m_H L - 2m_t (a_1 + a_2 + b_1)) - \\ &\quad - 2m_s a_1 L \cos(\theta_S - \theta_{NS}) \end{aligned}$$

$$\mathbf{Q}^- = \begin{bmatrix} Q_{11}^+ & Q_{12}^+ \\ Q_{21}^+ & Q_{22}^+ \end{bmatrix} \quad (60)$$

$$\begin{aligned} Q_{11}^+ &= -m_s(l_t + b_1)^2 - m_t b_2^2 \\ Q_{12}^+ &= (m_s(b_1 + l_t) + m_t b_2)L \cos(\theta_S - \theta_{NS}) \\ Q_{21}^+ &= (m_s(b_1 + l_t) + m_t b_2)L \cos(\theta_S - \theta_{NS}) - \\ &\quad - m_s(l_t + b_1)^2 - m_t b_2^2 \\ Q_{22}^+ &= (m_s(b_1 + l_t) + m_t b_2)L \cos(\theta_S - \theta_{NS}) - \\ &\quad - (m_t + m_s + m_H)L^2 - m_s a_1^2 - m_t(l_s + a_2)^2 \end{aligned}$$

Matice (59) a (60) sú rozmeru 2×2 , pretože kolenný kĺb v momente tesne po dopade švihovej nohy na podložku sa odblokuje a má rovnakú uhlovú výchylku aj rýchlosť ako uhol $\theta_{NS}(t^+)$, čiže $\theta_K(t^+) = \theta_{NS}(t^+)$ a $\dot{\theta}_K(t^+) = \dot{\theta}_{NS}(t^+)$.

Keďže je potrebné určiť uhlové rýchlosti všetkých uvažovaných uhlov kráčajúceho robota a matice (59) a (60) sú rozmeru len 2×2 , je nutné doplniť maticu $(\mathbf{Q}^-(\theta))^{-1} \mathbf{Q}^-(\theta)$ o nulový stĺpec:

$$\begin{bmatrix} \dot{\theta}_{NS}^+ \\ \dot{\theta}_S^+ \end{bmatrix} = \begin{bmatrix} (\mathbf{Q}^-(\theta))^{-1} \mathbf{Q}^-(\theta) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta}_{NS}^- \\ \dot{\theta}_S^- \\ \dot{\theta}_K^- \end{bmatrix} \quad (61)$$

Opäť je potrebné realizovať modul $D3$, keďže pre každú diskretnú udalosť je postup odvodenia rovnaký. Výsledná prepínacia mapa Φ_1 pre túto diskretnú udalosť bude mať tvar:

$$\mathbf{x}^+(t) = \Phi_2(\mathbf{x}^-(t)) = \begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \mathbf{x}^-(t) \quad (62)$$

kde

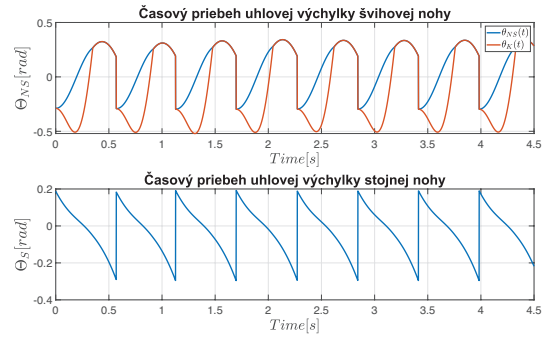
$$\mathbf{H}^+ = \mathbf{Q}^{+1} \mathbf{Q}^- \quad (63)$$

Výsledný vygenerovaný hybridný matematický model kráčajúceho robota s kolenami bude mať nasledovný tvar:

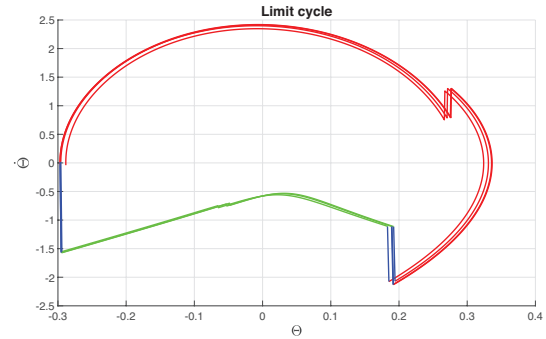
$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad \text{ak} \quad \zeta(t) \notin s_i \\ \zeta^+ &= \Phi(\mathbf{x}, \mathbf{u}, x_d, u_d, t^-) \quad \text{ak} \quad \zeta(t) \in s_i \end{aligned} \quad (64)$$

Riešenie modulu $A1$ analýzy kráčajúcich robotov poskytuje pohľad na časové priebehy natočenia nôh. Časový priebeh uhlových výchyliek $\theta_{NS}(t)$, $\theta_S(t)$ a $\theta_K(t)$ spolu s aplikovaním diskretných rovníc k spojitej dynamike je zobrazený na Obr. 13.

V module $A2$ je možné získať analýzu kráčajúceho robota vo fázovej rovine. Predpokladáme, že kráčajúci robot tvorí uzavretý medzný cyklus. Obr. 14 znázorňuje medzný cyklus kráčajúceho robota s kolenami. Uhol natočenia predkolenia nie je na fázovom portréte zobrazený. Realizáciou modulu $A3$ získame Poincarého mapu pre medzný cyklus, ktorý vytvára kráčajúci robot s kolenami. Poincarého mapa uvedená v dizertačnej práci dokazuje, že medzný cyklus vytváraný kráčajúcim robotom s kolenami je asymptoticky stabilný.



Obr. 13. časový priebeh uhlových výchyliek kráčajúceho robota s kolenami



Obr. 14. Zjednodušenie medzného cyklu do 2-rozmerného priestoru pre kráčajúceho robota s kolenami

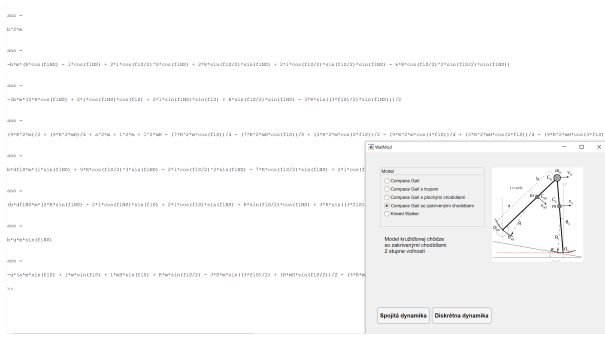
3) *Prípadová štúdia 3:* Jedným z prvých možných rozšírení základného modelu kružidlovej chôdze je pridanie zakrivených chodidiel. Zakrivené chodidlá nie sú plne biologicky inšpirované, pretože len ťažko nájdeme v živej prírode živočíchov so zakrivenými chodidlami pevne upevnenými na konce nôh. Pridanie zakrivených chodidiel je viac technická záležitosť, pretože pri tvorbe laboratórných modelov je ľahšie vytvoriť kráčajúceho robota so zakrivenými chodidlami [20].

Oproti prípadovej štúdii 1, základný kráčajúci robot kružidlovej chôdze, pridáva táto prípadová štúdia zakrivené chodidlá, ktoré majú overiť použitie metodiky aj pre zložitejšie kráčajúce roboty, ktoré modifikujú kružidlovú chôdzu. Kráčajúci robot so zakrivenými chodidlami obsahuje podobne ako kráčajúci robot kružidlovej chôdze dva stupne voľnosti.

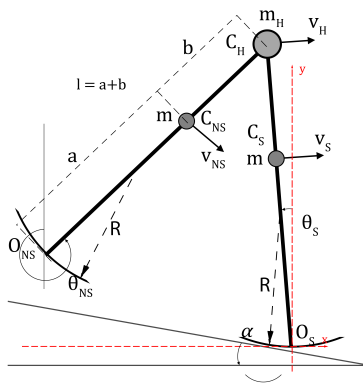
Spojité dynamiky kráčajúceho robota so zakrivenými chodidlami $\mathbf{f}_{x_d}(\mathbf{x}, \mathbf{u}, t)$ bude pozostávať len z jednej spojitely dynamiky popísanej diskretnou premennou $x_d = 1$:

$$\dot{\mathbf{x}}(t) = \mathbf{f}_1(\mathbf{x}, \mathbf{u}, x_d, t) \quad (65)$$

Podľa metodiky predstavenej je nutné realizovať modul $P1$ a určiť typ kráčajúceho robota, v tomto prípade je to kráčajúci robot so zakrivenými chodidlami. Ďalším predpokladovým modulom je modul $P2$, riešením ktorého získame súradnice hmotných bodov, dôležité pre získanie spojitely aj diskretnely dynamiky. Pre kráčajúceho robota kružidlovej chôdze so zakrivenými chodidlami, ktorý je znázornený na Obr. 16 sú



Obr. 15. Ukážka vygenerovaných pohybových rovníc v ODE forme pre kráčajúceho robota so zakrivenými chodidlami



Obr. 16. Kráčajúci robot s chodidlami typu kružidlovej chôdze

do metodiky ako vstup vložené súradnice hmotných bodov a koncov nôh nasledovné:

$$\begin{aligned}
 C_{NS} &= [-l \sin(\theta_S) - R\theta_S + b \sin(\theta_{NS}) \\
 &\quad l \cos(\theta_S) + R \sin((\pi - \theta_S)/2) \sin(\theta_S)^2 - b \cos(\theta_{NS})] \\
 C_S &= [-a \sin(\theta_S) - R\theta_S a \cos(\theta_S) + R \sin((\pi - \theta_S)/2) \sin(\theta_S)^2 \\
 &\quad a \cos(\theta_S)] \\
 C_H &= [-l \sin(\theta_S) - R\theta_S l \cos(\theta_S) + R \sin((\pi - \theta_S)/2) \sin(\theta_S)^2 \\
 &\quad l \cos(\theta_S)] \\
 O_{NS} &= [-l \sin(\theta_S) - R\theta_S + l \sin(\theta_{NS}) \\
 &\quad l \cos(\theta_S) + R \sin((\pi - \theta_S)/2) \sin(\theta_S)^2 - l \cos(\theta_{NS})] \\
 O_S &= [0 \ 0]
 \end{aligned} \tag{66}$$

Ďalším krokom metodiky je modul P3, teda definovanie počtu dynamík kráčajúceho robota a prechodov medzi nimi. Z konfigurácie robota vyplýva, že obsahuje iba jednu spojitú dynamiku a jednu diskretnú udalosť rovnako ako kráčajúci robot kružidlovej chôdze bez chodidiel a bez kolien.

Spojité dynamiky kráčajúceho robota so zakrivenými chodidlami

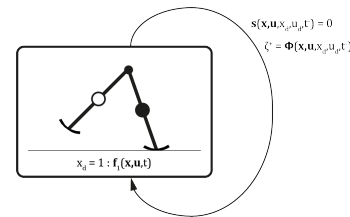
Ďalej pokračujeme modulmi modelovania spojitých dynamík C1 až C5. Prvým modulom modelovania spojitých dynamík je modul C1, v ktorom je nutné derivovať súradnice hmotných bodov a určiť rýchlosti hmotných bodov. S využitím vzťahov 66 je možné prejsť do modulu C2 a získať celkovú kinetickú energiu E_K a celkovú potenciálnu energiu E_P . Kvôli značnej zložitosti vzťahov pre kinetickú a potenciálnu en-

ergiu ich neuvádzame v tejto dizertačnej práci. Z týchto dvoch energií popisujúcich systém je ďalej nutné prejsť do modulu C3 a získať Lagrangián, ktorý je rozdielom kinetickej energie E_K a potenciálnej energie E_P . z Lagrangiánu následnými deriváciami získame pohybové rovnice (modul C4). Kvôli značnej zložitosti nelineárnych pohybových rovníc ich neuvádzame, uvádzame len výstup softvérového nástroja *WalMod*, vytvoreného v rámci riešenia dizertačnej práce, ktorý je na Obr. 15.

Automatické generovanie matice ODE formy pre kráčajúceho robota so zakrivenými chodidlami je naprogramované v simulačnom prostredí MATLAB vo funkcii *underactuated_modeling_feet()*.

Diskrétna dynamika kráčajúceho robota so zakrivenými chodidlami

Ďalším súborom krokov v modelovaní kráčajúcich robotov je podľa navrhutej metodiky modelovanie diskretnej dynamiky kráčajúceho robota, ktoré je zastrešené modulmi D1 až D3. Keďže kráčajúci robot so zakrivenými chodidlami vychádza zo základného robota kružidlovej chôdze, obsahuje len jednu diskretnú udalosť a to dopad švihovej nohy na podložku.



Obr. 17. Prechodový diagram kráčajúceho robota so zakrivenými chodidlami

Pri chôdzi kráčajúceho robota so zakrivenými chodidlami sa aplikuje jedna prechodová podmienka a to v momente dotyku švihovej nohy s podložkou. Táto diskretná udalosť je rovnaká ako v prípade systému kružidlovej chôdze bez chodidiel, pričom podmienka vyplýva z geometrie modelu:

$$s(\theta) : \theta_{NS}(t) + \theta_S(t) + 2\alpha = 0 \tag{67}$$

Postup v realizovaný v moduloch D1 a D2 je rovnaký, ako v prípade kráčajúceho robota kružidlovej chôdze, rozdielom sú len iné predpoklady a rozdielne súradnice hmotných bodov. Prepínacia mapa $\zeta^+ = \Phi(\zeta^-)$ popisujúca dopad švihovej nohy na podložku je nasledovná:

$$x^+(t) = \Phi_1(x^-(t)) = \begin{bmatrix} J & 0 \\ 0 & H \end{bmatrix} x^-(t) \tag{68}$$

kde

$$H^+ = Q^+(\gamma)^{-1} Q^-(\gamma) \tag{69}$$

Význam jednotlivých matíc je popísaný v prípadovej štúdií modelovania jednoduchého kráčajúceho robota kružidlovej chôdze, jednotlivé prvky matíc Q^+ a Q^- je možné získať použitím nástroja *WalMod*.

Tabuľka II znázorňuje význam a numerické hodnoty fyzikálnych parametrov potrebných pre simuláciu robota

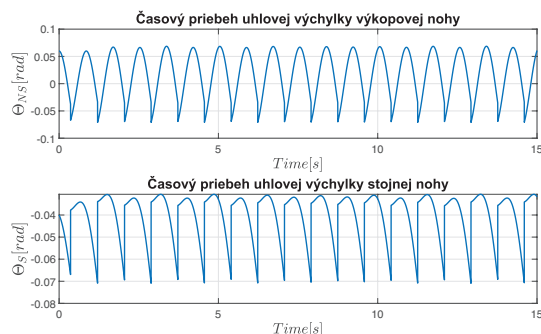
krúžidlovej chôdze so zakrivenými chodidlami. Simulácie boli vykonané s nasledovnými počiatočnými podmienkami: $\theta_{NS}(0) = 0.06$, $\theta_S(0) = -0.04$, $\dot{\theta}_{NS}(0) = -0.014$, $\dot{\theta}_S(0) = -0.03$

a	dĺžka od konca nohy po ťažisko	0.5m
b	dĺžka od ťažiska nohy po jej koniec	0.5m
R	polomer zakrivených nôh	0.3m
m	hmotnosť nohy	5kg
m_H	hmotnosť v bedrovom kĺbe	10kg
α	sklon podložky	-0.1048rad
l	$l = a + b$	1m

Tab. II

PARAMETRE MODELU KRÚŽIDLOVEJ CHÔDZE SO ZAKRIVENÝMI CHODIDLAMI

Modul analýzy A1 umožňuje získať analýzu v časovej oblasti. Obrázok 18 znázorňuje časový priebeh uhlových výchyliek θ_{NS} a θ_S . Napriek takmer rovnakému modelu existujú rozdiely, ktoré sú ovplyvnené použitím chodidiel. Najväčším rozdielom je zmenšenie amplitúd uhlových výchyliek $\theta_{NS}(t)$ a $\theta_S(t)$ pre obidve nohy.

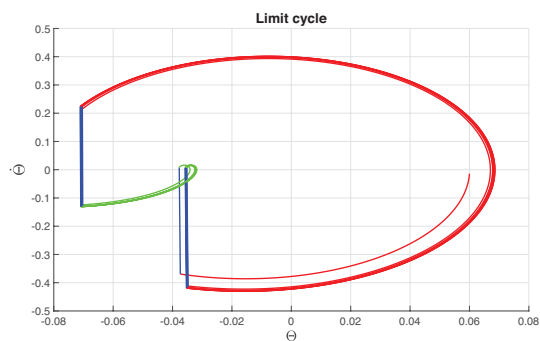


Obr. 18. Časový priebeh uhlových výchyliek kráčajúceho robota so zakrivenými chodidlami

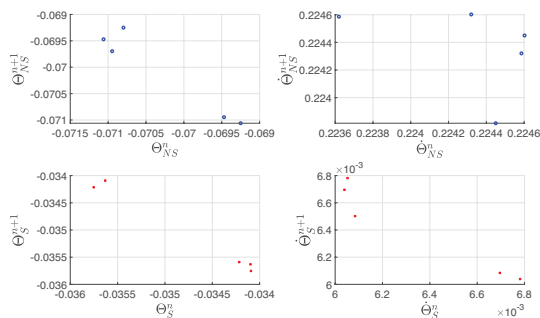
Ďalším modulom analýzy je A2, ktorý umožňuje analyzovať kráčajúceho robota vo fázovej rovine. Obrázok 19 zobrazuje fázový portrét kráčajúceho robota so zakrivenými chodidlami, ktorý vytvára uzavretý medzný cyklus. Podobne ako v simuláciách základného kráčajúceho modelu krúžidlovej chôdze dochádza k striedaniu stojnej a švihovej fázy. Medzi tými fázami dochádza k prepínaniu, čiže výskytu diskretných udalostí, ktoré sú na Obr. 19 ukázané modrou farbou.

Posledným modulom analýzy je A3, ktorý využíva na analýzu Poincarého mapy. Poincarého mapa je jedným z možných nástrojov na analýzu stability robotického chôdze. Údaje na Obr. 20 zobrazujú, že jednotlivé iterácie, ktoré prechádzajú Poincarého rovinou vytvárajú jeden bod len s malými odchýlkami spôsobenými nelineárnym správaním modelu.

Prípadová štúdia overila použitie metodiky aj pre zložitejšie konfigurácie kráčajúceho robota krúžidlovej chôdze.



Obr. 19. Fázový portrét (medzný cyklus) kráčajúceho robota so zakrivenými chodidlami



Obr. 20. Poincarého mapa kráčajúceho robota so zakrivenými chodidlami

III. RIADENIE KRÁČAJÚCICH ROBOTOV

Pohyb nôh robota pri pasívnej chôdzi po naklonenej rovine sa riadi prirodzenou dynamikou s hnacou silou v podobe gravitačného zrýchlenia g . Prirodzená dynamika systému určuje časový priebeh uhlových výchyliek $\theta(t)$, z ktorých vieme určiť pohyb nôh [21].

Existuje niekoľko riadiacich metód pre kráčajúce roboty. Každá zmena v modeli si vyžaduje zmenu v metóde riadenia. Metódy je však možné zoskupiť do menšieho počtu metód, ktoré majú spoločného menovateľa. Základné delenie riadiacich metód môže byť pre plne aktuované a pre podaktuované roboty.

A. Parciálna spätnoväzobná linearizácia kráčajúcich robotov

Princíp spätnoväzobnej linearizácie spočíva v pridaní takej spätnej väzby, ktorá zabezpečí, že nelineárny systém sa bude správať ako lineárny systém. Pre kráčajúce roboty sme zvolili metódu čiastočnej spätnoväzobnej linearizácie (partial feedback linearization - PFL), ktorá môže byť použitá pre podaktuované systémy. Pri použití parciálnej spätnoväzobnej linearizácie je nutné poznať referenčné trajektórie pre natočenie aktuovaných kĺbov [17].

Princíp parciálnej spätnoväzobnej linearizácie

Parciálnu spätnú väzbu je možné získať tým, že sa zohľadní podaktuovanosť vzhľadom na švihovú nohu $\theta_{NS}(t)$ alebo stojnú nohu $\theta_S(t)$. V prípade kráčajúceho robota bude uvažovaná podaktuovanosť vzhľadom na stojnú nohu $\theta_S(t)$,

to znamená, že stojná noha $\theta_S(t)$ nie je aktuovaná. Inak povedané, aktuátor sa nachádza len v bedrovom kĺbe, členok sa voľne otáča na podložke.

Pre výpočet zákona riadenia pomocou metódy PFL je vhodné uvažovať systém kružidlovej chôdze v štandardnej minimálnej forme podľa (5) [17].

Rozpísaním ODE formy (5) je možné získať sústavu dvoch nelineárnych diferenciálnych rovníc:

$$M_{11}\ddot{\theta}_{NS}(t) + M_{12}\ddot{\theta}_S(t) + N_{12}\dot{\theta}_S(t) + P_1 = u_{NS} \quad (70)$$

pre švihovú nohu

$$M_{21}\ddot{\theta}_{NS}(t) + M_{22}\ddot{\theta}_S(t) + N_{21}\dot{\theta}_{NS}(t) + P_2 = 0 \quad (71)$$

pre stojnú nohu

Z rovnice (71) je potrebné vyjadriť uhlové zrýchlenie stojnej nohy $\ddot{\theta}_S(t)$.

$$\ddot{\theta}_S(t) = -M_{22}^{-1}(M_{21}\ddot{\theta}_{NS}(t) - N_{21}\dot{\theta}_{NS}(t) - P_2) \quad (72)$$

Substitúciou (72) do (70) je možné získať vnútornú spätnú väzbu, ktorá transformuje nelineárny systém na parciálne linearizovaný.

$$u_{NS} = (M_{11} - M_{12}M_{22}^{-1}M_{21})\ddot{\theta}_{NS}(t) - M_{12}M_{22}^{-1}(N_{21}\dot{\theta}_{NS}(t) + P_2) + N_{12}\dot{\theta}_S(t) + P_1 \quad (73)$$

Linearizujúca spätná väzba (73) umožňuje predpísať zákon riadenia pre švihovú nohu $\theta_{NS}(t)$ v tvare:

$$\ddot{\theta}_{NS}(t) = v(t) \quad (74)$$

$$v(t) = \ddot{\theta}_{NS}^*(t) + K_1 e(t) + K_2 \dot{e}(t)$$

$v(t)$ je kombináciou lineárneho PD regulátora $K_1 e(t) + K_2 \dot{e}(t)$, ktorý zabezpečuje spätnoväzobné sledovanie referenčných trajektórií a doprednej zložky riadenia $\ddot{\theta}_{NS}^*(t)$, ktorá pôsobí ako referenčný vstup. PD regulátor je zložený z proporcionálnej zložky: $e(t) = \theta_{NS}^*(t) - \theta_{NS}(t)$ a derivačnej zložky: $\dot{e}(t) = \dot{\theta}_{NS}^*(t) - \dot{\theta}_{NS}(t)$, pričom $\theta_{NS}^*(t)$ a $\dot{\theta}_{NS}^*(t)$ sú referenčné trajektórie pre natočenie švihovú nohy $\theta_{NS}(t)$ a pre jej uhlovú rýchlosť švihovú nohy $\dot{\theta}_{NS}(t)$. Referenčné hodnoty môžu byť použité vybratím jedného kroku z medzného cyklu na Obr. 10 podľa [17].

Výsledný zákon riadenia pre kráčajúceho robota kružidlovej chôdze popísaného nelineárnymi diferenciálnymi rovnicami (20) a (21) má nasledovný tvar:

$$u_{NS}(t) = \left(b^2 m + \frac{b^2 l^2 m^2 \cos(\theta_{NS}(t) - \theta_S(t))^2}{a^2 m + l^2 m + l^2 m_H} \right) (\ddot{\theta}_{NS}^*(t) + K_1 e(t) + K_2 \dot{e}(t)) + u_{pfl} \quad (75)$$

kde

$$u_{pfl} = b g m \sin(\theta_{NS}(t)) - b \dot{\theta}_S^2(t) l m \sin(\theta_{NS}(t) - \theta_S(t)) + \left[b l m \cos(\theta_{NS}(t) - \theta_S(t)) \left(b l m \sin(\theta_{NS}(t) - \theta_S(t)) \dot{\theta}_{NS}^2(t) + (a m + l m + l m_H) g \sin(\theta_S(t)) \right) \right] (a^2 m + l^2 m + l^2 m_H)^{-1} \quad (76)$$

Periodická referenčná trajektória pre natočenie aktuovanej švihovú nohy pre podaktuovanú robotickú chôdzu môže byť vygenerovaná riešením rovnice:

$$\ddot{\theta}^*(t) = -[M(\theta^*(t))]^{-1}(N(\theta(t), \dot{\theta}(t))\dot{\theta}(t) + P(\theta(t))) \quad (77)$$

kde referenčné trajektórie θ^* je možné získať zo stabilného medzného cyklu [17] pasívneho robota kráčajúceho po naklonenej rovine. Referenčné trajektórie spolu s diskretnou dynamikou spĺňajú Lagrangeov systém rovníc (ODE forma):

$$M(\theta(t))\ddot{\theta}(t) + N(\dot{\theta}(t), \theta(t))\dot{\theta}(t) + P(\theta(t)) = 0 \quad (78)$$

Metóda riadenia založená na PFL je závislá na voľbe parametrov K_1 a K_2 a tiež na referenčných trajektóriách natočenia švihovú nohy $\theta_{NS}^*(t)$, jej rýchlosti $\dot{\theta}_{NS}^*(t)$ a zrýchlenia $\ddot{\theta}_{NS}^*(t)$, ktoré spolu s online aktuálnou trajektóriou robota určujú výsledný riadiaci zásah do podaktuovaného modelu kráčajúceho robota [17].

Pre riadenie robota kráčajúceho robota kružidlovej chôdze pomocou metódy založenej na PFL boli za referenčné trajektórie $\theta_{NS}^*(t)$, $\dot{\theta}_{NS}^*(t)$ a $\ddot{\theta}_{NS}^*(t)$ zvolené trajektórie $\theta_{NS}(t)$, $\dot{\theta}_{NS}(t)$ a $\ddot{\theta}_{NS}(t)$ keď robot pasívne kráčal po naklonenej rovine $\alpha = -0.1048 \text{ rad}$.

B. Dvojbodová okrajová úloha v plánovaní trajektórie

Plánovanie trajektórie kráčajúcich robotov rieši problém nájdenia dynamicky uskutočniteľného pohybu, ktorý zabezpečí požadované správanie. Dynamicky uskutočniteľný znamená, že kombinácia hybridného stavového vektora $\zeta(t)$ a radiaceho vstupu $u(t)$ existuje a je to platné riešenie pre hybridný dynamický systém [22].

Kľúčovou črtou plánovaného pohybu podaktuovanej robotickej chôdze je periodicitu, čo je rozdiel oproti plánovaniu trajektórie pre jednoduché typové podaktuované systémy, ktoré slúžia pre overenie metód riadenia pre robotickú chôdzu [12]. Jeden periodický pohybový vzor sa môže opakovať viackrát, napr. na prekonanie dlhej vzdialenosti, bez toho, aby sa zvyšovala náročnosť úlohy plánovania. Preto musí byť periodicitu zohľadnená pri plánovaní trajektórie [22].

Vygenerované požadované trajektórie sú periodické nielen v spojitých časti hybridného systému, ale aj v diskretných, preto pre hybridný stavový vektor musí platiť:

$$\zeta(t) = \zeta(t + T) \quad (79)$$

kde $T > 0$ je dĺžka periódy trajektórie. Periodickú trajektóriu pre aktuovaný kĺb je možné získať výberom periodických

požadovaných trajektórií a algoritmu riadenia, ktorý zabezpečí sledovanie požadovaných trajektórií. Dvojbodová okrajová úloha (boundary value problem - BVP) musí zabezpečiť periodické správanie aj pre neaktuovaný kĺb a tiež pre diskrétnu stavovú premennú $x_d(t)$.

Ak je systém hybridnej robotickej chôdze popísaný nelineárnou diferenciálnou rovnicou:

$$\dot{\theta}(t) = f(\theta(t)) \quad (80)$$

s počiatocnými podmienkami $\theta(t_0) = \theta_0$ a $\theta(t) \in \mathbb{R}^n$, potom príslušajúca dvojbodová okrajová úloha má tvar:

$$\dot{\theta}(t) = f(\theta(t)) \quad \varrho(\theta(t_0), \theta(t_f)) = 0 \quad (81)$$

kde $\varrho(\theta(t_0), \theta(t_f)) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ združuje okrajové podmienky. To znamená, že n podmienok, ktoré predstavujú správanie systému, sú rozdelené medzi počiatocné a koncové podmienky a umožňujú definovať počiatocné podmienky a koncové podmienky. Pre riešenie dvojbodovej okrajovej úlohy (81) môže byť použitý numerický solver, napríklad *bvp4c* integrovaný v programovom prostredí *MATLAB* [22].

Pre požadovanú trajektóriu chodiaceho robota kružidlovej chôdze, ktorý je popísaný nelineárnymi diferenciálnymi rovnicami (20) a (21) spolu s diskretnou dynamikou popísanou rovnicou (68), je potrebné definovať periodickú funkciu ako:

$$\theta_{NS}^*(t) = L_{CG} \cos(\omega t + t_0) \quad (82)$$

kde L_{CG} predstavuje dĺžku kroku a ω frekvenciu chôdze a t_0 predstavuje fázový posun, ktorý je vypočítaný algoritmom BVP.

Riešenie dvojbodovej okrajovej úlohy pre kráčajúce roboty odpovedá na otázku kedy a s akými počiatocnými podmienkami musí byť chôdza iniciovaná, aby sa získalo stabilné periodické správanie [22]. Preto je najprv potrebné definovať riadiaci zákon, ktorý umožňuje sledovanie požadovanej trajektórie $\theta_{NS}^*(t)$ pre natočenie švihovej nohy. Ako riadiaci zákon bol zvolený parciálny spätnoväzobný linearizačný prístup s PD regulátorom (74). Pohyb podaktuovaného kĺbu predstavuje vnútornú dynamiku v rámci PFL. Použitím parciálnej spätnoväzobnej linearizácie získame systém v tvare:

$$\begin{pmatrix} \ddot{\theta}_S(t) \\ \ddot{\theta}_{NS}(t) \end{pmatrix} = \begin{pmatrix} f_{int}(\theta(t), v(t)) \\ v(t) \end{pmatrix} \quad (83)$$

lineárny zákon riadenia (74) je možné prepísať do nasledovného tvaru

$v = \ddot{\theta}_{NS}^*(t) + K_1(\theta^*(t) - \theta(t)) + K_2(\dot{\theta}^*(t) - \dot{\theta}(t))$, ktorý asymptoticky sleduje požadované trajektórie pre švihovú nohu $\theta_{NS}^*(t)$. Zložka $\dot{\theta}_{NS}^*(t)$ predstavuje doprednú zložku riadenia.

Ďalej je potrebné definovať okrajové podmienky. Počiatocný, ale stále neznámy čas pre chôdzu je označený t_0 , koncový čas $t_f, t_f = t_0 + \frac{\pi}{\omega}$. Predpokladá sa, že krok sa začína v čase, keď dochádza k výmene stojnej a švihovej nohy a obe sú v kontakte so zemou na veľmi krátky čas. Pre túto konfiguráciu sú uhly švihovej $\theta_{NS}(t)$ a stojnej nohy $\theta_S(t)$ spojené rovnicou $\theta_S(t_0) = -\theta_{NS}(t_0)$, ktorá vychádza

zo vzťahu $\theta_S(t_0) - \theta_{NS}(t_0) + (\theta_{NS}(t_0) - \theta_{NS}(t_0)) = \pi$, ktorý vyplýva z vnútorného trojuholníka kružidlovej chôdze. Výpočet dvojbodovej okrajovej úlohy začína na začiatku kroku a končí v moment dopadu švihovej nohy na podložku, periodicita správanie je zabezpečená zvolenými okrajovými podmienkami.

Počiatocnú neznámu konfiguráciu $\theta_0 = \theta(t_0)$ je možné zapísať nasledovne: $\theta_0 = (\theta_{NS_0}, -\theta_{NS_0}, \dot{\theta}_{NS_0}, \dot{\theta}_{S_0})$ s neznámymi parametrami $\theta_{NS_0} = \theta_{NS}(t_0)$, $\dot{\theta}_{NS_0} = \dot{\theta}_{NS}(t_0)$ a $\dot{\theta}_{S_0} = \dot{\theta}_S(t_0)$. Navyše, ako je spomenuté vyššie, počiatocný čas t_0 nie je známy. Riešenie dvojbodovej okrajovej úlohy je takto definované vektorom parametrov $\sigma = (t_0, \theta_{NS_0}, \dot{\theta}_{NS_0}, \dot{\theta}_{S_0})$.

Okrajové podmienky $\varrho(\theta(t_0), \theta(t_f), \sigma) = 0$, ktoré sú potrebné na dosiahnutie periodického správania sú nasledovné:

$$\begin{aligned} \varrho_1(\theta(t_0), \theta(t_f), \sigma) &= \theta_{NS}(t_0) - \theta_{NS_0} \\ \varrho_2(\theta(t_0), \theta(t_f), \sigma) &= \theta_S(t_0) + \theta_{NS}(t_0) \\ \varrho_3(\theta(t_0), \theta(t_f), \sigma) &= \dot{\theta}_{NS}(t_0) - \dot{\theta}_{NS_0} \\ \varrho_4(\theta(t_0), \theta(t_f), \sigma) &= \dot{\theta}_S(t_0) - \dot{\theta}_{S_0} \\ \varrho_5(\theta(t_0), \theta(t_f), \sigma) &= \theta_{NS}(t_f) + \theta_{NS_0} \\ \varrho_6(\theta(t_0), \theta(t_f), \sigma) &= l \sin \theta_S(t_f) + l \sin \theta_{NS}(t_f) \\ \varrho_7(\theta(t_0), \theta(t_f), \sigma) &= \dot{\theta}_{NS}(t_f) + \dot{\theta}_{NS_0} \\ \varrho_8(\theta(t_0), \theta(t_f), \sigma) &= \dot{\theta}_S(t_f) + \dot{\theta}_{NS}(t_f) - \dot{\theta}_{S_0} \end{aligned} \quad (84)$$

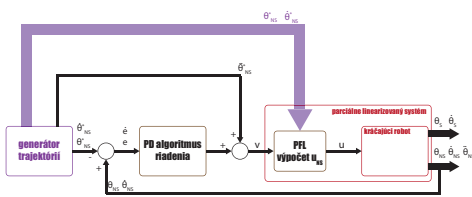
Opäť platí, že neexistuje žiadna explicitná podmienka pre $\theta_S(t_f)$, pretože $\theta_S(t_f)$ je implicitne definovaná z ϱ_5 . V okrajových podmienkach $\varrho(\theta(t_0), \theta(t_f), \sigma) = 0$ sú $\dot{\theta}_{NS}(t_f)$ a $\dot{\theta}_S(t_f)$ rýchlosti okamžite po dopade švihovej nohy. Okrajové podmienky ϱ_6, ϱ_7 a ϱ_8 vedú na symetrické správanie pretože $\theta_0 = \Phi(\theta^-)$. Aby bola zachovaná súmernosť, musia byť splnené podmienky ϱ_5 a ϱ_6 .

Nelineárne diferenciálne rovnice (20) a (21) vedú na systém diferenciálnych rovníc štvrtého rádu, čo umožňuje definovať štyri okrajové podmienky. 4-rozmerný vektor parametrov σ umožňuje definovať ďalšiu sadu štyroch okrajových podmienok. Parametre L_{CG} a ω sú vstupné parametre do dvojbodovej okrajovej úlohy, ktoré sú volené expertom. Ak existuje riešenie, tak pre každý pár $\sigma_{in} = (L_{CG}, \omega)$ je určené riešenie $\sigma = (t_0, \theta_{NS_0}, \dot{\theta}_{NS_0}, \dot{\theta}_{S_0})$ [22].

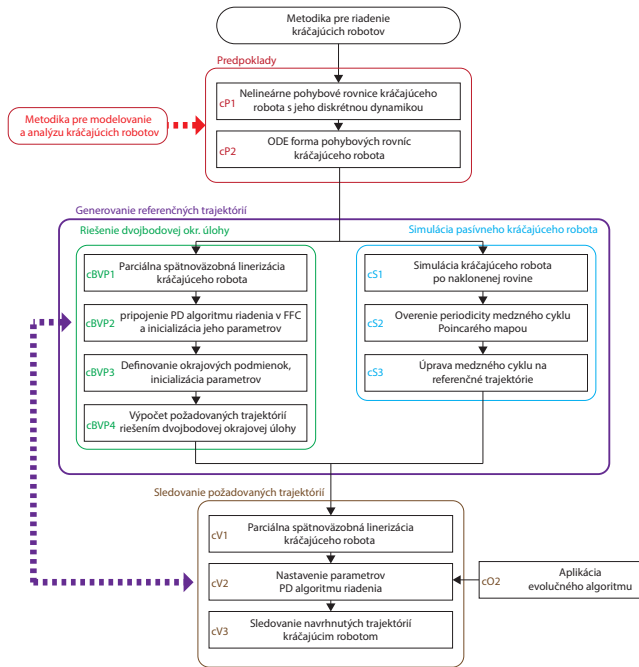
C. Návrh metodiky pre riadenie kráčajúcich robotov

Metodika pre návrh riadenia kráčajúcich robotov uvažuje s využitím doprednej riadiacej štruktúry (Feedforward Control - FFC), ktorej všeobecná schéma je ukázaná na Obr. 21. Metodika je overená niekoľkými prípadovými štúdiami, ktoré budú obsahovať nie len model základného kráčajúceho robota kružidlovej chôdze, ale napríklad aj kráčajúceho robota s kolenami.

Metodika pre návrh riadenia je rozdelená na tri moduly. V prvom module sú popísané predpoklady, ktoré musí splniť kráčajúci robot. Druhý modul metodiky sa zaoberá získaním referenčných trajektórií pre kráčajúceho robota a v poslednom module je predstavená verifikácia referenčných trajektórií ich



Obr. 21. Schéma doprednej riadiacej štruktúry pre kráčajúce roboty



Obr. 22. Schéma navrhutej metodiky pre syntézu riadenia kráčajúcich robotov

sledovaním kráčajúcim robotom. Metodika je navrhovaná pre podkatuované kráčajúce roboty.

Moduly *cP1* a *cP2* priamo nadväzujú na moduly *C4*, *C5* a *CD*, ktoré boli predstavené v metodike modelovania kráčajúcich robotov. Moduly *cP1* a *cP2* sú predpokladové moduly, ktoré musia byť splnené a vykonané na to, aby bolo možné ďalej použiť metodiku pre návrh riadenia - je nutné získať pohybové rovnice v ODE forme, ktoré popisujú kráčajúceho robota spolu s jeho hybridným modelom.

Modul *cBVP1* zabezpečuje vykonanie parciálnej spätňoväzobnej linearizácie kráčajúceho robota, aby bola zabezpečená lineárna dynamika kráčajúceho robota. Parciálna linearizácia kráčajúceho robota sa vykoná vzhľadom na švihovú nohu, ktorá je riadená. V module *cBVP2* je použitý PD algoritmus riadenia k doprednej zložke a k parciálne linearizovanému kráčajúcejmu robotovi ako súčasť parciálnej linearizácie a zabezpečenie sledovania požadovaných trajektórií v rámci algoritmu BVP. Tiež je nutné nastaviť parametre PD algoritmu riadenia.

V module *cBVP3* je potrebné definovať správne okrajové podmienky, ktoré zabezpečia periodickú dynamiku kráčajúceho robota. V module *cBVP3* je tiež potrebné nas-

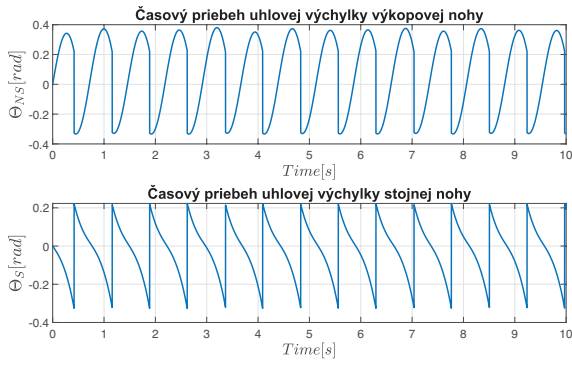
taviť vstupné parametre pre riešenie dvojbodovej okrajovej úlohy. Modul *cBVP4* obsahuje algoritmus riešenia dvojbodovej okrajovej úlohy, pričom je možné využiť napríklad funkciu programového prostredia MATLAB *bvp4c*. Tento krok tiež obsahuje prepis a výpočet referenčných trajektórií z výstupných parametrov riešenia dvojbodovej okrajovej úlohy.

Druhý podmodul generovania trajektórií je zaštrešený modulmi *cS1* až *cS3*. Modul *cS1* je zhodný s modulom, ktorý zabezpečujú simuláciu hybridného modelu kráčajúceho robota. Modul *cS2* overuje, či časové priebehy uhlových výchyliek získaných v module *cS1* vytvárajú požadovaný uzavretý medzný cyklus. Na overenie sa využíva analýza pomocou Poincarého máp. Ak časové priebehy uhlových výchyliek nevytvárajú uzavretý medzný cyklus, je potrebné zopakovať modul *cS1* s inými počiatočnými podmienkami simulácie a/alebo fyzikálnymi parametrami. Modul *cS3* vytvára z časových priebehov uhlových veličín získaných v module *cS1* referenčné trajektórie. Trajektórie uhlových výchyliek pasívneho kráčajúceho robota, ktorý kráča po naklonenej rovine, zabezpečia chôdzu pre riadeného kráčajúceho robota, ktorý bude kráčať po rovine bez sklonu.

Posledný modul metodiky pre návrh riadenia kráčajúcich robotov slúži pre overenie navrhnutého riadenia a sledovanie navrhnutých referenčných trajektórií. Modul *cV1* je zhodný s modulom *cBVP1*, jeho vykonaním je získaný parciálne spätňoväzobne linearizovaný model kráčajúceho robota. Modul *cV2* slúži pre nastavenie parametrov PD algoritmu riadenia, ktorý bol vybraný pre asymptotické spätňoväzobné sledovanie referenčných trajektórií. Parametre proporciálno-derivatívneho algoritmu riadenia môžu byť nastavené empiricky, alebo s využitím optimalizačných algoritmov. V predstavovanej metodike je využitý evolučný algoritmus. Posledný modul *cV3* overuje sledovanie referenčných trajektórií navrhnutých pre kráčajúceho robota. Overenie referenčných trajektórií prebieha s využitím Poincarého máp.

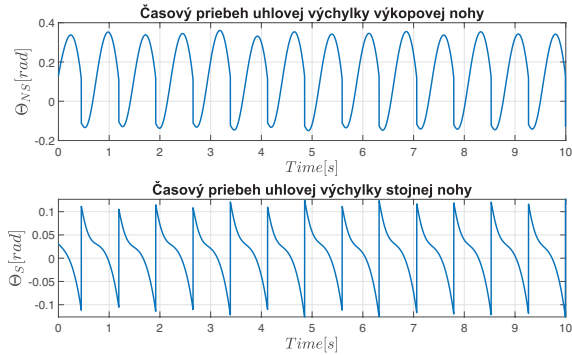
D. Overenie metodiky pre návrh riadenia kráčajúcich robotov

Predstavená prípadová štúdia: Kráčajúci robot chodiaci po rovine overuje navrhnutú metodiku pre riadenie kráčajúceho robota pričom je založená hlavne na využití poznatkov z teórie parciálnej spätňoväzobnej linearizácie. Prípadová štúdia podľa modulov *cP1* a *cP2* je uvažovaná pre kráčajúceho robota kružidlovej chôdze, ktorého nelineárne pohybové rovnice sú uvedené v prípadovej štúdii 1. Pre riadenie základného kráčajúceho robota kružidlovej chôdze pomocou metódy založenej na PFL spolu s PD algoritmom riadenia boli za referenčné trajektórie $\theta_{NS}^*(t)$, $\dot{\theta}_{NS}^*(t)$ a $\ddot{\theta}_{NS}^*(t)$ zvolené trajektórie kráčajúceho robota kružidlovej chôdze $\theta_{NS}(t)$, $\dot{\theta}_{NS}(t)$ a $\ddot{\theta}_{NS}(t)$, ktorý kráčal pasívne po naklonenej rovine so sklonom $\alpha = -0.1048rad$. Podrobný popis odvodenia a generovania nelineárnych diferenciálnych pohybových rovníc spojitej dynamiky a vzťahov diskretné dynamiky je uvedený v prípadovej štúdii 1. Tým sú vykonané moduly *cS1* a *cS2*, pričom referenčné trajektórie, ktoré sú výstupom modulu *cS3* sú ukázané na Obr. 23. Overenie navrhnutého riadenia, referenčných trajektórií a konštánt PD regulátora je vykonané

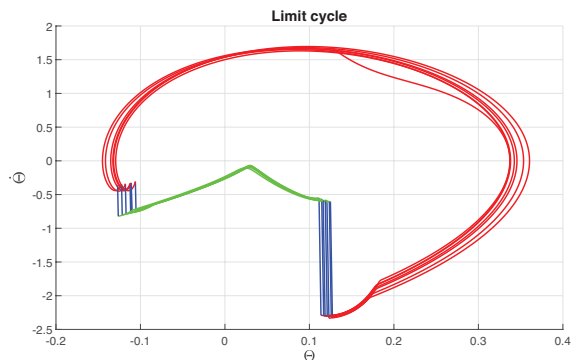


Obr. 23. Referenčný časový priebeh uhlových výchyliek $\theta_{NS}(t)$ a $\theta_S(t)$

vv rámci krokov $cV1$, $cV2$ a $cV3$. Konštanty PD regulátora boli volené empiricky a mali hodnotu $K_1 = 118$ a $K_2 = 18.8$. Riadenie kráčajúceho robota bolo následne vykonané pri kráčaní robota na rovnej podložke s nulovým sklonom, teda $\alpha = 0rad$. Výsledné časové priebehy aktuovaného kĺbu $\theta_{NS}(t)$ aj podaktuovaného kĺbu $\theta_S(t)$ riadeného kráčajúceho robota sú na Obr. 24.

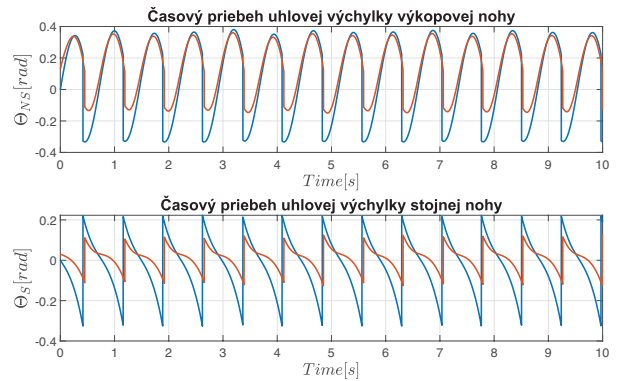


Obr. 24. Časový priebeh uhlových výchyliek $\theta_{NS}(t)$ a $\theta_S(t)$



Obr. 25. Medzný cyklus kružidlovej chôdze základného modelu riadeného pomocou algoritmu využívajúceho PFL

Medzný cyklus, ktorý vytvárajú trajektórie $\theta_{NS}(t)$ a $\theta_S(t)$ riadeného kráčajúceho robota sú na Obr. 25. Obr.



Obr. 26. Porovnanie referenčných uhlových výchyliek (modrá farba) a uhlových výchyliek získaných zo systému riadeného pomocou PFL (červená farba)

26 znázorňuje priame porovnanie referenčných trajektórií získaných pomocou pasívnej chôdze po naklonenej rovine a trajektórií získaných využitím algoritmu riadenia založeného na PFL spolu s PD regulátorom. Pre simulácie boli použité rovnaké parametre, ako v prípade pasívneho kráčajúceho robota.

Pre nastavenie parametrov PD regulátora K_1 a K_2 pre kráčajúceho robota sa literatúre nachádza len málo metód, aj keď sú tieto hodnoty kľúčové pre stabilitu a efektivitu kráčajúceho robota. V rámci metodiky pre riadenie kráčajúcich robotov, ktorá je predstavená v tejto dizertačnej práci, sme využili metódu umelej inteligencie pre nájdenie suboptimálnych parametrov K_1 a K_2 . Metóda PFL môže byť zovšeobecnená nie len pre model kráčajúceho robota kružidlovej chôdze ale aj pre ostatné zložitejšie modely kráčajúcich robotov, ktoré majú hybridnú dynamiku, napríklad pre kráčajúceho robota so zakrivenými chodidlami alebo pre kráčajúceho robota s kolenami.

1) *Generovanie trajektórií pre kráčajúci robot kružidlovej chôdze:* Prípadová štúdia: Generovanie trajektórií pre kráčajúci robot kružidlovej chôdze má za cieľ implementovať metodiku pre návrh riadenia kráčajúcich robotov. Kráčajúci robot je popísaný nelineárnymi diferenciálnymi rovnicami (20) a (21). Týmto sú splnené moduly $cP1$ a $cP2$.

Pred definovaním okrajových podmienok pre kráčajúceho robota je nutné najskôr parciálne linearizovať model kráčajúceho robota. Tento krok je zabezpečený vykonaním modulu $cBVP1$. Následne je ešte nutné podľa kroku $cBVP2$ zapojiť k parciálne linearizovanému robotovi PD regulátor, ktorý zabezpečí spätnoväzobné sledovanie trajektórií. Rovnice popisujúce kráčajúceho robota kružidlovej chôdze spolu vytvárajú systém štvrtého rádu, čo umožňuje definovať štyri okrajové podmienky, toto je obmedzenie použitého algoritmu riešenia dvojbodovej okrajovej úlohy. Ak pridáme 4-rozmerný vektor parametrov σ , umožňuje nám to definovať ďalšiu sadu štyroch okrajových podmienok. S týmito predpokladmi vieme splniť modul $cBVP3$ definovaním sady ôsmich okrajových podmienok (81).

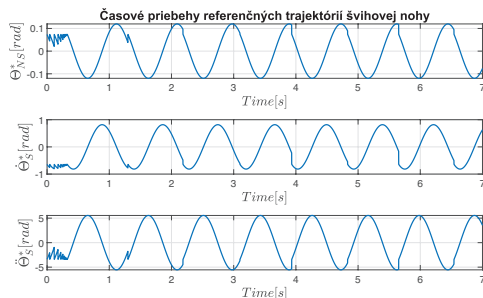
Pokračujeme modulom *cBVP4*. Vstupné parametre pre dvojbodovú okrajovú úlohu sú L_{CG} a ω . Vstupné parametre je nutné zvoliť manuálne, podľa skúseností experta. Ak bude existovať riešenie dvojbodovej okrajovej úlohy pre zvolené vstupné parametre, tak pre každý pár vstupných parametrov $\sigma_{in} = (L_{CG}, \omega)$ je určené riešenie $\sigma = (t_0, \theta_{NS_0}, \dot{\theta}_{NS_0}, \dot{\theta}_{S_0})$ [22].

V literatúre, napríklad v [22], je popísaných niekoľko numerických riešení dvojbodovej okrajovej úlohy pre kráčajúceho robota kružidlovej chôdze s tromi sadami vstupných parametrov $\sigma_{in} = (L_{CG}, \omega)$. Po získaní riešenia σ je potrebné skontrolovať, či potrebné okrajové podmienky sú dodržané.

Pre numerické riešenie dvojbodovej okrajovej úlohy bude využitá funkcia *bvp4c*, ktorá je implementovaná v programovom prostredí Matlab. Funkcia *bvp4c* implementuje trojstupňový linearizačný algoritmus Lobatto IIIa [23]. Časová doména riešenia je rozdelená na subdomény, pričom samotné riešenie sa aproximuje pomocou kubických splinov. Kubické spliny sú určené tak, že výsledná trajektória je spojitá a spĺňa riešenie zadanej diferenciálnej rovnice v bodoch mriežky a tiež spĺňa okrajové podmienky [24]. Riešenie je nájdené iteratívne pozdĺž gradientu, v prípade potreby sa body mriežky prepočítajú a celý postup sa opakuje.

Podľa východísk je potrebné najskôr určiť amplitúdu natočenia aktuovanej švihovej nohy a uhlovú frekvenciu rýchlosti zmeny požadovanej trajektórie. Pre kráčajúceho robota kružidlovej chôdze boli zvolené parametre nasledovne: $L_{CG} = 0.12$ a $\omega = 6.8$ a parametre PD regulátora boli zvolené nasledovne: $K_1 = 71.71$ a $K_2 = 9.06$. Pre zvolenú dvojicu vstupných parametrov boli získané nasledované výstupné parametre $\sigma = (t_0, \theta_{NS_0}, \dot{\theta}_{NS_0}, \dot{\theta}_{S_0})$:

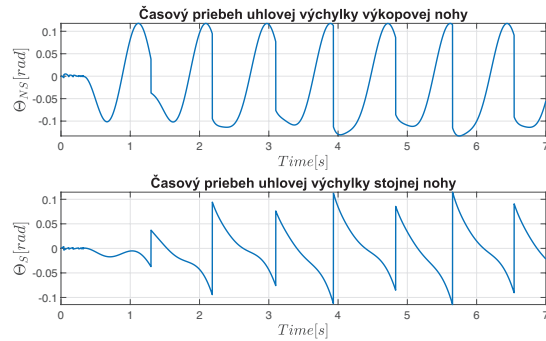
$$t_0 = 0.9109 \quad \theta_{NS_0} = 0 \quad \dot{\theta}_{NS_0} = 0 \quad \dot{\theta}_{S_0} = 0.0960 \quad (85)$$



Obr. 27. Referenčné trajektórie získané riešením algoritmu dvojbodovej okrajovej úlohy pre kráčajúceho robota

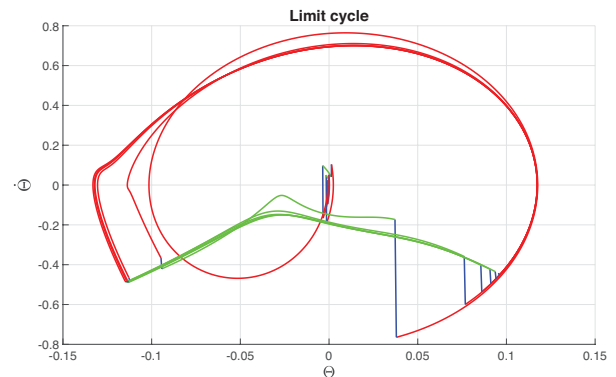
S využitím výstupných parametrov riešenia dvojbodovej okrajovej úlohy (85) sme získali referenčné trajektórie natočenia švihovej nohy. Tieto referenčné trajektórie sú ukázané na Obr. 27. Následne overíme navrhnuté trajektórie modulom *cV3*. Využijeme moduly *A1* a *A2* metodiky pre modelovanie kráčajúcich robotov, ktoré nám poskytnú časový priebeh uhlových výchyliek a tiež analýzu vo fázovej rovine.

Obr. 28 poskytuje časový priebeh uhlových výchyliek jednotlivých nôh.



Obr. 28. Časový priebeh uhlových výchyliek $\theta_{NS}(t)$ a $\theta_S(t)$ robota kružidlovej chôdze, ktorý sledoval referenčné trajektórie navrhnuté pomocou algoritmu BVP ($K_1 = 71.71$, $K_2 = 9.06$)

Obr. 29 znázorňuje zjednodušený fázový portrét aktívneho kráčajúceho robota kružidlovej chôdze.



Obr. 29. Fázový portrét - medzný cyklus kráčajúceho robota (BVP)

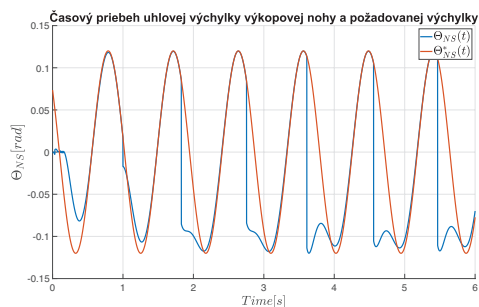
Okrem výstupov z modulov *A1* a *A2* sme vytvorili aj porovnanie požadovanej trajektórie a reálneho natočenia švihovej nohy, ktoré je ukázané na Obr. 30. Obr. 31 zobrazuje porovnanie kinetickej a potenciálnej energie kráčajúceho robota, ktorý sleduje požadované trajektórie. Obr. 31 tiež znázorňuje celkovú energiu robota, ktorá sa rovná súčtu kinetickej a potenciálnej energie: $E_{celk} = E_K + E_P$.

Správnym návrhom okrajových podmienok bolo možné navrhnuť periodickú trajektóriu, ktorá sa môže opakovať bez zmeny a robot dokáže prejsť aj dlhé vzdialenosti [22].

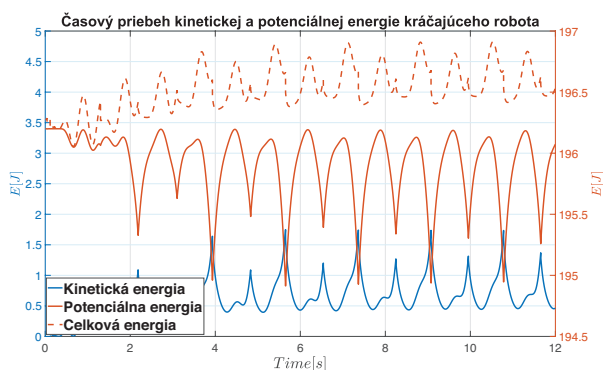
IV. ZÁVER

Článok sumarizuje riešenie cieľov dizertačnej práce a to hlavne modelovanie a analýzu kráčajúcich robotov v rámci formalizmu hybridných systémov. Dizertačná práca tiež pojednávala o návrhu riadenia pre vygenerované hybridné modely kráčajúcich robotických systémov.

Výskumné ciele dizertačnej práce je možné rozdeliť do dvoch hlavných častí. Prvá časť *Modelovanie kráčajúcich robotických systémov s hybridnou dynamikou* uvádza súčasný



Obr. 30. Porovnanie požadovanej trajektórie a reálneho natočenia švihovej nohy



Obr. 31. Porovnanie kinetickej a potenciálnej energie kráčajúceho robota

stav v modelovaní a analýze hybridných modelov typových kráčajúcich robotov. Na základe prehľadu súčasného stavu sa dizertačná práca zaoberala návrhom metodiky pre modelovanie a analýzu kráčajúcich robotov a jej overenia prípadovými štúdiami, ktoré overili metodiku rôzne zložitými typmi kráčajúcich robotov.

Druhá časť dizertačnej práce *Riadenie kráčajúcich robotických systémov s hybridnou dynamikou* sa venovala prehľadu v súčasnom stave riadenia typových kráčajúcich robotov so zameraním na metódy vhodné pre podaktuované kráčajúce roboty. Na základe súčasného stavu bola predstavená metodika pre riadenie typových modelov kráčajúcich robotov, ktorá bola overená prípadovými štúdiami s využitím modelov generovaných v rámci prvej časti dizertačnej práce.

Dizertačná práca sa zaoberala aj úlohami riešenými predovšetkým v rámci experimentu ALICE na LHC v CERNe. Predstavená bola distribuovaná architektúra riadiaceho systému detektora ALICE v CERNe, ktorá bola navyše porovnaná s DSR na školiacom pracovisku KKUI FEI TUKE. Koncept hybridných systémov je v dizertačnej práci využitý aj pre riešenie jednej z výskumných úloh v rámci experimentu ALICE na LHC v CERNe, ktorá spočívala v návrhu programových modulov pre aplikáciu ovládania a monitorovania elektroniky detektora ITS tvorenú operátorskými panelmi vo WinCC OA s využitím ALFRED systému pre komunikáciu s detektorovou elektronikou.

POĎAKOVANIE

Dizertačná práca bola riešená v rámci týchto vedeckovo-výskumných projektov:

- **Výskumné laboratórium nelineárnych podaktuovaných systémov (Research laboratory for nonlinear underactuated systems)**, grant FEI TUKE 2015-33, (1/2017 - 12/2017).
- **Univerzitný vedecký park Technicom pre inovačné aplikácie s podporou znalostných technológií- II.fáza**, kód ITMS:313011D232, spolufinancovaný zo zdrojov Európskeho fondu regionálneho rozvoja, "Centrum pre nedeštruktívnu diagnostiku technologických procesov so štandardným softvérovým vybavením pre riadenie a komunikáciu"(aktivita 3.1, PP7), (2015 2017),
- **Experiment ALICE na LHC v CERN: Štúdium silno interagujúcej hmoty v extrémnych podmienkach**, ALICE KE FEI TUKE(0222/2016 - 2/2016-DOT), (2016-2020),
- **Implementácia výsledkov vedeckého výskumu v oblasti modelovania a simulácie kyberfyzikálnych systémov do výučby – tvorba moderných vysokoškolských učebníc**, KEGA 072 TUKE – 4/2018, (20182020),

REFERENCES

- [1] Tech. Rep.
- [2] S. Feng, X. Xinjilefu, C. G. Atkeson, and J. Kim, "Optimization based controller design and implementation for the atlas robot in the darpa robotics challenge finals," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, 2015, pp. 1028–1035.
- [3] L. Koska, S. Jadlovská, D. Vošček, and A. Jadlovská, "Inverted pendulum with linear synchronous motor swing up using boundary value problem," *Acta Polytechnica*, vol. 59, no. 5, pp. 458–466, 2019.
- [4] S. Gupta and A. Kumar, "A brief review of dynamics and control of underactuated biped robots," *Advanced Robotics*, vol. 31, no. 12, pp. 607–623, 2017.
- [5] A. Goswami, B. Espiau, and A. Keramane, "Limit cycles in a passive compass gait biped and passivity-mimicking control laws," *Autonomous Robots*, vol. 4, no. 3, pp. 273–286, 1997.
- [6] H. Lin, P. J. Antsaklis *et al.*, "Hybrid dynamical systems: An introduction to control and verification," *Foundations and Trends® in Systems and Control*, vol. 1, no. 1, pp. 1–172, 2014.
- [7] A. J. Van Der Schaft and J. M. Schumacher, *An introduction to hybrid dynamical systems*. Springer London, 2000, vol. 251.
- [8] S. Jadlovský and J. Sarnovský, "Modelling of classical and rotary inverted pendulum systems—a generalized approach," *Journal of Electrical Engineering*, vol. 64, no. 1, pp. 12–19, 2013.
- [9] S. Jadlovská and J. Sarnovský, "Modelling of classical and rotary inverted pendulum systems—a generalized approach," *Journal of Electrical Engineering*, vol. 64, no. 1, pp. 12–19, 2013.
- [10] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics, Vol. I: The new millennium edition: mainly mechanics, radiation, and heat*. Basic books, 2011, vol. 1.
- [11] A. Jadlovská and S. Jadlovská, *Moderné metódy modelovania a riadenia nelineárnych systémov*. elfa s.r.o., Letná 9, Košice, 2013.
- [12] L. Tedrake, Russel, *Underactuated Robotics: Learning, Planning, and Control for Efficient and Agile Machines*, 2009.
- [13] S. Jadlovská, M. Kentoš, L. Koska, and J. Sarnovský, "Program tools for modeling and control of underactuated mechanical systems," in *Technical Computing*, 2015.
- [14] S. J. Goldstein H., Poole C.P., *Classical Mechanics: Pearson New International Edition*. Pearson Education Limited, 2014.
- [15] S. Čelikovský, J. Zikmund, and C. Moog, "Partial exact linearization design for the acrobot walking," 2008.
- [16] H. V. F. Chen, "Passive dynamic walking with knees: A point foot model," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

- [17] S. Kochuwila, S. Tripathi, and T. S. B. Sudarshan, "Control of a compass gait biped robot based on partialfeedback linearization," vol. 29, no. 3, pp. 117–127, 2012.
- [18] P. X. M. La Hera, A. S. Shiriaev, L. B. Freidovich, U. Mettin, and S. V. Gusev, "Stable walking gaits for a three-link planar biped robot with one actuator," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 589–601, 2013.
- [19] T. McGeer, "Powered flight, child's play, silly wheels and walking machines," in *Proceedings, 1989 International Conference on Robotics and Automation*, May 1989, pp. 1592–1597 vol.3.
- [20] E. R. Westervelt, J. W. Grizzle, C. Chevallereau, J. H. Choi, and B. Morris, *Feedback control of dynamic bipedal robot locomotion*. CRC press, 2007, vol. 28.
- [21] S. Jadlovská, "Modelovanie a optimálne riadenie nelineárnych podaktuovaných dynamických systémov," Ph.D. dissertation, Technická univerzita v Košiciach, 2015.
- [22] M. Sobotka, "Hybrid dynamical system methods for legged robot locomotion with variable ground contact," Ph.D. dissertation, Technische Universität München, 2007.
- [23] J. Kierzenka and L. F. Shampine, "A bvp solver based on residual control and the matlab pse," *ACM Transactions on Mathematical Software (TOMS)*, vol. 27, no. 3, pp. 299–316, 2001.
- [24] L. F. Shampine, J. Kierzenka, M. W. Reichelt *et al.*, "Solving boundary value problems for ordinary differential equations in matlab with bvp4c," *Tutorial notes*, vol. 2000, pp. 1–27, 2000.

Smart grid research

Maksym Oliinyk

Department of Electric Power Engineering
Faculty of Electrical Engineering and Informatics
Technical University of Kosice
Kosice, Slovakia
maksym.oliinyk@tuke.sk

Jaroslav Džmura

Department of Electric Power Engineering
Faculty of Electrical Engineering and Informatics
Technical University of Kosice
Kosice, Slovakia
jaroslav.dzmura@tuke.sk

Abstract— Smart grids are used to refer to a group of technologies that work together to create the next generation electrical grid. This paper explores the creation of an intelligent parking network. During the simulation, the connection of various elements of the intelligent network and their joint work is realized, as well as the main advantages from the introduction of a full-fledged smart network.

Keywords— Smart grids; electric vehicles; electric charging; demand management systems.

I. INTRODUCTION

Smart Grids are an automated system that independently monitors and distributes electricity flows to achieve maximum energy efficiency. Unlike the current distribution network, the smart grid will provide electricity from suppliers to consumers through two-way digital technology and an intelligent monitoring system to monitor the entire flow of electricity in the system. The aim of network modernization is to ensure the reliability of network operation, maintain accessibility, provide consumers with greater control to withstand physical strain, cyber-attacks. The reasons for modernization are the intensification of global competition, the extensive use of renewable energy sources, the reduction of carbon dioxide emissions in the world and the introduction of innovative technologies in the electricity sector. At the same time, environmental policies in many countries require higher levels of renewable energy to reduce carbon emissions. Renewable energy sources, such as photovoltaic cells and wind turbines, cannot ensure stable energy production because they are highly dependent on weather and climatic conditions. As a result, distributed renewable energy sources cause small (but on a large scale with increasing power) fluctuations in energy systems for which the conventional energy system was not intended. In order for the conventional energy system to adapt to such working conditions, a restructuring of the market structure and operating parameters is needed. On the other hand, smart grids consisting of flexible loads, storage systems and advanced control systems are able to integrate more intermittent renewable energy sources into the system at the local level. They are able to coordinate work between different distributed energy sources (DER) and effectively balance the demand for electricity. The intelligent energy system will include various technologies and tools that will allow the energy system to work much more efficiently. The US Department of Energy has identified five basic technologies needed to build next-generation networks [1][1][2].

These technologies include [3]:

- Integrated communication combining components with an open architecture for real-time information and control.
- Sensing and measurement technologies to support faster and more accurate response.
- Advanced components in superconductivity, power storage, power electronics and diagnostics.
- Advanced monitoring methods for key components for rapid diagnostics.
- Advanced interfaces and decision algorithms.

The European Union has identified the following important areas for development [4]:

- Intelligent network management.
- Demand-side control (DSM).
- Integrated Production and Storage (DG&S).
- Electric mobility.
- Extensive integration of renewable energy.

The new technological structure dramatically changes all processes in society, dictating the need for active introduction of electric vehicles into the transport and communication complex of countries with the creation of suitable infrastructure. Almost all global manufacturers plan to produce new cars that will run on electricity. In addition, according to many experts, electric vehicles have a future. The benefits of electric vehicles will undoubtedly make them the transportation of the future:

1. Reduce costs - An electric vehicle is a great way to save fuel.
2. Local reduction of environmental pollution (within the city) - the running engine does not release any harmful gases or other substances.
3. Noise reduction - electric motors are quite capable of providing quiet and smooth acceleration, while being able to achieve high acceleration.
4. Safety - road safety is a top priority for every healthy driver.

The main limiting factor in mass use is the high price of electric vehicles. However, progress in the field of electric vehicles is very rapid and in the near future their price will not differ much from the cost of hybrid and traditional cars. China will remain the largest market for electric vehicles. The government supports their dissemination in order to reduce CO₂ emissions. In table no. 1 shows the forecasts of various companies, which show the rapid penetration of electric vehicles into our lives.

TABLE I. PREDICTION OF THE NUMBER OF ELECTRIC VEHICLES BY TYPE OF COMPANY [5][6][7][8][9]

Company	Year	A million pieces	Percentage of the total number of cars
Frost & Sullivan	2025	25	≈20-22%
Bloomberg New Energy Finance	2025	–	≈50% of buses
Deloitte	2030	21	≈20%
BNEF	2030	30	–
International Energy Agency	2030	21,5	–
J.P. Morgan	2030	–	≈39%

II. MATHEMATICAL MODELS

Modeling is one of the most common methods of examining and studying the properties of a real object (an element of an electrical energy system) using a model that reflects with high degree of accuracy the processes and phenomena occurring in a real device. Modeling is divided into two groups depending on the models used:

1. Object modeling, which takes into account material models of objects
2. abstract modeling that takes into account theoretical models of objects.

A mathematical model is a mathematical description of the behavior of a real object - an element of the electric power system. The elements of the electricity system include electrical equipment intended for the production, transformation, transmission, storage, distribution or consumption of electricity.

A. Mathematical model of a photovoltaic panel

A photovoltaic cell is actually a large-area semiconductor diode (internal photoelectric effect) that converts solar energy incident in the form of radiation directly into electrical energy. This means that it works on the physical principle of the flow of electric current between two interconnected semiconductors with different electrical properties, which are affected by light radiation. It contains a **p - n** junction called a diode, a photocurrent generator representing the generation of current from light, and two resistors, one in series and the other in parallel, which describes the Joule effect and the loss of recombination [10].

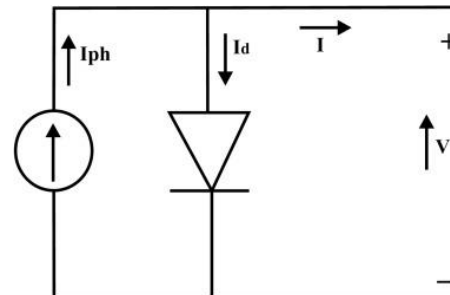


Fig. 1 Simplified equivalent circuit of the PV module. [10]

The photovoltaic panel can be considered as an ideal photovoltaic panel with a current source (I_{ph}) which are parallel to the diode, as shown in Fig. 1 and using Kirchhoff's first law, the output current of an ideal photovoltaic cell is described by equation (1) [10]:

$$I = I_{ph} - I_d \tag{1}$$

From semiconductor theory, the basic mathematical equation that describes the VAC (VACH) of a photovoltaic cell is known as the Shockley diode current equation, as shown in Equation (2) [11][12].

$$I_d = I_s \cdot [\exp((q \cdot U_{oc}) / (N_s \cdot K \cdot A \cdot T_0)) - 1] \tag{2}$$

Where I_s - saturation current is a fraction of the reverse current in a semiconductor diode caused by the diffusion of minority carriers from neutral regions in the depletion region. This current is practically independent of the reverse voltage. In the formula. (1) produces the output current I of an ideal photovoltaic cell as described in formula (3).

$$I_d = I_{ph} - I_s \cdot [\exp((q \cdot U_{oc}) / (N_s \cdot K \cdot A \cdot T_0)) - 1] \tag{3}$$

Ideally, the photovoltaic cell provides a very good approximation of the current generated by the photons, which is directly proportional to the intensity of the illumination. There are several parameters that are not taken into account in the ideal cell model and that actually or in practice affect the performance of the photovoltaic system [10][11][12].

B. Mathematical model of load

The process of electricity consumption is identified with the concept of electrical load, which is characterized by power and energy. The load can be one electrical appliance, a group of the same type of electrical appliances or a set of different electrical appliances - a mixed load. The value of active and reactive power of the load is required for tasks of analysis of steady state modes of electrical systems. The physical nature of energy consumption by electrical loads is such that its active and reactive forces depend on the supplied voltage and frequency in the electrical system. Such dependencies are called static load characteristics in terms of frequency and voltage. Different types of electrical loads have different static characteristics. In the sum of the different types of electrical consumers, the static characteristics of the mixed load are

taken into account. Here we will take into account the dependence of the load power only on the voltage - static voltage characteristics of the load. In this case, we will consider the frequency in the EES as a constant value. Modeling of electrical loads with static voltage characteristics in steady state calculations is considered to be the most accurate way to take into account the energy consumption of the load. However, experimental studies are needed to obtain true static characteristics, and the load composition must be known to select typical static characteristics, which can vary widely over time. In addition, in this case, the calculations should certainly take into account the action of voltage regulators, which greatly complicates the preparation of data and requires knowledge of legal regulations. Therefore, in most cases they use the simplest load model - constant values of active and reactive power: $P = \text{const}$, $Q = \text{const}$. Consider the electrical circuit in which the load is represented as the resistance Z_z . In general, this resistance is a variable - a nonlinear electrical circuit is obtained. Although we assume that the power consumed by the load is constant, the resistance will vary depending on the voltage according to the formula [13]:

$$Z_z = \frac{U_z^2}{S_z^*} \quad (4)$$

$$Z_z = \frac{U_z^2}{P_z(u_z) - jQ_z(u_z)} \quad (5)$$

The load can be represented in the form of two equivalent circuits: with series and parallel connection of elements (Fig. 2).

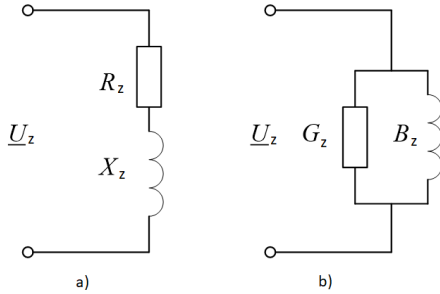


Fig. 2 Alternative load scheme [13]

With serial connection:

$$\underline{Z}_z = \frac{U_z^2}{S_z^*} \cdot (\cos \varphi + j \sin \varphi) = R_z + jX_z \quad (6)$$

parallel:

$$\underline{Y}_z = \frac{S_z}{U_z^2} \cdot (\cos \varphi - j \sin \varphi) = G_z + jB_z \quad (7)$$

C. Mathematical model of the battery

The voltage and current model, which describes how the voltage at the battery terminals varies according to current, is the most important submodel of the battery for the study of electrical systems. The best known volt-ampere model for unidirectional discharge is the Shepard model:

$$V_{batt} = E_0 - K \left[\frac{Q}{Q - i_t} \right] \cdot i - R_0 \cdot i \quad (8)$$

where E_0 is the open circuit voltage (OCV) of the battery at full capacity, the coefficient of resistance (Ω), Q is the capacity of the battery (Ah), i is the current of the battery (A), R is the internal resistance $a = \int i dt$. The second member is associated with losses of polarization and ohmic voltage and the last member is associated with losses on internal resistance.

III. PARKING HOUSE MODEL

The parking house will be used for internal purposes of company A with the possibility of using the parking spaces by the public outside working hours. The parking house will contain a Li-ion battery with parameters: Power +/- 120 kW and capacity +/- 120 kWh. A steel structure will be placed on the top floor of the parking house, on which photovoltaic panels used as roofing with a total output of 140 kW will be placed. The Python programming language was used for the construction of computational models and visualization. Visualization and drawings of the parking house are given below.

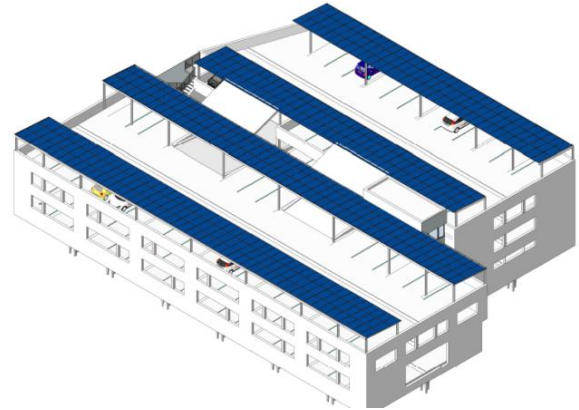


Fig.3 Visualization of a parking house

A. Modeling of PV plants

A photovoltaic panel from Solar was chosen for the created model of the photovoltaic panel, which will have real parameters. The technical parameters of this panel are listed below. It is estimated that the parking house will have approximately 520 modules. Which corresponds to 140 kW of installed power. All solar modules were divided into 4 groups, the installed power of each group is 35 kW.

TABLE II. TYPE AND NUMBER OF SELECTED PHOTOVOLTAIC PANELS [14]

Panel type	Quantity
PV Solar Modul MD P60PX 275 W (polycrystalline)	520

Data on the energy produced from a real photovoltaic panel were used to create a simulation of the work. Several weeks in winter and summer were selected for analysis. The energy production graph is shown in Fig. 4 and in more detail in Fig. 5. As can be seen from the figure in the selected time period, there are days with a fairly good generation and days without it. This choice is due to the fact that it allows you to simulate the operation of a parking house in different conditions. According to the visualization of the parking house model, it can be assumed that the angle of inclination will not be the most optimal and the efficiency of the actual PV plants will differ.

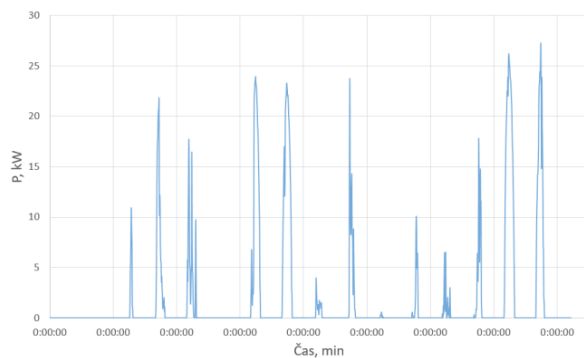


Fig. 4 Graph of output produced at one group of PV plants in the winter period

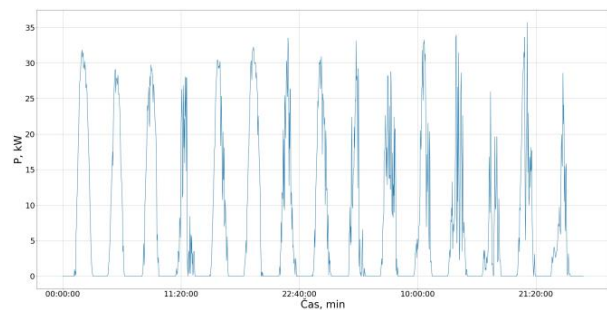


Fig. 5 Graph of output produced on one group of PV plants in the summer period

B. Modeling of electric vehicles

8 types of different electric vehicles were selected to simulate the operation of a parking house. The large selection of electric vehicles is due to the fact that different electric vehicles have different charging curves, different charging power, different battery capacity. The technical parameters of all selected electric vehicles are listed in Table 3.

TABLE III. TYPE AND NUMBER OF SELECTED PHOTOVOLTAIC PANELS [15]

Type of vehicle	Battery, kWh	3-phase 32A (22 kW)	3-phase 32A (22 kW), A	CCS (50 kW DC) avr, kW
Volkswagen e-Golf	32	7.2	2x16	39
Volkswagen e-Up	32.3	7.2	2x16	30
Skoda CITIGOE IV	32.3	7.2	2x16	30
Nissan Leaf	36	6.6	1x29	46
Hyundai Kona Electric	39.2	11	3x16	35
Hyundai IONIQ Electric	38.3	7.2	1x31A	34
SEAT Mii Electric	36.8	7.2	2x16A	30
Renault Kangoo Maxi ZE 33	31	7,4	1x32A	0

When charging with 3-phase charging, the electric vehicle sometimes creates a load for only 2 phases (Volkswagen e-Golf, Škoda CITIGOE IV) and some 1-phase (Nissan Leaf, Renault Kangoo Maxi ZE 33, etc.). As part of the experiment, two electric vehicles (Nissan Leaf and Volkswagen Golf) were charged at the same time. In FIG. 26 shows the current measurement in phase 1 and phase 2. As you can see, the current in phase 1 is almost three times larger, due to the fact that Nissan Leaf loads only one phase and Volkswagen Golf two phases. This phenomenon is dangerous in that it can cause an uneven load, and thus cause an asymmetry of voltages and currents, which negatively affects the operation of the electrical network.

C. Modeling of electric chargers

The project plans to use three types of electric chargers of different designs:

1. AC 22 kW
2. DC 25 kW
3. DC 50 kW

The number and total power of each type of charger are given in Table 4.

TABLE IV. NUMBER AND TYPES OF SELECTED ELECTRIC CHARGES

Charger	Power [kW]	Number of pieces	Total power [kW]	Connector
AC 22 kW wallbox	22	20	440	Typ 2
DC 25 kW wallbox	25	9	225	CCS + Chademo
DC 50 kW + 22 AC stojan	50 + 22	1	72	CCS + Chademo + Typ2

D. Battery information

A battery from SOCOMEC was chosen to simulate the battery. The main technical parameters are listed in Table 5.

TABLE V. PARAMETERS OF THE SELECTED ELECTRIC BATTERY

Useful battery capacity	100 – 120 kWh
Output power	33.3 kW
Minimum service life	6000 cycles
Ideal operating temperature	20°C– 25°C

E. Input data for creating a charging profile for electric vehicles

Article [16] presents (Fig. 6) the distribution of the daily distance of an electric vehicle based on a database. As already mentioned, the focus is on working days, because it is these days that it is clear that it is better to use network models, and where the work presented here is best applied. The total daily distance from 1 km to 40 km is the most common and the probability is constantly decreasing on daily journeys longer than 40 km.

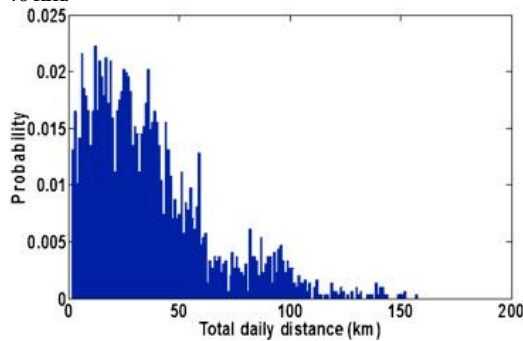


Fig. 6 Daily distance distribution scheme [17]

Article [17] deals with the deviation of simulated electric vehicle charging profiles from actual data. In this article, the authors created a model that generates charging profiles for electric vehicles based on real data. Figure 7 shows the result of the model. This model used 2,000 simulated vehicles out of 40 real vehicles tested over many years.

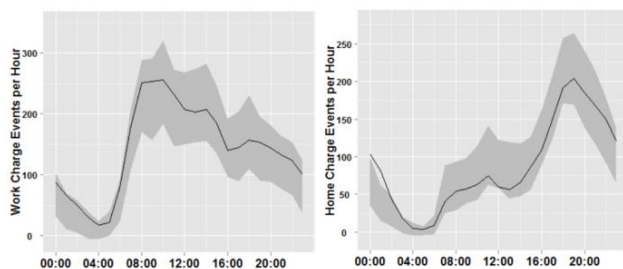


Fig. 7 Graph of the distribution of the number of charging processes for electric vehicles over time. A) Distribution of charging processes in the robot B) Distribution of charging processes at home [18]

Based on the collected data, several conclusions can be drawn:

- The average distance the car travels daily is between 20 - 40 km, which is equal to the consumption of 3.9 - 7.8 kWh.

- The most probable time of arrival at the workplace is the period 7:30 - 10:00 and the arrival home is the period 16:30 - 21:30.

IV. SIMULATION RESULTS

4 scenarios were created for the analysis of different working conditions. Each of the scenarios differs from each other during the charging start period. So in scenario 1, all cars will start charging between 7:00 and 10:00 in the morning and from 17:00 to 22:00. A description of each scenario is given in Tables 6.

- Scenario 1 - A scenario in which electric vehicles come to and from work at a narrow time interval.
- Scenario 2 - A scenario in which electric vehicles are charged mainly during low tariffs.
- Scenario 3 - A scenario in which electric vehicles can start charging over a wide period of time.
- Scenario 4 - A scenario in which all electric vehicles begin to charge simultaneously, at a specific time.

TABLE VI. DESCRIPTION OF SCENARIOS

Scenario	Charging time a day	Evening charging day	Photovoltaics	Battery
Scenario 1	7:00 – 10:00	17:00 – 22:00	+	+
Scenario 2	9:30 – 18:30	19:30 – 8:30	+	+
Scenario 3	6:00 – 14:00	16:00 – 02:00	+	+
Scenario 4	6:30 – 7:00	17:30 – 18:00	+	+

Fig. 8 shows the average of 10 simulations for each scenario. The lowest average load is in scenario 3, as this scenario lengthens over time. Scenarios 1 and 2 have approximately the same distribution, while scenario 4 most often reaches a maximum value.

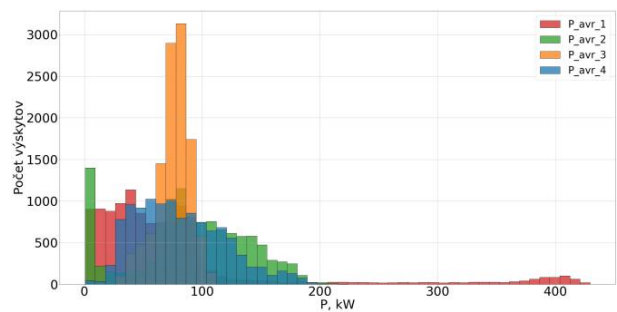


Fig. 8 Distribution of the total performance of the parking house over time

After generating the load data, the currents in each section of the power line, and the voltage values in each node of the parking network's electrical network were calculated. The simulation lasted a full 15 days. Figure 9. maximum current

values in the network. Figure 10 shows the minimum mains voltage values. Due to the fact that there are no voltage measurements in the node where the connection of the parking house is planned, a constant voltage of 230 V was used for the simulation. As can be seen from the graphs, the electrical equipment is not overloaded even at the maximum possible load. However, it is worth noting that the maximum load scenario is unlikely in real life. This simulation already includes the operation of the energy storage systems as well as the production of energy on photovoltaic panels. This will slightly reduce the maximum current in the network.

As part of the implementation of the parking house, the location of the electricity storage system is planned. When simulating the operation of a parking garage, this battery is mainly used to prevent the flow of electricity due to high production on photovoltaic panels, as well as to reduce peak currents in the grid. Figures 13 and 14 show graphs of the maximum current in the network. The graphs show that using a battery will reduce the maximum mains current by about 15-20%. The battery can also be used to charge several electric vehicles (the number depends on the power of the installed inverter) in the event of a power failure.

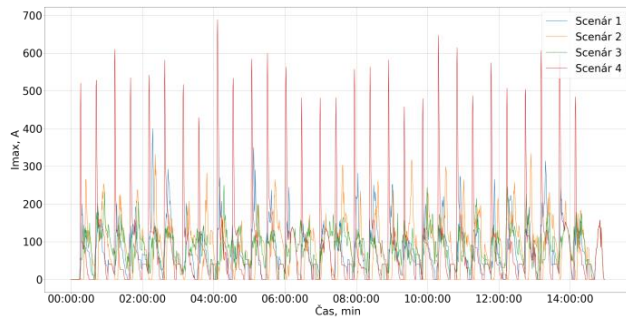


Fig. 9 Graph of maximum current in the network

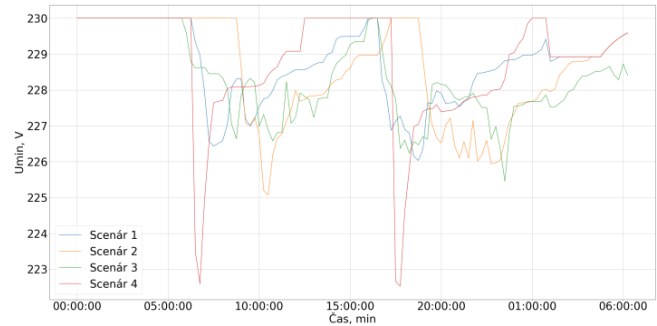


Fig. 12 Graph of minimum min voltage

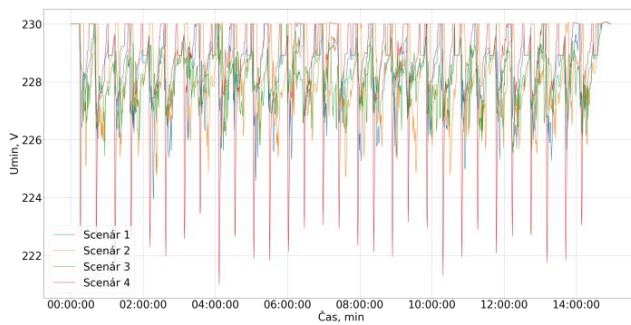


Fig. 10 Graph of minimum min voltage

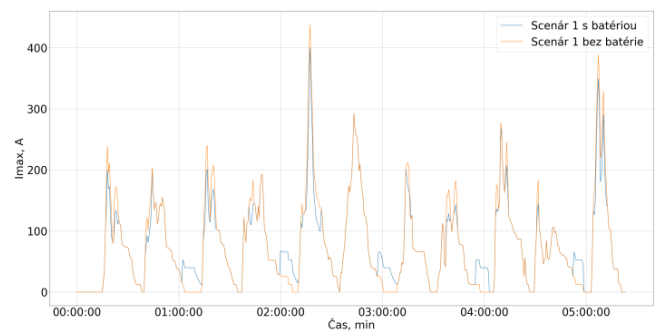


Fig. 13 Graph of maximum current in the network

Figures 11 and 12 show a one-day simulation. The graph shows the relationship between the current value and how narrow the switching-on time of the charging stations is. This relationship can be used to design or optimize charging strategies to reduce energy costs or load balancing at the grid node.

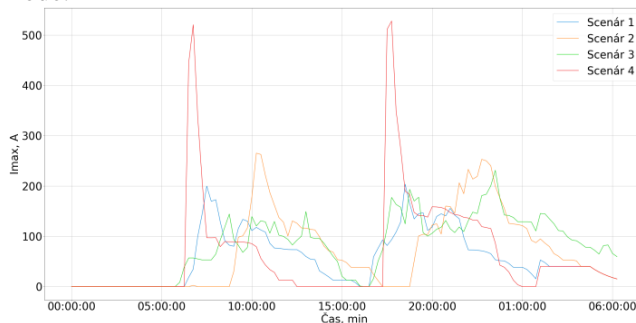


Fig. 11 Graph of maximum current in the network

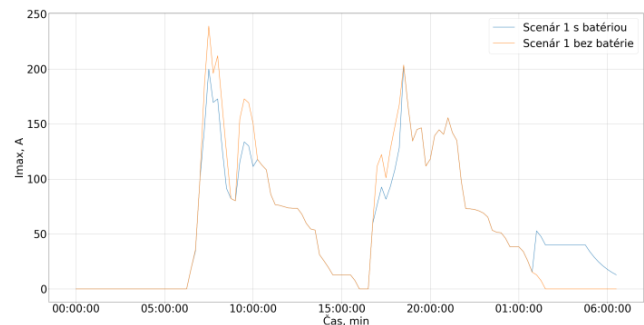


Fig. 14 Graph of maximum current in the network

V. CONCLUSION

In the article, a model of the electric part of a smart parking was created. For this model, there were generated charging profiles of electric vehicles that generate loads based on real

data. According to the results obtained, the power equipment of the parking lot is able to withstand the maximum possible load, subject to all existing standards. The created model can be used to predict the amount of energy consumed by parking, select power equipment, and create different charging strategies to maximize the use of low tariffs.

ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences by the Projects VEGA No. 1/0372/18 and by the Slovak Research and Development Agency under the Contract No. APVV-19- 0576.

REFERENCES

- [1] M. Shamshiri, C. K. Gan, a C. W. Tan, "A review of recent development in smart grid and micro-grid laboratories", v 2012 IEEE International Power Engineering and Optimization Conference Melaka, Malaysia, 2012, s. 367–372.
- [2] G. M. Shafiqullah, Amanullah M. T. Oo, A. B. M. Shawkat Ali, Peter Wolfs : «Grids 2030». A National Vision for Electricity's Second 100 years [Online]. Office of Electric Transmission and Distribution of USA Department of Energy, 2003. Available on the Internet: https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/Electric_Vision_Document.pdf
- [3] European Commission Directorate-General for Research Information and Communication Unit European Communities: European Technology Platform Smart Grids, Vision and Strategy for Europe's Electricity Networks of the future [Online]. Luxembourg, European Communities, 2006. Available on the Internet: https://ec.europa.eu/research/energy/pdf/smartgrids_en.pdf ISBN 92-79-01414-5
- [4] Smart Grid System report 2018, Report to Congress. United States Department of Energy Washington, DC 20585 [online], Available on the Internet: https://www.energy.gov/sites/prod/files/2019/02/E9/Smart%20Grid%20System%20Report%20November%202018_1.pdf
- [5] Mengolini A., "Smart grid projects outlook 2017", Luxembourg, 2017, ISSN 1018-5593 [Online]. Available on the Internet: https://ses.jrc.ec.europa.eu/sites/ses.jrc.ec.europa.eu/files/u24/2017/sgp-outlook_2017-online.pdf
- [6] Global EV Outlook 2019, [Online]. Available on the Internet: <https://webstore.iea.org/global-ev-outlook-2019>
- [7] Electric Vehicle Outlook Data. Wood Mackenzie. August 2018., [Online]. Available on the Internet: <https://www.woodmac.com/reports/power-markets-ev-charging-infrastructure-development-global-market-sizing-and-forecasts-29627>
- [8] IEA, Global EV Outlook 2018, May 2018, [Online]. Available on the Internet: <https://webstore.iea.org/global-ev-outlook-2018>
- [9] Ball J., "The Global Electric-Car Showdown Is Officially on in China", 2019 [Online]. Available on the Internet: <https://fortune.com/2019/03/22/electric-car-showdown-china/>
- [10] J.P. Morgan Driving into 2025: The Future of Electric Vehicles, October 2018 [Online]. Available on the Internet: <https://www.jpmorgan.com/global/research/electric-vehicles>
- [11] Villalva M.G., Gazoli J.R., Ruppert Filho E. Comprehensive approach to modeling and simulation of photovoltaic arrays, IEEE Trans. Power Electron., 24 (5) (2009), pp. 1198-1208. Available on the Internet: [10.1109/TPEL.2009.2013862](https://doi.org/10.1109/TPEL.2009.2013862)
- [12] Sera D., Teodorescu R., Rodriguez P. PV panel model based on datasheet values. Industrial Electronics, IEEE (2007), pp. 2392-2396. Available on the Internet: [10.1109/ISIE.2007.4374981](https://doi.org/10.1109/ISIE.2007.4374981)
- [13] Bellini A., Bifaretti S., Iacovone V., Comaro C. Simplified model of a photovoltaic module, Applied Electronics, IEEE (2009), pp. 47-51
- [14] Golubev M.L. Расчет токов короткого замыкания в электросетях 0,4-35кВ. - 2-е издание переработанное и дополненное. - М.: Энергия, 1980.
- [15] Solar panel technical specifications Modul MD P60PX [Online]. Available on the Internet: <http://www.pvsolarsys.sk/en/pvsolarsys-panel-polycrystalline-md-p60-px>
- [16] Data_base_EV [Online]. Available on the Internet: <https://ev-database.org>
- [17] Hill G., Blythe P.T., Higgins C., Deviations in Markov chain modeled electric vehicle charging patterns from real world data Intelligent transportation systems (ITSC), 2012 15th international IEEE conference on, 16–19 September (2012), pp. 1072-1077
- [18] Weldon P., Morrissey P., Brady J., O'Mahony M., An investigation into usage patterns of electric vehicles in Ireland, Transportation Research Part D: Transport and Environment, Volume 43, 2016, Pages 207-225, ISSN 1361-9209. Available on the Internet: <https://doi.org/10.1016/j.trd.2015.12.013>.

Virtuálno-reality technológie a inteligentné používateľské rozhrania

Martin Sivý

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
martin.sivy@tuke.sk

Branislav Sobota

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
branislav.sobota@tuke.sk

Abstrakt— Tento článok je sumarizáciou dizertačnej práce, zaoberá sa návrhom a implementáciou univerzálnych používateľských rozhraní pre ovládanie inteligentných prostredí a zariadení vo virtuálnej a zmiešanej realite. Hlavným cieľom práce bolo navrhnúť štandard pre opis rozhraní zariadení a delegáciu požiadaviek vytváraných používateľovou interakciou s prostredím. Bol navrhnutý multiplatformový webovo orientovaný systém pre virtualizáciu generických používateľských rozhraní pre rôzne typy inteligentných zariadení. Riešenie sa orientuje najmä na návrh rozhraní a ich softvérovú implementáciu. Implementácia riešenia prebiehala využitím technológií dostupných v laboratóriu LIRKIS Technickej univerzity v Košiciach. Pomocou zvolených metrik bolo následne riešenie overené v praxi z hľadiska používateľského zážitku metrikou SUS.

KLúčové slová—virtuálna realita, zmiešaná realita, SMART prostredie

Abstract— This article is a summary of the PHD thesis, its main focus is the design and implementation of universal user interfaces for controlling intelligent environments and devices in virtual and mixed reality. The main goal of this work was to design a standard for the description of device interfaces and the delegation of requests created by user interaction with the environment. A multiplatform web-oriented system for virtualization of generic user interfaces for various types of intelligent devices was designed. The solution focuses mainly on the design of interfaces and their software implementation. The implementation of the solution took place using technologies available in the LIRKIS laboratory at the Technical University in Košice. Using the selected metrics, the solution was subsequently verified in practice in terms of user experience with SUS metrics.

Keywords— virtual reality, mixed reality, smart interface

I. ÚVOD

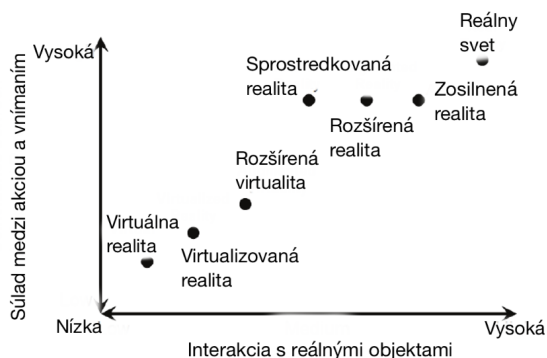
SMART prostredia a virtuálna realita sa s narastajúcou mierou digitalizácie a dostupnosťou technológií širokej verejnosti dostávajú do popredia. Oba tieto smery informačných technológií sa snažia obohatiť život a potreby používateľov. Rôzne prvky SMART zariadení je už postupom času možné nájsť takmer v každej domácnosti. Virtuálna realita, zmiešaná realita a aj rozšírená realita už nie sú pre človeka veľkou neznámou, pomaly sa stávajú súčasťou jeho

každodenného života. No napriek pokrokovej dobe a rovnakému účelu mnohých SMART zariadení, nie je ich ovládanie jednotné a vo veľkej miere závisí od špecifickosti zariadenia a jeho ekosystému, ktoré definuje výrobca. Je možné tvrdiť, že každé SMART zariadenie sa ovláda inak a je od výrobcu dodávané so svojim vlastným ovládacím prvkom, či už hardvérovým alebo softvérovým. Používatelia sú nútení učiť sa rôzne vzory a druhy ovládania pre zariadenia s rovnakým účelom. Problém ovládania takýchto zariadení môže najmä nastať v prípade, že ich používateľ chce používať v jednom SMART prostredí a tieto zariadenia nie sú medzi sebou kompatibilné. Tento článok sa zaoberá návrhom systému, ktorý zjednocuje ovládanie rôznych typov zariadení vo virtuálnej alebo zmiešanej realite bez ohľadu na ich špecifickosť alebo pôvod. Navrhnutý štandard upúšťa od rozdielov ovládania a snaží sa ich zjednotiť v jednom UI vo VR. Následne sa tento článok venuje testovaniu navrhnutého riešenia v praxi. Testovaním boli skúmané hardvérové nároky pre vytvorenie klientskú časť a to v počte snímkov za sekundu (FPS) v odozve na používateľove vstupy ako aj celkovú reakciu systému. V neposlednom rade bolo navrhnuté riešenie overené používateľmi so zameraním na jednoduchosť a intuitivnosť ovládania SMART zariadení vo VR pomocou SUS metricky.

II. VIRTUÁLNA REALITA

Rozhrania medzi reálnym svetom a úplne virtuálnou realitou je možné rozdeliť do niekoľkých kategórií. Milgram [8] tieto reality rozdelil podľa pomeru medzi tým akú rolu zohrávajú objekty reálneho sveta voči virtuálnym objektom vo vnímaní používateľa. Taktiež je tieto druhy realít možné kategorizovať podľa korelácie súladu medzi akciami, vnímaním a mierou interakcie s reálnymi objektami, čo je znázornené na Obr. 1.

Novšie radenie realít opisuje Margetis [9], sa ktorý snaží existujúce reality (VR, MR, AR) zlúčiť do jednej väčšej a inovatívnejšej oblasti X-reality (XR). „X“ v tomto prípade reprezentuje variabilné využitie príslušnej reality. V súčasnej dobe je trend vo vývoji týchto aplikácií zjednocovanie systémov na úrovni XR. Článok sa na príslušné XR systémy odkazuje podľa jednotlivých realít (VR, MR, AR) kvôli lepšiemu prehľadu.

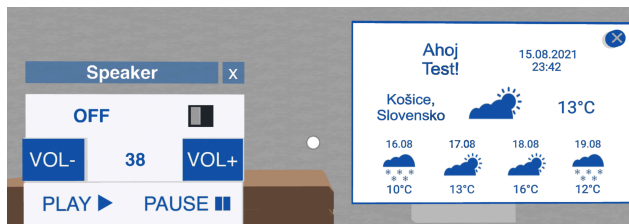


Obr. 1 Závislosť realít

III. NÁVRH OVLÁDANIA REÁLNYCH ZARIADENÍ A DÁTOVÝ MODEL KOMUNIKÁCIE

Ovládaniu reálnych zariadení pomocou rozhrania virtuálnej reality sa venovalo v minulosti viacero publikácií. Z prehľadu, ktorý vytvoril Zaidan [6] so zameraním na komunikačné komponenty pre IoT v chytrých domácnostiach je zrejme, že téma SMART prostredí je medziročne stále populárnejšou, a najviac článkov sa venuje vývoju a dizajnu. Anderson a spol. [5] navrhli spôsob synchronizácie udalostí medzi VR a IoT zariadeniami. Pre IoT senzory, ktoré by boli na nedostupných alebo nebezpečných miestach. Každá akcia vyvolaná vo virtuálnom prostredí sa odzrkadlí aj na zariadení reálnom, a naopak. V procese implementácie sa sústredili iba na integrovanie zariadenia Raspberry Pi. Podobne ako v tejto práci, aj autori v [5] potrebovali vytvoriť štandard pre zaznamenávanie zmien na zariadeniach. Každá zmena na zariadení používateľom vyvolá akciu s inštrukciou, ktorá aktualizuje stav zariadenia v reálnom aj virtuálnom svete. Dáta sa medzi platformami preposielajú v podobe JSON objektov, ktoré obsahujú informácie potrebné pre vykonanie zmeny, ako sú napríklad token používateľa, druh akcie a ak je to potrebné tak aj hodnota, na ktorú má byť daný parameter zariadenia nastavený. V práci Hamdana [7] boli navrhnuté dva elementy ovládania IoT pomocou mobilného telefónu, išlo iba o prepínače a posúvače, ale bolo možné počet takýchto zariadení ľubovoľne rozširovať.

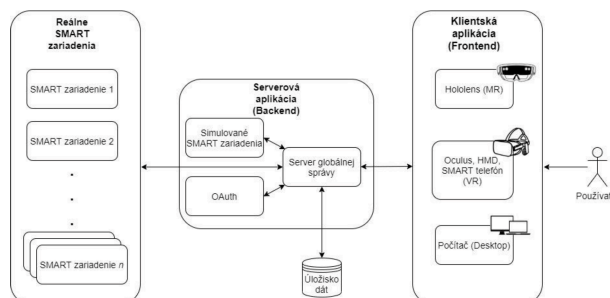
Tento článok sumarizuje vytvorenie generického návrhu komunikácie, pomocou ktorého bude možné opísať ovládanie ľubovoľného SMART zariadenia. Riadenie požiadaviek je centralizované serverom a teda je možné do systému pridávať a odoberať SMART zariadenia v ľubovoľnom počte. Rovnaký spôsob komunikácie je použitý pre tok dát zo servera ku SMART zariadeniam ako aj od servera ku klientským aplikáciám. Každé SMART zariadenie v systéme je reprezentované svojim generickým grafickým rozhraním, ktoré reprezentuje virtualizáciu ovládacieho rozhrania a jednotlivé akcie prvkov ovládania. Takýto ovládací panel má 2D formu v 3D priestore. Podporovanými ovládacími elementami sú: *timer*, *input*, *slider*, *checkbox*, *radio button*, *button*, *label* a *image*. Z týchto prvkov je možné vyskladať ovládací panel, aký je zobrazený na Obr. 2.



Obr. 2 Vzor ovládacieho panelu

IV. APLIKÁCIA

Prezentované riešenie je možné rozdeliť do troch samostatných celkov ako na Obr. 3. Prvým z nich je server (backend), ktorý je zodpovedný za celú logiku a ovládanie ostatných častí systému, s ktorými komunikuje pomocou technológií REST a websocket. Taktiež zabezpečuje autorizovaný prístup používateľov k dátam. Druhou časťou je súbor viacerých SMART zariadení, ktoré komunikujú výlučne zo serverom pomocou špecifickej dátovej štruktúry. Poslednou hlavnou súčasťou systému je klientská aplikácia (frontend), pomocou ktorej používatelia riadia celý systém. Klientská aplikácia je multiplatformová a teda dokáže podporovať viacero módo, virtuálnu realitu (2D/3D) a zmiešanú realitu využívajúc viacero virtuálno-reálnych zariadení dostupných v rámci laboratória LIRKIS KPI. Zameranie celkového systému je na používateľov zrakový vnem.



Obr. 3 Konceptuálny model systému

A. Serverová aplikácia

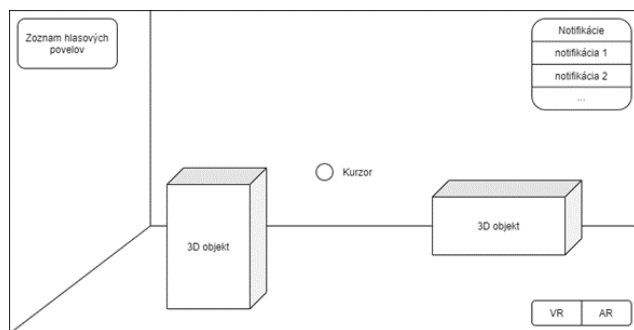
Serverová aplikácia je medzníkom komunikácie medzi klientskou aplikáciou a SMART zariadeniami napojenými do tohoto systému. Server ponúka vystavené API služby, na ktoré je ďalej možné napájať a vyvíjať ľubovoľné klientské aplikácie. Zaoberá sa správaním SMART prostredí a používateľov, validovaním vstupov smerom ku klientským aplikáciám a taktiež k SMART zariadeniam. Ďalšou možnosťou využitia systému je prototypovanie zariadení a simulácia ich stavu aj bez reálne pripojených SMART zariadení, čím sa stáva testovanie klientských aplikácií jednoduchším. Medzi hlavné systémy výhody patrí vytvorenie komunikačného štandardu pre ovládanie ľubovoľného počtu SMART zariadení prostredníctvom jednej klientskej aplikácie. Systém sa snaží o prepojenie viacerých typov zariadení, ktoré nemusia byť od jedného výrobcu, a nie je potrebné aby boli vzájomne kompatibilné. Teda ich prepojenie a integrovanie do

jedného SMART priestoru bez použitia navrhovaného systému môže byť náročné ba priam až nemožné. Taktiež si používateľ môže naplánovať jednotlivé vykonávanie úloh na zariadeniach s možnosťou pravidelných opakovaní presne definovaných scenárov, a teda takto reťaziť vykonávanie aj zariadení z rôznych ekosystémov a prispôbiť ich svojim potrebám.

B. Klientská aplikácia

Na základe zozbieraných poznatkov počas doktorandského štúdia bola táto aplikácia vytváraná s ohľadom na cross-platformovosť a webovú-orientáciu. Riešenie je teda nasadené na serveri a používateľ sa nemusí starať o žiadnu lokálnu údržbu alebo sťahovanie nových aktualizácií a obávať sa o kompatibilitu so systémom.

Návrh používateľského rozhrania je tvorený s ohľadom na čo najmenej rušivých prvkov grafického rozhrania a jeho návrh je zobrazený na Obr. 4. Model finálneho riešenia grafického rozhrania je zobrazený na Obr. 7 prezentovaním jedného SMART prostredia.



Obr. 4 Návrh prvkov klientského rozhrania

Ovládanie kurzora je pri použití dátovej prilby riešené natáčaním celej hlavy (nie len pohybom očí). Kurzor hlavového pohľadu je bod, ktorý sa pripája na koniec vektora pohľadu smerujúceho zo stredu zorného poľa používateľa.

Keďže táto jedna aplikácia je určená pre viacero zariadení, budú následne opísané možnosti ovládania avatara používateľa v priestore. Režimy zobrazenia sa dajú meniť v pravej dolnej časti obrazovky a systém automaticky rozpozná typ HMD zariadenia.

Ovládanie v 2D režime:

- Pohyb avatara je riešený pomocou smerových šípok na klávesnici,
- Rotácia pohľadu kamery je pomocou pohybu myši a stlačeného ľavého tlačidla,
- Kolízia s objektami scény je zapnutá.

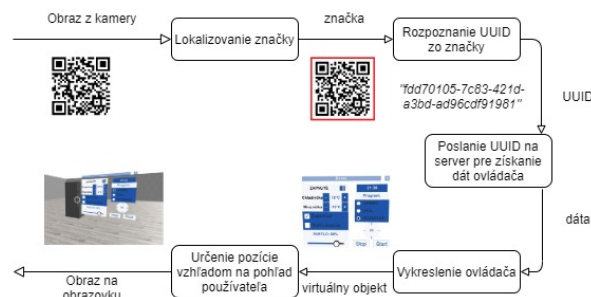
Ovládanie v 3D režime:

- Pohyb avatara je možné uskutočniť pomocou viacerých prístupov. Prvým prístupom je snímanie pohybu zariadenia pomocou obrazu z kamery a gyroskopu zariadenia. Druhý

prístup je pomocou vyznačených virtuálnych smerových šípok umiestnených pod avatarom, ak používateľ nasmeruje kurzor na danú šípku avatar sa pohybuje v danom smere šípky. Ďalšou možnosťou je v scéne rozmiestniť kontrolné body a pri nasmerovaní kurzora na takýto bod je avatar teleportovaný na danú pozíciu.

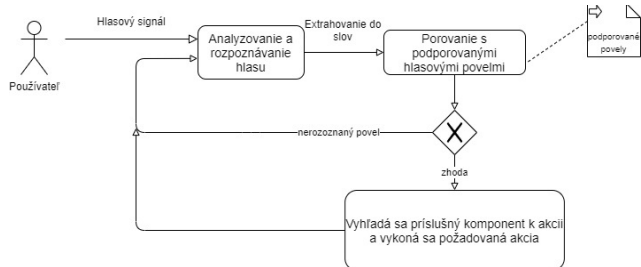
- Rotácia pohľadu kamery je ovládaná reálnym pohybom hlavy používateľa (pri použití dátovej prilby)
- Kolízia s objektami je vypnutá, nakoľko by dochádzalo k neúmernému pohybu medzi používateľom v reálnom a virtuálnom svete.

Rozpoznávanie objektov z obrazu sa venovalo už viacero prác [2][3][4]. Autori sa zameriavali aj na rozpoznávanie objektov z obrazu použitím HoleLens zariadenia. No napriek tomu tieto publikácie neprinášajú spôsob ako rozpoznať 2 identické zariadenia toho istého typu. Proces rozpoznávania reálnych objektov pre zmiešanú realitu bol v tejto práci navrhnutý pomocou QR značiek s jedinečným identifikátorom každého zariadenia (pomyselným sériovým číslom). Tým pádom je možné rýchlo a jednoducho pomocou kamery a QR kódu identifikovať zariadenie. Postup od identifikácie až po vykreslenie virtuálneho rozhrania je popísaný na Obr. 5.



Obr. 5 Proces od snímania QR značky až po jej vykreslenie v UI

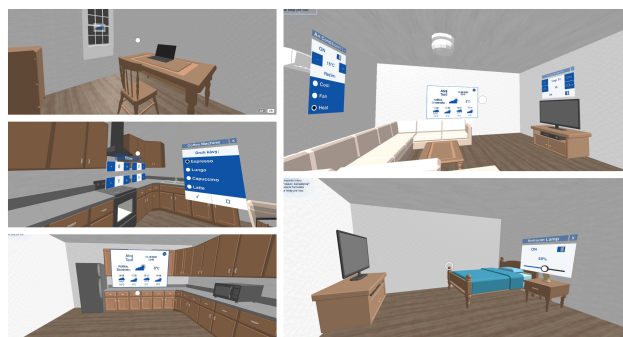
Taktiež boli do aplikácie pridané hlasové povely. Hlasové povely rozširujú možnosti interakcie používateľa so systémom a v tomto riešení sú použité pre menej časté funkcie systému. Proces rozpoznávania hlasových povelov je zobrazený na Obr. 6. a tabuľka so základnými povelmi **Tab. 1**. Nevýhodou použitej knižnice pri implementácii je obmedzenie kompatibility len na Google Chrome prehliadač.



Obr. 6 Diagram procesu rozpoznávania hlasových povelov

Tab. 1 podporované hlasové povel

Povel	Opis
Show menu	Zobrazí pred používateľom základné menu systému
Show +“názov zariadenia“	Zobrazí menu konkrétneho zariadenia podľa názvu
Turn on/off + “názov zariadenia“	Ovládanie zapnutia a vypnutia zariadenia (ak je táto funkcia k dispozícii)
Show settings	Zobrazí pred používateľom panel s ovládaním SMART prostredia
Show users	Zobrazí pred používateľom panel s používateľmi a ich právami v SMART prostredí
Show appliances	Zobrazí pred používateľom panel so všetkými registrovanými zariadeniami v SMART prostredí a umožní ich ďalej spravovať

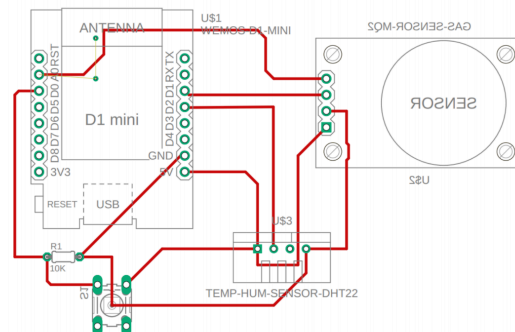


Obr. 7 Model navrhnutého grafického rozhrania klientskej aplikácie

C. SMART zariadenia

Navrhnutý systém dokáže pri využití protokolu posilať požiadavky ďalej na príslušné zariadenie a vice versa. Pre overenie funkčnosti boli niektoré modelové zariadenia pre overenie funkcionality skonštruované na báze Arduino dosiek.

Príkladom modelového zariadenia je „chladnička“, ktorá disponovala svetlom a senzorom pre „nezatvorené dvere“ s možnosťou nastavenia teploty. Taktiež bolo navrhnuté zariadenie SMART teplomer, ktoré meralo a hlásilo teplotu a vlhkosť, alebo zariadenie „pohybový senzor“. Model zapojenia „SMART teplomera“, je zobrazený na Obr.



Obr. 8 Model inteligentného teplomera vytvorený pomocou Arduino komponentov

Niektoré súčiastky boli vytlačené aj 3D tlačiarňou, zobrazené na Obr. 9



Obr. 9 Skonštruované zariadenie na báze Arduino dosky

V. EXPERIMENTÁLNA ČASŤ

Navrhnuté a implementované riešenie systému bolo overené pomocou objektívnych a subjektívnych metrik. Medzi objektívne patrilo meranie FPS a latencie systému podľa počtu pripojených používateľov v zdieľanom prostredí a subjektívne testovanie prebehlo pomocou metódy SUS.

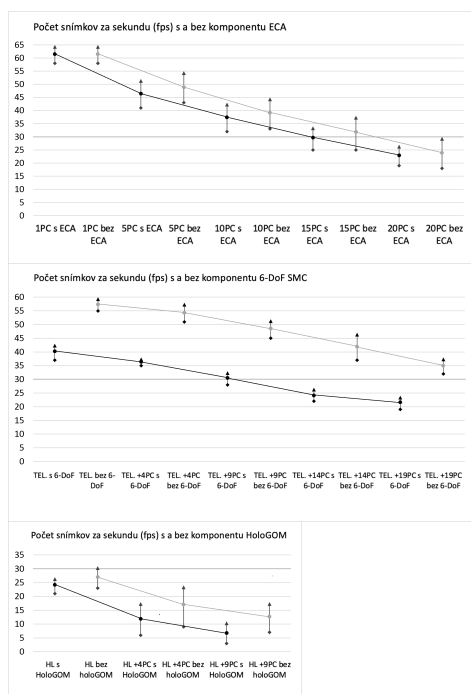
A. Meranie FPS a latencii systému s narastajúcou záťažou viacerých používateľov

V tomto experimente prebehlo meranie výkonu klientskej aplikácie v závislosti od počtu pripojených a vizualizovaných ostatných používateľov [10]. Taktiež sa meral výkon pri zapnutých a vypnutých navrhnutých komponentoch. Navrhnuté komponenty boli 3. Prvý s názvom ECA [11] (Enhanced Client

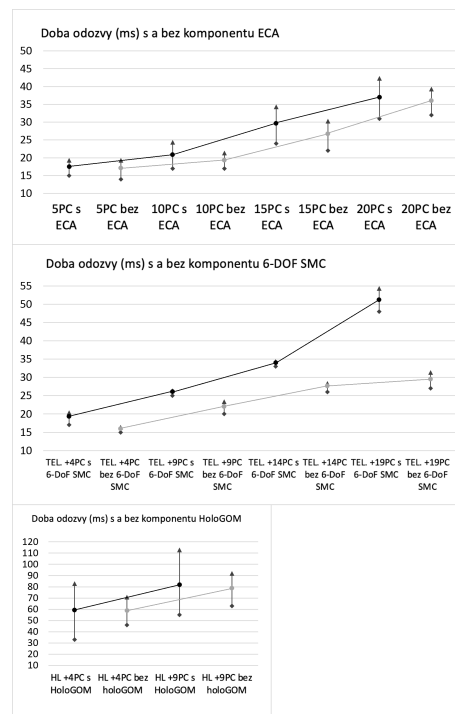
Access), ktorý rozširuje klasickejší komponent NAF (od A-frame). Ďalším komponentom je 6-Dof SMC (Smartphone Motion Capture), tento komponent zabezpečil využitie smartfónu ako prostriedku pre pohyb vo virtuálnej scéne. Posledný komponent bol navrhnutý pre podporu rozpoznávania gest a pohybu zariadenia HoloLens s názvom HoloGOM.

Testovaním sa overovala plynulosť používateľského zážitku, kde cieľom bolo mať hodnotu FPS aspoň 30/s a hodnotu odozvy pod 100ms. Testovali sa jednotlivé komponenty zariadení s pripojením ďalších používateľov do zdieľanej scény a to 5, 10, 15, 20.

Výsledky testovania sú zobrazené na nasledujúcich grafoch zobrazujúcich počet FPS v závislosti od zariadenia/komponentu a počtu pripojených používateľov Obr. 10. a graf pre dobu odozvy taktiež v závislosti od zariadenia/komponentu a počtu pripojených používateľov Obr. 11.



Obr. 10 Graf znázorňujúci nameraný počet snímkov pre jednotlivé prípady



Obr. 11 Graf znázorňujúci nameranú dobu odozvy pre jednotlivé prípady

B. Meranie výkonnosti systému z pohľadu latencii

Tento experiment [1] sa venuje porovnaniu latencii servera pri použití troch rôznych prístupov ovládania prostredia. Skúma interakciu človeka s prostredím z hľadiska plynutosti používateľského zážitku. Pri všetkých 3 metódach sa využíva chytrý telefón ako podriadené zariadenie slúžiace ako ovládač kurzora scény v režimoch:

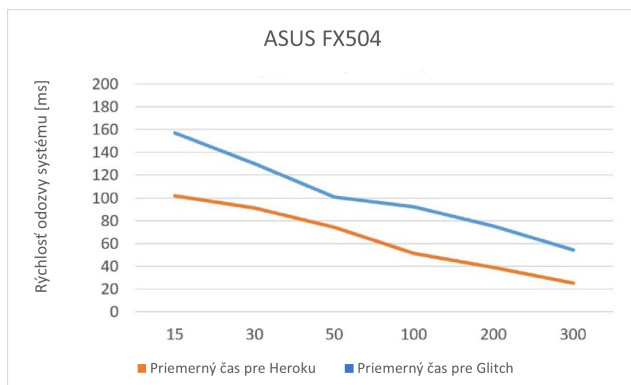
- Raycaster,
- VR Manipulator,
- Touch VR Joystick.

Použitelnosť daného prostredia a rýchlosť splnenia úlohy používateľmi bola testovaná na dvoch rôznych cloudových platformách, Heroku a Glitch. Tieto prostredia na pozadí využívajú služby Amazon Web Services. Testovací server mal konfiguráciu: CPU Intel Xeon E5-2686 v4 a RAM 2GB.

Testovanie prebehlo na vzorke 20 participantov. Cieľom bolo zmerať trvanie odozvy medzi používateľovou interakciou a vykonaní vizuálnej odozvy v scéne.

Navrhnuté rozhranie sa najprv testovalo na 3 rôznych zariadeniach slúžiacich ako „master“ a to: ASUS FX504, Microsoft HoloLens a Oculus Quest. Ako „slave“ zariadenie slúžil inteligentný telefón Xiaomi Redmi Note 7. Testovanie prebiehalo s rôznou mierou frekvencie odosielania polohových údajov a to od 15 do 300 milisekúnd.

Z grafov je možné vidieť, že najlepšia odozva bola dosiahnutá použitím najmenšieho intervalu pre odosielanie nových dát a to 15 milisekúnd (66Hz). Tiež z grafov vyplýva, že server Heroku dosahoval lepšie výsledky Obr. 12.



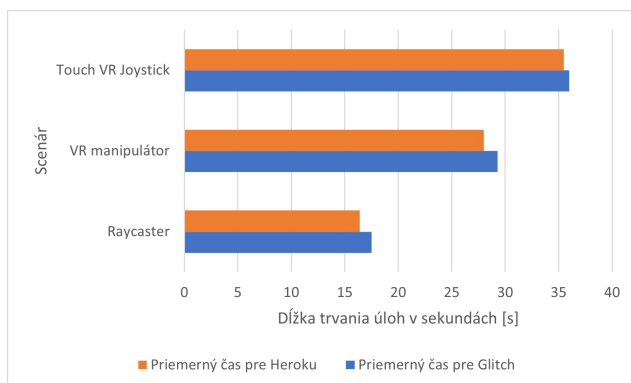
Obr. 12 porovnanie latencií použitím notebooku ASUS FX504

Okrem testovania rýchlosti odozvy servera, bolo navrhnuté riešenie testované reálnymi používateľmi. Skúmala sa rýchlosť, akou dokážu splniť zadané úlohy pre tieto tri typy ovládania objektov vo virtuálnom priestore.

Prvý scenár (Raycaster) bol zameraný priamo na interakciu kurzora. V scéne sa postupne zobrazovali 4 farebné kocky na, ktoré musel používateľ ukázať kurzorom a kliknúť, po prevedení kliku nad objektom kocky sa zobrazila ďalšia až pokiaľ sa nezobrazili všetky 4. Scenár sa považuje za kompletný ak používateľ klikne na poslednú zobrazenú kocku. V Druhom scenári (VR Manipulator) mal používateľ za úlohu pomocou tohoto ovládača navigovať kocku do vyznačenej oblasti scény. Keď používateľ dopravil kocku do určeného miesta, scenár sa považoval za splnený.

Posledný spôsob ovládania (Touch VR Joystick), v ňom mal používateľ pohybovať avatarom po vopred vyznačenej trase až do cieľového bodu. Scenár sa považuje za splnený, keď sa avatar dostal do koncového bodu.

Výsledkom experimentu bolo, že čas odozvy vplýva na rýchlosť akou dokáže používateľ splniť definované úlohy. Tabuľka rýchlosti splnenia daných úloh je zobrazená na Obr. 13. Využitie inteligentného telefónu ako ovládača v režime „slave“ sa javí ako vhodné riešenie v prípade ak je možné dosiahnuť vizualizáciu používateľskej akcie serverom pod hranicu 100ms.



Obr. 13 Dĺžka riešenia zadaných úloh v závislosti od cloudovej služby

C. Používateľské testovanie (SUS metrika)

Navrhované riešenie bolo tiež pripravené na testovanie s reálnymi používateľmi. Cieľom bolo získať reálnu spätnú väzbu od používateľov, ich pripomienky, postrehy a možnosti vylepšení v ďalších iteráciách. Z veľkej časti toto testovanie bolo obmedzené pandemickou situáciou vírusu SARS-CoV-2 a obmedzeniami zhromažďovania sa štátom Slovenskej republiky. Vzniknutá situácia obmedzila používateľské testovanie, keďže sa nemohlo uskutočniť osobne a teda vo väčšom počte testovať na zariadení Microsoft HoloLens.

Testovanie prebiehalo formou video hovoru. Používateľom bola na začiatku testovania odprezentovaná funkcionálna a zameranie testovaného systému. Následne si mohli používatelia samostatne vyskúšať systém a prezrieť rozmiestnenie elementov v SMART prostredí. Následne sa prechádzali body scenáru a meral sa čas dĺžky vykonávania jednotlivých úloh, po skončení testovania prebehla s testermi diskusia. Body testovaného scenára pre virtuálnu realitu sú:

1. prihlásiť sa prostredníctvom údajov meno - test, a heslo - heslo,
2. zasvietiť svetlo v obývacej miestnosti,
3. nastaviť výkon mrazničky na mínus 10 stupňov Celzia,
4. prepnúť TV kanál na STV 1,
5. hlasovým povelením otvoriť menu sporáku (show stove),
6. nastaviť a spustiť časovač mikrovlnnej rúry na 45 sekúnd a zastaviť časovač presne na 30 sekundách,
7. zistiť informácie o počasí,
8. odhlásiť sa so systémom.

Otázky v štandardnom SUS dotazníku sú nasledovné:

1. Myslím si, že by som používal tento systém často.
2. Systém sa mi zdal zbytočne zložitý.
3. Myslím si, že systém je ľahko použiteľný.
4. Myslím si, že aby som mohol používať tento systém, potreboval by som podporu technicky zdatného človeka.
5. Zistil som, že rôzne funkcie v tomto systéme sú dobre integrované.
6. Myslel som si, že v tomto systéme je príliš veľa nezrovnalostí.
7. Vedel by som si predstaviť, že väčšina ľudí by sa naučila používať tento systém veľmi rýchlo.
8. Systém sa mi zdal veľmi ťažkopádny.
9. Pri používaní systému som sa cítil veľmi sebaisto.
10. Musel som sa naučiť veľa vecí, skôr ako som mohol začať pracovať s týmto systémom.

Testovanie sa podarilo uskutočniť na vzorke 20 ľudí vo veku od 22 do 60 rokov.

Tab. 2 Priemerné výsledky SUS dotazníka

otázka	Priemerné hodnotenie
1	4,25
2	1,5
3	4
4	1,5
5	3,75
6	2
7	4,25
8	2
9	4,25
10	2

Tab. 3 Priemerný čas vykonávania úloh

úloha	Priemerný čas používateľov (s)
1	0:42
2	0:27
3	0:42
4	0:38
5	0:20
6	0:41
7	0:15
8	0:47

Plnenie úloh testovacieho scenára prebiehalo bez väčších komplikácií a nedorozumení. Napriek tomu je ťažké z tak malej testovacej vzorky vyvodit' konštruktívne závery. No aj napriek tomu by bolo niektoré z vyplývajúcich pripomienok vhodné zaradiť do ďalšej iterácie systému. Je to hlavne návrh lepšieho algoritmu pre zobrazovanie používateľského rozhrania zariadenia, nakoľko sa mnohokrát stalo to, že sa rozhranie načítalo cez kurzor používateľa a ten vyvolal nechtiac akciu kliku nad elementom, kde kurzorom ukazoval počas načítavania rozhrania. Taktiež pri využívaní elementov typu button, ktorých akciu je možné vyvolávať opakovane počas držania kurzoru nad elementom, by bolo vhodné predĺžiť interval jednotlivých volaní, nakoľko s touto funkcionalitou nevedeli používatelia časovo presne fungovať, niektorých prekvapilo to, že sa hodnota menila počas súvislého pohľadu na element a mali tendenciu po každom „kliku“ hýbať kurzorom preč z elementu a späť. Taktiež pri graficky menších elementoch sa stávalo to, že ich používatelia nevedeli rýchlo zamerať a teda ich „klikanie“ trvalo dlhšie. Vhodným riešením

by teda bolo zväčšiť hitzóny malých elementov, prípadne nastaviť minimálnu hitzónu pre každý objekt vo virtuálnom priestore. Táto situácia nastávala najčastejšie pri zariadení lampy alebo vyvolaní informácií o počasi. Taktiež komunikácia pomocou hlasových povelov nebola bezproblémová a používatelia mali problém takto ovládať systém. 14 účastníkov malo problém pochopiť zadávanie týchto povelov, aj keď sú zobrazené v ľavom hornom rohu používateľského rozhrania. Väčšinou išlo o starších používateľov a taktiež im zadávanie komplikoval anglický jazyk, kde samotná výslovnosť nebola vždy správne rozpoznaná. Navrhne sa aby sa systém dal plne ovládať iba pohľadom a hlasové povely by predstavovali iba alternatívnu možnosť ovládania, nie nosnú časť pre niektoré typy prístupov, ako je panel s hlavným menu aplikácie.

Výsledky SUS metriky boli v priemere 78.75 bodov. Výsledná hodnota, aj keď na malej vzorke používateľov, sa považuje za dostatočujúcu a nadpriemernú nakoľko hranica priemeru je 68 bodov. Väčšina účastníkov prejavila o testované rozhranie veľký záujem a vedľa si predstaviť koncept jeho používania v budúcnosti, keď sa hardvér zariadenia a dátových prílieb pre zmiešanú realitu posunie na ďalšiu úroveň.

VI. ZÁVER

Táto práca sa venovala tvorbe a návrhu inteligentných používateľských rozhraní so zameraním na virtuálnu a zmiešanú realitu, s prevedením objektívnych a subjektívnych testov. Zameranie bolo taktiež na komunikáciu s rôznymi druhmi SMART zariadení so zameraním na zjednotenie ich ovládacích prvkov do jedného unifikovaného používateľského rozhrania s rovnakým prístupom k dátam.

Výsledok testovania v zdieľanom prostredí je vo väčšine testovacích scenárov pri použitých komponentoch dostatočná. Scenáre, kde bol testovaný výkon na smartfóne alebo počítači, poskytovali dostatočujúci používateľský zážitok. Horšie výsledky boli namerané pri testovaní HoleLens, kde už pri 4 ďalších používateľoch začal výkon rapídne klesať na hodnotu okolo 16 FPS a pri 9 ďalších používateľoch klesol na hodnotu 6 FPS. Testovať HoleLens pri viacerých používateľoch už nemalo význam a prinieslo by to iba horšie výsledky. Tento problém by bolo možné vyriešiť novším a výkonnejším hardvérom zariadenia HoleLens 2. generácie.

Ďalší test, ktorý meral latencie systému pri použití rôznych typov ovládania je možné považovať za úspešný. Používatelia slovné subjektívne kladne ohodnotili systém a radi by tento koncept používali aj ďalej. Pri testoch bolo možné stabilne držať hodnotu odozvy systému pod 100 ms čo je podľa [12] hraničná hodnota pre plynulý používateľský zážitok.

Posledný test, ktorý sa zamerával na testovanie definovaného scenára použitia a metriky SUS je taktiež možné považovať za úspešný. Hodnota z SUS bola 78.75 bodu, hodnoty nad 68 sú považované za dostatočné. Z vyhodnotenia SUS a z rozhovorov s testermi vyplýva, že má zmysel podobný koncept ovládania vo virtuálnej a zmiešanej realita naďalej rozvíjať.

ACKNOWLEDGMENT

Táto práca bola podporená nasledujúcimi projektami:

- APVV grant no. APVV-16-0202 “Enhancing cognition and motor rehabilitation using mixed reality”
- KEGA grant No. 035TUKE-4/2019: „Virtual-reality technologies and handicapped people education“.

REFERENCES

- [1] Hudák, Martin Sivý, and Branislav Sobota. Uniform smartphone controller for web-based virtual reality purposes. In *Acta Electrotechnica et Informatica*. - Košice (Slovensko) : Fakulta elektrotechniky a informatiky., 2021.
- [2] Haythem Bahri, David Krčmařík, and Jan Kočí. Accurate object detection system on hololens using yolo algorithm. In *2019 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO)*, pages 219–224. IEEE, 2019.
- [3] Alessandro Farasin, Francesco Peciarolo, Marco Grangetto, Elena Gianaria, and Paolo Garza. Real-time object detection and tracking in mixed reality using microsoft hololens. In *15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2020*, volume 4, pages 165–172. SciTePress, 2020.
- [4] M Gomathy Nayagam and K Ramar. Reliable object recognition system for cloud video data based on ldp features. *Computer Communications*, 149:343–349, 2020.
- [5] Anderson Augusto Simiscuka and Gabriel-Miro Muntean. Synchronisation between real and virtual-world devices in a vr-iot environment. In *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5. IEEE, 2018.
- [6] Aws Alaa Zaidan, Bilal Bahaa Zaidan, MY Qahtan, OS Albahri, AS Albahri, Mussab Alaa, Fawaz Mohammed Jumaah, Mohammed Talal, Kian Lam Tan, WL Shir, et al. A survey on communication components for iot-based technologies in smart homes. *Telecommunication Systems*, 69(1):1–25, 2018.
- [7] Omar Hamdan, Hassan Shanableh, Inas Zaki, AR Al-Ali, and Tamer Shanableh. Iot-based interactive dual mode smart home automation. In *2019 IEEE international conference on consumer electronics (ICCE)*, pages 1–2. IEEE, 2019.
- [8] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. International Society for Optics and Photonics, 1995.
- [9] G Margetis, G Papagiannakis, and Constantine Stephanidis. Realistic natural interaction with virtual statues in x-reality environments. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2019.
- [10] Štefan Korečko, Marián Hudák, Branislav Sobota, Martin Sivý, Matúš Pleva, and William Steingartner. Experimental performance evaluation of enhanced user interaction components for web-based collaborative extended reality. *Applied Sciences*, 11(9):3811, 2021.
- [11] Marián Hudák, Štefan Korečko, and Branislav Sobota. Lirkis global collaborative virtual environments: Current state and utilization perspective. *Open Computer Science*, 11(1):99–106, 2021.
- [12] RAAEN, K. – KJELLMO, I.: Measuring latency in virtual reality systems. In: *International Conference on Entertainment Computing*. Springer, Cham, 2015. p. 457–462.

Metódy strojového učenia na rozpoznávanie objektov v počítačovom videní

Martin Štancel

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
martin.stancel@tuke.sk

Liberios Vokorokos

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovensko
liberios.vokorokos@tuke.sk

Abstrakt—Práca sa zaoberá analýzou oblastí softvérového inžinierstva, ktoré riešia moderný trend v rámci využívania neuronových sietí na klasifikáciu obrazov a detekciu objektov. Preto konkrétne v práci sú analyzované umelé neuronové siete, ktoré sa v posledných dekádach ukázali ako najvhodnejšia metóda pri klasifikácii obrazov a detekcii objektov. Využívajú metodiky strojového učenia na prácu s obrovským množstvom dát a počítačového videnia, ktorá je najdôležitejšia súčasť pri získavaní dát z digitálnych obrazových informácií a videosekvencií. Boli spomenuté aj v súčasnosti významné metódy na prácu s dátami v tejto nami riešenej doméne. Práca ďalej rieši potrebu vytvorenia nového prístupu na tvorbu obrazových dátových sád, kde generuje potrebné dáta na obohacovanie pôvodných obrazových dátových sád alebo vytváranie úplne nových obrazových dátových sád. Tento prístup je dôležitý hlavne z pohľadu časovej zložitosti a nepraktickosti práve pri tvorbe takýchto dátových sád, kde označovanie a lokalizovanie objektov záujmu je robené prevažne manuálne. Z toho dôvodu opisuje vznik novej metódy automatizovaného generovania potrebných údajov na vytváranie obrazových dátových sád, ktoré sú nevyhnutné na tréningový proces umelých neuronových sietí a následné vytváranie tréningových modelov z nich pre dosiahnutie korektných výsledkov na objektových detektoroch. V závere práce je porovnaná súčasná obrazová dátová sada COCO s rovnakou obrazovou dátovou sadou obohatenou o nové dáta pomocou vytvorenej metódy, kde sú spísané aj vyhodnotenia a prínosy vytvoreného riešenia.

Kľúčové slová—*umelé neuronové siete, klasifikácia obrazu, detekcia objektov, segmentácia obrazu, obrazová dátová sada, počítačové videnie, strojové učenie, generovanie dát*

Abstract— The work deals with the analysis of such software engineering areas that are connected with the modern trend in the use of neural networks for image classification and object detection. Therefore, specifically in the work, artificial neural networks are analyzed, which in recent decades have proven to be the most suitable method for the image classification and the object detection. They use machine learning methodologies to work with vast amounts of data and computer vision, which is the most important part of extracting data from digital image information and video sequences. Current important methods for working with the data like that in this domain have also been mentioned. The work further addresses the need to create a new approach to the creation of image datasets, where it generates the necessary data to enrich the original image datasets or create

completely new image datasets. This approach is important mainly from the point of view of time complexity and impracticality in the creation of such datasets, where the labeling and localizing of objects of interest is done mainly manually. For this reason, it describes the emergence of a new method of automated generation of data that are necessary for the creation of image datasets, which are later used for the training process of artificial neural networks and subsequent creation of training models from them to achieve correct results on object detectors. At the end of the work, the current image dataset COCO is compared with the same image dataset that have been enriched with new data obtained from the created method, where the evaluations and benefits of the created solution are also written.

Keywords— *artificial neural network, image classification, object detection, image segmentation, pattern recognition, image dataset, computer vision, machine learning, data generation*

I. ÚVOD

V posledných dvoch dekádach vedci a výskumníci v oblasti počítačového videnia, strojového učenia a neuronových sietí pociťujú veľký nárast popularity týchto odvetví. Stalo sa tak najmä preto, že technologicky sa hardvérové, ale aj softvérové predpoklady dnešných počítačov posunuli výrazne vpred, čo umožnilo vykonávať rozsiahle algoritmické operácie a prácu s gigantickými dátovými sadami.

Boli analyzované umelé neuronové siete, čo je vlastne podoblasť strojového učenia, ktoré sú v dnešnej dobe najvhodnejšou metódou (angl. state-of-the-art) pri klasifikácii obrazu a rozpoznávaní objektov z naučených, natrénovaných dátových sád.

Učenie takejto neuronovej siete sa spolieha na dátové sady, z ktorých čerpá dáta, a na základe ktorých potom vie pracovať s novými, dovtedy pre ňu neznámymi, dátami. Umelé neuronové siete využívajú metodiky strojového učenia a počítačového videnia. Počítačové videnie súvisí s procesom spracovania obrazu v tom zmysle, že taktiež rieši znižovanie šumu, zmenu jasu alebo vylepšenie obrazu rôznymi technikami. Na druhej strane, strojové učenie je flexibilné, pretože môže byť použité buď v počítačovom videní, spracovaní obrazu alebo iných oblastiach informatiky. V krátkosti jednotlivé odvetvia riešia:

- *Predspracovanie obrazu:* cieľom je zlepšiť alebo komprimovať obrazové alebo video informácie. Používa operácie s pixelmi na aplikovanie rôznych transformácií ako napríklad rotácia alebo zmena veľkosti, no v konečnom dôsledku sa z toho nevie vyvodit' nejaká zmysluplná informácia pre ďalší vývoj.
- *Počítačové videnie:* cieľom je získať čo najväčší počet informácií z obrazu alebo videosekvencie. Neriešia sa tu len operácie s pixelmi, ale aj oveľa zložitejšie, komplexnejšie a výpočtovo náročnejšie úlohy ako je napríklad práca s detektormi na rozpoznávanie objektov. Ide o určenie konkrétnych objektov alebo rozpoznávanie určitých vzorov v scéne.
- *Strojové učenie:* cieľom je optimalizovať diferenciálne parametre, aby sa minimalizovala určitá funkcia strát/nákladov. Funguje na báze pamätania si predchádzajúcich, naučených situácií, na čo využíva v dnešnej dobe umelé neurónové siete a dátové sady.

Vzťah medzi týmito oblasťami môže byť dosť zložitý. Napríklad konvolučné neurónové siete používajú všetky tri techniky: konvulcia je zo spracovania obrazu, pretože pracuje na báze pixelov, potreba extrahovať obrazový obsah je z počítačového videnia, zatiaľ čo parametre jadra (aplikácia rôznych filtrov) sú upravené pomocou metód strojového učenia.

V práci boli taktiež opísané najdôležitejšie prínosy v oblasti vývoja konvolučných neurónových sietí. Tento typ sietí sa stal najpopulárnejším v oblasti klasifikácie obrazov a detekcii objektov. Boli analyzované jej najrýchlejšie algoritmy a detektory, ktoré sú založené práve na takýchto sieťach.

V neposlednom rade sa pri vývoji navrhovanej metódy v práci využíva aj segmentácia obrazu, ktorá bližšie určuje rozpoznané objekty na obraze na úrovni pixelov. Táto technológia umožňuje detegovať objekty v presnej miere, v akej sa na obraze vyskytujú z rôznych oblastí záujmu. V metóde, ktorá je opísaná v práci, sa segmentácia obrazu používa pri spracovaní dynamických obrazových informácií, kde je pozadie snímaného obrazu (definované snímaním statickej kamery bez objektov v popredí) a popredie, kde sú určené pohybujúce sa objekty, ktoré chceme segmentovať, rozpoznávať a zaznamenávať.

Na základe analyzovaných faktov, dosiahnutých záverov a skúmaných dát bola v tejto práci navrhnutá a implementovaná metóda zlepšovania alebo vytvárania nových natrénovaných modelov za pomoci vygenerovaných dát práve z tejto metódy. Takýto prístup výrazne urýchľuje tvorbu obrazových dátových sád, ktoré sú kritické pre tréningovú fázu neurónových sietí na rozpoznávanie objektov. Zlepšovanie obrazových dátových sád je v tejto práci opísané ako obohacovanie pôvodných obrazových dátových sád, čo neskôr v hodnotení vytvorenej metódy prinieslo pozitívne výsledky.

Metóda je vhodná nielen na obohacovanie pôvodných obrazových dátových sád, ale aj na vytváranie úplne nových, doménovo-spezifických dátových sád, ktoré budú neskôr

použitú na vytvorenie tréningových modelov pre následné rozpoznávanie objektov v určenej doméne detektormi.

II. KLASIFIKÁCIA OBRAZU, ROZPOZNÁVANIE OBJEKTOV A SEGMENTÁCIA OBRAZU

V našom výskume sa budeme zaoberať automatizovanou tvorbou generovania doménovo-spezifických dátových sád, ktorá je úzko spojená s klasifikáciou obrazu a detekciou objektov, pričom konvolučné neurónové siete [1][2] (angl. convolutional neural networks) sa ukázali byť najpoužívanejšou a najlepšou metódou pri klasifikácii obrazu a detekcii objektov. Využíva vrstvy [3], rôzne špecifické operácie [4][5][6]. Tieto siete sa rapídne vyvíjajú a celý prehľad ich vývoja je zhrnutý v nasledujúcej publikácii [7].

Klasifikácia obrazu zahŕňa predvídanie triedy jedného objektu v obraze. Lokalizácia objektu sa týka identifikácie umiestnenia jedného alebo viacerých objektov na obraze a nakreslenia ohraničujúceho boxu okolo ich rozsahu. Detekcia objektov kombinuje tieto dve úlohy a lokalizuje a klasifikuje jeden alebo viac objektov v obraze [8][9].

Vo všeobecnosti, týmito oblasťami softvérového inžinierstva sa zaoberá počítačové videnie. Počítačové videnie (angl. computer vision) je časťou softvérového inžinierstva, ktorá sa zaoberá spracovaním obrazu. Je to tak rozsiahla téma, že každý si jej definíciu interpretuje podľa toho, aký konkrétny problém v rámci počítačového videnia rieši. Počítačové videnie je odvetvie výpočtovej techniky a vývoja softvéru, ktoré sa zaoberá vytváraním zariadení a metód na to, aby boli schopné získavať informácie z obrazu alebo videosekvencie [10].

Klasifikácia obrazu je jednou z najpopulárnejších úloh algoritmov počítačového videnia a strojového učenia, čiže určenie toho, ktorý objekt sa na obraze nachádza [11]. Klasifikácia obrazu sa zvyčajne vzťahuje na obrázky, na ktorých sa objaví a analyzuje práve jeden objekt. Naproti tomu detekcia objektov zahŕňa úlohy klasifikácie aj lokalizácie a používa sa na analýzu realistickejších prípadov, keď v obraze môže existovať viac objektov.

Historicky najznámejšie konvolučné neurónové siete sú AlexNet [2], ZF Net [12], GoogLe Net [13], VGG Net [14], ResNet [15].

Detekcia objektov je definovaná ako podmnožina klasifikácie obrazu. Zatiaľ čo klasifikáciu obrazu rieši len klasifikovanie objektov na obraze, detekcia rieši aj ich lokalizáciu (Obrázok 2). Najpoužívanejšie algoritmy v rámci detekcie objektov sú R-CNN [16], Fast R-CNN [17], Faster R-CNN [18], SSD [19] alebo YOLO [20].

Pri analýze obrazov v počítačovom videní sa ľudia často zaujímajú iba o určité časti obrazovej informácie. Tieto časti sa často nazývajú ciele alebo popredia a vo všeobecnosti zodpovedajú konkrétnym jedinečným oblastiam obrazu. S cieľom identifikovať a analyzovať takýto cieľ je potrebné tieto príslušné oblasti oddeliť a extrahovať. Segmentácia obrazu sa týka techniky a procesu rozdelenia obrazu na charakteristické oblasti a extrakcie predmetov záujmu (chcených objektov).

Existuje množstvo techník segmentovania od jednoduchších, klasických prístupov po zložité, ktoré využívajú

aj spomínané neurónové siete a vyžadujú veľkú výpočtovú silu. Segmentácia obrazu sa ďalej klasifikuje do skupín [21], ktorými sú sémantická segmentácia a segmentácia inštancií. Nasledujúci obrázok (Obrázok 1) znázorňuje hlavné rozdiely medzi spomenutými skupinami segmentácie obrazu.



Obrázok 1. Skupiny segmentácie obrazu

Pri segmentácii obrazu chceme zistiť, ktoré pixely patria presne do určitej triedy pozorovaného objektu. V podstate ide o vymedzenie hraníc zmysluplných objektov. Všetky tieto oblasti, ktoré sa nachádzajú v rámci hranice, by teda mali mať rovnaké označenie a patria do rovnakej triedy. Každý pixel teda získa sémantickú triedu, a potom sa generujú pixelové štítky. Princíp takéhoto fungovania segmentácie je taktiež úzko spätý so spracovaním zvuku, kde sa analogicky deje niečo podobné pri analyzovaní spektrogramov. Pri hodnotení výkonnosti sémantickej segmentácie existujú dve hlavné kritériá: presnosť, alebo inak povedané, úspešnosť algoritmu; a výpočtová zložitosť, pokiaľ ide o požiadavky na rýchlosť a pamäť. V tejto časti sú zanalyzované tieto dve kritériá osobitne. Meranie výkonnosti segmentácie môže byť komplikované hlavne preto, že sa dajú merať dve odlišné hodnoty. Prvou je klasifikácia, ktorá jednoducho určuje pixelové štítky tried a druhá je lokalizácia. Toto meranie je založené na presnosti pixelov [22] a priesečníkom množiny v rôznych variantách [23].

Segmentácia inštancií je náročná výpočtová úloha, pretože vyžaduje správne oddelenie všetkých objektov v obraze a zároveň sémantické segmentovanie každej inštancie na úrovni pixelov. Rozdiel medzi týmito spôsobmi segmentácie je možné vidieť na predchádzajúcom obrázku (Obrázok 1). Objekty na obraze patria do pevnej skupiny sémantických kategórií, počet inštancií sa však líši. Výsledkom je, že sémantická segmentácia sa dá ľahko formulovať ako problém s hustou klasifikáciou podľa pixelov, zatiaľ čo v rámci segmentácie inštancií je náročné predpovedať popisy priamo podľa rovnakej paradigmy [24].

III. DÁTOVÉ SADY

Dátová sada je skupina štruktúrovaných údajov, z ktorej môžeme získavať jednotlivé entity dátovej sady pomocou konkrétnych štruktúrovaných inštrukcií pre dátovú sadu. Majú tieto vlastnosti [25]:

- *štruktúrované*: údaje musia byť štruktúrované, a preto musí byť vzťah medzi prvkami jasne a jednoznačne určený,
- *obsahujú dáta*: znamená to, že obsah je digitálne kodifikovaný,
- *majú odkazy alebo inštrukcie*: dátové sady môžu byť online dostupné cez odkazy alebo programovo

dostupné ako celok, do ktorého sa dostaneme inštrukciami,

- *tvoria celok*: dátové sady sú vtedy, ak k nim vieme prísť ako k celku. Ak prístup k dátovej sade nie je úplný, vtedy nemôžeme hovoriť o dátovej sade,
- *sú aktualizované*: dátové sady zvyknú byť pravidelne aktualizované o nové dáta.

Dnes existujú rôzne dátové sady, no my ukážeme práve tie, ktoré sa zaoberajú a sú využívané v nami skúmanej doméne, teda strojovom učení. V strojovom učení, podľa oblasti, do ktorej dátová sada spadá, sa používajú [26]: hudobné [27], obrazové (neskôr podrobnejšie rozobrané v článku), textové, hlasové a ďalšie.

V poslednej dobe sa spopularizovali hlavne obrazové dátové sady, ktoré sú využívané v neurónových sieťach na tréning dát na následnú klasifikáciu obrazu a detekciu objektov (kapitola III).

Dnes už existuje veľké množstvo dátových sád, ktoré sú dostupné pre verejnosť. Napríklad korporátne (COCO [28]) alebo akademické (ImageNet [29]) dátové sady, ktoré sú používané v rôznych súťažach na testovanie stále nových a nových algoritmov, ale aj v iných oblastiach.

Tvorba dátových sád je časovo zložitá operácia, keďže ich význam je nadobudnutý vtedy, ak obsahujú obrovské množstvo dát. Získavanie týchto dát sa robí ručne alebo automatizovane. Konkrétne, obrazové dátové sady, ktoré sa využívajú pri klasifikácii obrazu a detekcii objektov sú vytvárané pomocou označovania objektov v nich a presného lokalizovania, napríklad rámčekom, ako to môžeme vidieť na nasledujúcom obrázku (Obrázok 2).



Obrázok 2. Príklad klasifikácie a lokalizácie objektu na obraze

Práve toto ručné označovanie a lokalizovanie objektov v obrazoch je časovo zložitá operácia, ak chceme takto spracovať tisíce alebo desať tisíce obrazov, a preto komunita ľudí nevytvára nové dátové sady, ale prispieva do už vytvorených dátových sád. Najpoužívanejšie z nich sú: COCO [28], ImageNet [30][31], PASCAL VOC [32] alebo CIFAR-10 [33][34].

Označovanie objektov na obrazoch robí tvorbu dátových sád časovo náročnou úlohou. Ak hovoríme o detekcii objektov podľa Wardena [35], vtedy je 1000 obrazov na jednu triedu (príklad je Obrázok 2, kde trieda objektu je *mačka*) dostačujúci na natrénovanie neurónovej siete na následné detegovanie objektov z vybranej triedy. Ručné označovanie je nepraktické,

respektíve zaberie roky práce na tvorbu jednej konkrétnej obrazovej dátovej sady.

Už vytvorené a najčastejšie používané obrazové dátové sady síce obsahujú státisíce, respektíve milióny obrazov, no majú nevýhodu v tom, že neobsahujú všetky triedy, triedy používané v konkrétnych doménach alebo súčasný stav obsiahnutých obrazov je nedostačujúci v rámci vysoko pravdepodobnostného detegovania určitej triedy objektu.

Cieľom tejto práce je preto navrhnuť novú metódu, ktorá by prispievala k obohacovaniu súčasných obrazových dátových sád automatizovaným generovaním potrebných dát na vytváranie nových natrénovaných modelov pre následne vylepšené detegovanie objektov. Táto metóda taktiež prispieje k redukovaniu značného času pri procese tvorby obrazových dátových sád a zároveň poskytne priestor pre detekciu objektov v doménovo-špecifických prostrediach. Návrh tejto metódy je spísaný v nasledujúcej kapitole (kapitola IV).

IV. NÁVRH VYLEPŠENEJ METÓDY TVORBY OBRAZOVÝCH DÁTOVÝCH SÁD

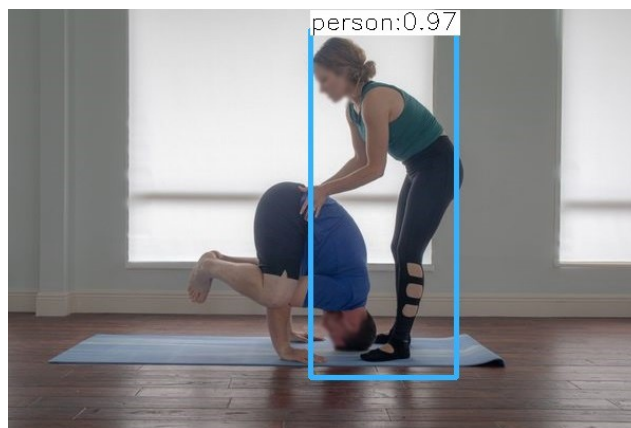
Cieľom tejto práce a je plne automatizovať generovanie potrebných dát na obohatenie existujúcej obrazovej dátovej sady s tým, že človek už nebude potrebný pri jej spracovávaní a tréningovom procese, alebo vytvorenie úplne novej obrazovej dátovej sady s možnosťou ju bližšie doménovo špecifikovať.

Myšlienka vylepšenia a zdokonalenia fázy spracovania a tvorby obrazových dátových sád spočíva v kombinácii detegovania objektov v reálnom čase za pomoci detektora objektov s výberom vhodnej dátovej sady spolu so spracovávaním pohybujúcich sa objektov na statickom plátne (pozadí). Pohybujúce sa objekty na statickom pozadí si môžeme predstaviť ako kameru snímajúcu prostredie, v ktorom sa v reálnom čase pohybujú objekty (ľudia, autá atď.), pričom tie pohybujúce sa objekty tvoria popredie obrazu a sú tzv. „chcené objekty“ (sú súčasťou regions of interest). Formálne zápisy takejto segmentácie a spracovania obrazu v reálnom čase sú špecifikované v implementačnej časti tejto práce.

Z doterajších bádání a experimentov, kde bola navrhnutá metóda na zjednodušenie celkovej tvorby obrazových dátových sád [36] pomocou generovania syntetických obrazov na základe nakonfigurovaných parametrov vyšlo najavo, že síce je možné vytvoriť a uľahčiť proces pri tvorbe takýchto dátových sád, no stále to nie je plne automatizované. V spomenutej práci sa nám podarilo vytvoriť polo automatizovaný systém, kde človek je naďalej potrebný pri určovaní klasifikácie resp. aj lokalizácie daných objektov. Síce celkový čas pri vytváraní takejto dátovej sady sa skrátil, no stále je tam priestor na jeho skrátenie a v konečnom dôsledku aj úplne eliminovanie pomoci človeka pri tvorbe takejto dátovej sady. Výsledky a experimenty polo automatizovaného systému na tvorbu obrazových dátových sád je možné vidieť v spomenutej práci.

Obohatenie existujúcej dátovej sady spočíva v tom, že v priebehu času budú automaticky generované obrázky podľa špecifických kritérií daného detektora a k nemu príslušnej neurónovej siete, kde budú na základe spomenutej kombinácie technológií zozbierané práve tie obrázky, ktoré príslušný detektor na základe danej obrazovej dátovej sady nevie

rozpoznať, respektíve by ich rozpoznať nesprávne alebo s horšou mierou sebavedomia. To bude prebiehať za pomoci segmentácie, ktorá bude oddeľovať pohybujúce sa objekty v popredí snímaného obrazu na statickom pozadí. Pôjde hlavne o objekty v neštandardných polohách pri pohybe, prekrytých častiach objektov a podobne. Príklad je uvedený na nasledujúcom obrázku (Obrázok 3), kde osoba vľavo nebola detegovaná, pretože je vo veľmi neštandardnej polohe a osoba vpravo bola detegovaná s určitým sebavedomím (confidence). Samozrejme môže sa stať, že aj osoba v neštandardnej polohe bude detegovaná správne, no s výrazne menším sebavedomím, ako by bolo očakávané. Navrhovaná metóda má taktiež za cieľ tieto hodnoty sebavedomia detektorov zvyšovať.



Obrázok 3. Detegovanie len jedného z dvoch objektov triedy objektu *person*

Nedetegovaná osoba vľavo z predchádzajúceho obrázka sa v čase pohybovala do takého stavu, že zaujala neštandardnú pozíciu natoľko, že bola ďalej nedetegovateľná detektorom objektov. Skôr než bol dosiahnutý stav na uvedenom obrázku, osoba vľavo bola detektorom rozpoznaná a detegovaná presne tak ako osoba vpravo na obrázku. A presne v tomto spočíva myšlienka zdokonalenia detekcie obrazov dopĺňovaním práve takýchto objektov v neštandardných polohách do existujúcej obrazovej dátovej sady. Pôjde o kombinovanie detekcie objektov cez detektor založený na konvulčných neurónových sieťach, existujúcej obrazovej dátovej sade a segmentácie objektov v pohybe. Úvodný stav osôb z Obrázka 3 bol detegovateľný, pretože osoby neboli výrazne ovplyvnené neprírodnými polohami tela (boli prirodzene postavené).

Zlepšovanie súčasných, respektíve tvorba úplne nových obrazových dátových sád spočíva v snímaní videosekvencie, ktorého fundamentálna zložka je v rámci statického pozadia. To znamená, že videosekvencia je v statickom stave – zaznamenáva rovnaké súradnice s rovnakým pozadím. Avšak veci sa začnú stávať zaujímavými, keď sa pred týmto statickým pozadím začnú objavovať určité predmety, ktoré tam v minulosti neboli. Na predstavu vieme použiť príklad kamery, ktorá sníma chodník v rovnakom bode a chodci sa premávajú po chodníku, pričom vstupujú do snímaného plátna kamery a aj z neho vystupujú. Na takúto kameru vieme aplikovať algoritmus detekcie a lokalizácie objektov s tým, že premávajúce objekty budú neustále rozpoznávané s určitou pravdepodobnosťou rozpoznania (pričom pri súčasných

obrazových dátových sádach to nemusí byť správne rozpoznanie – napríklad pes bude detegovaný ako mačka a podobne). Pri takomto neustálom rozpoznávaní (rozpoznávanie v reálnom čase) môže dôjsť k anomáliám ako sú: odsekuté objekty, prekryté objekty, objekty sú v neštandardných polohách (ako príklad pád chodca pri zakopnutí atď.), nesprávnym určením objektu, správnym určením objektu s malou pravdepodobnosťou rozpoznania alebo dokonca žiadnym určením objektu, ktorý sa reálne na snímke vyskytuje.

Pri všetkých takýchto prípadoch opísaných vyššie môže navrhovaný algoritmus dopomôcť k lepším výsledkom či už pri obohatení pôvodnej obrazovej dátovej sady, alebo vytvorením úplne novej obrazovej dátovej sady, ktorá bude zameraná na určitú špecifickú doménu.

Algoritmus bude spočívať v kombinovaní pôvodného algoritmu detektora s pôvodnou obrazovou dátovou sadou a segmentácie obrazu na rovnakom obraze. Pôvodný detektor bude snímať obrazy a rozpoznávať objekty tak, ako by to robil v akejkolvek inej aplikácii. Avšak kombinovaním detektora a segmentovania obrazu vieme doceliť to, že ak detektor zaznamená nejakú z anomálií, ktoré sú spomenuté vyššie, tak potom pomocou segmentácie vieme takúto snímku ďalej spracovávať a následne uložiť na budúce tréningovanie či už obohateného modelu z pôvodnej obrazovej dátovej sady, alebo úplne nového modelu.

Takýto prístup nám nielenže pokryje nepresnosti, ktoré sa pri detegovaní objektov pri súčasných obrazových dátových sádach vyskytujú, ale takisto nám dáva priestor a nový rozmer na ozajstné automatické generovanie nových obrazových dátových sád. Automatické preto, lebo človek už nebude musieť ručne zaznamenávať akýkoľvek údaj, ktorý je potrebný pri tréningovaní modelu neurónovej siete a teda tvorbe takýchto dátových sád. Navrhovaný algoritmus je transformovaný aj do pseudokódu, ktorý je zobrazený v nasledujúcom fragmente kódu:

```
frames = opencv.inputVideo
for frame in frames:
    segmentation ()
    YOLOdetection ()
    if actualConfidence < settings["allowedConfidence"]:
        similarFrames = compareFrames(frame - 1, frame)
        if similarFrames:
            saveActualFrame (frame)
            saveActualFrameSegmentationData (frame)
            createFiles (frame) //train.txt, frame.txt
```

V. IMPLEMENTÁCIA NAVRHNUTEJ METÓDY

Na základe dôslednej analýzy v predchádzajúcich kapitolách bol pre implementáciu zvolený detektor YOLO, ktorý je spojený s konvolučnou neurónovou sieťou DarkNet [37]. Pohybujúce sa objekty na statickom plátne boli rozpoznávané pomocou segmentácie obrazu v knižnici počítačového videnia OpenCV [38]. Na experimenty a rôzne porovnania s predkladanou metódou bola použitá obrazová dátová sada COCO. V tejto časti práce je bližšie opísaná implementácia navrhovanej metódy. Myšlienka pozostáva z kombinácie detegovania objektov pomocou detektora YOLO a segmentácie obrazu. Tieto dva prístupy detekcie prebiehajú simultánne, pričom sú využité aspekty a výhody oboch algoritmov. Za najvhodnejší jazyk pre takúto kombináciu algoritmov bol zvolený Python a pochopiteľne vývojové

prostredie späté s touto platformou PyCharm. Celá implementovaná metóda pozostáva z niekoľkých programov a jedného konfiguračného súboru.

Segmentácia obrazu sa vypočítava ako absolútny rozdiel všetkých prvkov medzi dvoma poľami alebo medzi poľom a skalárom na základe nasledujúcich vzťahov:

- absolútny rozdiel medzi dvoma poľami, ak tieto polia majú rovnakú veľkosť a typ:

$$dst(I) = saturate(|src1(I) - src2(I)|) \quad (1)$$

- absolútny rozdiel medzi poľom a skalárom, ak sa druhé pole skladá zo skalárov alebo ak má rovnaký počet prvkov ako počet kanálov poľa $src1$:

$$dst(I) = saturate(|src1(I) - src2|) \quad (2)$$

- absolútny rozdiel medzi skalárom a poľom, ak sa prvé pole skladá zo skalárov alebo ak má rovnaký počet prvkov ako počet kanálov poľa $src2$:

$$dst(I) = saturate(|src1 - src2(I)|) \quad (3)$$

prícom I je viacrozmerný index prvkov poľa. V prípade viackanálových polí je každý kanál poľa spracovaný zvlášť. A jednotlivé parametre znamenajú: $src1$ – prvé vstupné pole, resp. skalár – snímok obsahujúci statické pozadie, $src2$ – druhé vstupné pole, resp. skalár – snímok pozostávajúci zo statického pozadia a oblastí záujmu (RoI - Regions of Interest) a dst – výstupné pole, ktoré má rovnakú veľkosť a typ ako vstupné polia – výsledný snímok obsahujúci oblasti záujmu.

V implementácii sú uvedené vzťahy použité v rámci funkcie `absdiff` v knižnici OpenCV takto:

```
dst = cv.absdiff(src1, src2[, dst])
```

Odstraňovanie tieňov v obraze sa vypočítava cez bitovú konjunkciu dvoch polí alebo poľa a skaláru ($dst = src1 \& src2$) pre každý prvok poľa. Formálne to vieme zapísať na základe nasledujúcich vzťahov:

- ak máme dve polia a $src1$ a $src2$ sú polia rovnakej veľkosti:

$$dst(I) = src1(I) \wedge src2(I) \text{ if } mask(I) \neq 0 \quad (4)$$

- pole a skalár, ak pole $src2$ sa skladá zo skalárov alebo má rovnaký počet prvkov ako má $src1$ kanálov - $src1.channels()$:

$$dst(I) = src1(I) \wedge src2 \text{ if } mask(I) \neq 0 \quad (5)$$

- skalár a pole, ak pole $src1$ sa skladá zo skalárov alebo má rovnaký počet prvkov ako má $src2$ kanálov - $src2.channels()$:

$$dst(I) = src1 \wedge src2(I) \text{ if } mask(I) \neq 0 \quad (6)$$

V prípade, že sa jedná o polia typu `float`, ich strojovo špecifická bitová reprezentácia je použitá pre túto operáciu. V prípade viackanálových polí, každý kanál je spracovaný zvlášť. V druhom a treťom prípade spomenutom vyššie je najprv skalár prevedený do poľa. A jednotlivé parametre znamenajú: $src1$ – prvé vstupné pole alebo skalár, $src2$ – druhé vstupné pole alebo skalár, dst výstupné pole rovnakej veľkosti a typu

ako vstupné polia, *mask* – nepovinná operácia. 8-bitové jednonanálové pole, ktoré špecifikuje elementy, ktoré majú byť vo výstupnom poli zmenené. Ak sú ako vstupné polia použité rovnaké snímky a je použitá maska, výsledkom je rovnaký snímok s aplikovanou maskou, čím je možné odstrániť určitý farebný rozsah.

V implementácii sú uvedené vzťahy použité v rámci funkcie *bitwise_and* v knižnici OpenCV takto:

```
dst = cv.bitwise_and(src1 , src2[, dst[, mask ]])
```

Na výpočet masky pri spracovávaní obrazu sa zisťuje, či sa prvky poľa nachádzajú medzi zvyšných dvoch polí, pričom formálne funkcia zisťuje tento rozsah na základe nasledujúcich vzťahov:

- pre každý prvok jednonanálového vstupného poľa:

$$dst(I) = lowerb(I)_0 \leq src(I)_0 \leq upperb(I)_0 \quad (7)$$

- pre dvojkanálove polia:

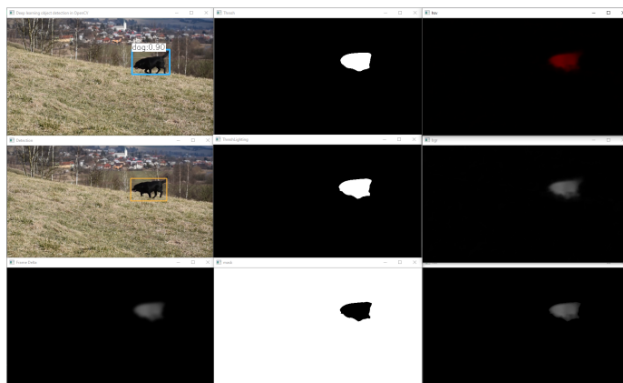
$$dst(I) = lowerb(I)_0 \leq src(I)_0 \leq upperb(I)_0 \wedge lowerb(I)_1 \leq src(I)_1 \leq upperb(I)_1 \quad (8)$$

- a všeobecne pre n kanálove polia:

$$dst(I) = lowerb(I)_{n-1} \leq src(I)_{n-1} \leq upperb(I)_{n-1} \wedge \dots \quad (9)$$

Ak sú spodné a/alebo horné ohraničenia parametrov skaláry, indexy (I) polí *lowerb* a *upperb* vo vzorcoch spomenutých vyššie by mali byť vynechané. A jednotlivé parametre znamenajú: *src* – prvé vstupné pole, *lowerb* – skalár alebo pole obsahujúce spodné ohraničenie rozsahu farby, *upperb* - skalár alebo pole obsahujúce horné ohraničenie rozsahu farby, *dst* – výstupné pole rovnakej veľkosti ako vstupné pole – tvorí masku, ktorá po aplikovaní na snímok za použitia *bitwise_and* odstráni farby dané rozsahom *lowerb* a *upperb*.

Nasledujúci obrázok (Obrázok 4) zobrazuje použitie vyššie opísaných vzťahov.



Obrázok 4. Rôzne druhy vizualizácií spracovania videosekvencie

VI. EXPERIMENTÁLNE OVERENIE IMPLEMENTOVANEJ METÓDY

Samotné experimenty prebiehali na novo natrénovanej obrazovej dátovej sade, ktorá sa skladala z pôvodných obrazoch z dátovej sady COCO a nami vygenerovanými obrazmi z navrhutej a implementovanej metódy. Takúto novú sadu bolo potrebné natrénovať na neurónovej sieti, k čomu poslúžila neurónová sieť DarkNet vo Windowsovom prevedení [39]. V sumáre bolo potrebné najprv:

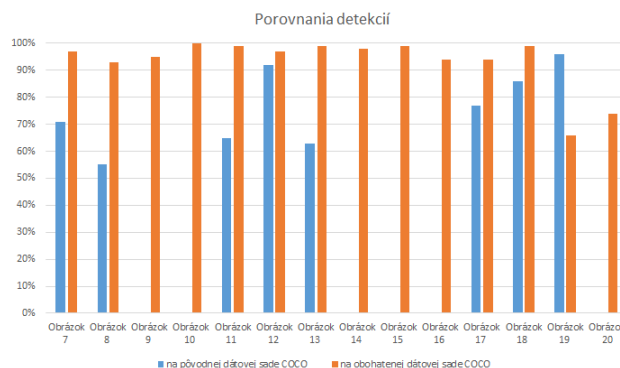
- pomocou implementovanej metódy vygenerovať potrebné súbory na tréningový proces neurónovej siete, aby vytvorila nový model obohatenej dátovej sady,
- pripraviť pôvodné obrázky z dátovej sady COCO cez ich API [40],
- vytvorenie novo natrénovaného modelu na základe pôvodných údajov z COCO a vygenerovaných súborov cez implementovanú metódu.

Experimenty prebiehali v rámci porovnávania originálnej detekcie YOLO na originálnej dátovej sade COCO s novovytvoreným natrénovaným modelom, ktorý obsahoval pôvodné obrázky psov, oviec a medveďov spolu s vygenerovanými obrazmi za použitia implementovanej metódy. Pri experimentoch nastali rôzne kombinácie udalostí, ktoré sú bližšie popísané v jednotlivých nasledujúcich odrážkach odkazujúce sa na obrázky, ktoré znázorňujú opísanú udalosť, pričom každý obrázok má v hornej časti a) detekciu YOLO na pôvodnej obrazovej dátovej sade COCO a v dolnej časti b) detekciu YOLO na obohatenej obrazovej dátovej sade COCO z obrazov vygenerovaných vytvorenou metódou. Orámovania, klasifikácia a detekcia tried objektov je automaticky generovaná, pričom textové údaje z obrázkov v čitateľnejšej forme sú aj v Tabuľke 1 a v grafe na Obrázku 5.

- Obohatená dátová sada pomocou vytvorenej metódy dosahuje vyššiu pravdepodobnosť úspešnej detekcie správnej triedy oproti pôvodnej dátovej sade COCO (Obrázok 7, Obrázok 8).
- Obohatená dátová sada pomocou vytvorenej metódy dosahuje výsledky pri detekcii aj v prípadoch, kde detekcia pri pôvodnej dátovej sade COCO nezaznamenala žiadne triedy objektov (Obrázok 9, Obrázok 10).
- Obohatená dátová sada pomocou vytvorenej metódy túto chybovosť eliminuje a dosahuje správne výsledky s vysokou pravdepodobnosťou výskytu správnej triedy objektu (Obrázok 11, Obrázok 12).
- Novo natrénovaný model za pomoci vytvorenej metódy má vyššiu pravdepodobnosť úspešnej detekcie správnej triedy aj pri odseknutých objektoch, respektíve pri ich dosť nepravdepodobnej polohe oproti pôvodnej dátovej sade COCO (Obrázok 13).
- Obohatená dátová sada pomocou vytvorenej metódy dosahuje výsledky pri detekcii aj v prípadoch, kde detekcia pri pôvodnej dátovej sade COCO nezaznamenala žiadne triedy objektov a to aj na objektoch, ktoré sú neúplné, odseknuté alebo v

neprirodzenej polohe. (Obrázok 14, Obrázok 15, Obrázok 16).

- Novo natrénovaný model za pomoci vytvorenej metódy má správne výsledky, teda správne určenie detegovanej triedy objektu aj pri odseknutých objektoch, respektíve pri ich dost' nepravdepodobnej polohe. Pričom pôvodná dátová sada COCO takého objekty určuje s vysokou pravdepodobnosťou nesprávne. (Obrázok 17, Obrázok 18).
- Obohatená dátová sada pomocou vytvorenej metódy dosahuje správne určenie triedy objektu oproti pôvodnej dátovej sade COCO aj pri odseknutých objektoch, ktoré majú vyše 50% svojho priestoru odseknuté (Obrázok 19).
- Obohatená dátová sada pomocou vytvorenej metódy dosahuje výsledky aj pri odseknutých objektoch, ktoré majú vyše 50% svojho priestoru odseknuté, kde detekcia pri pôvodnej dátovej sade COCO nezaznamenala žiadne triedy objektov (Obrázok 20).



Obrázok 5. Porovnanie detekcií pred a po aplikovaní vytvorenej metódy

V predchádzajúcej tabuľke a na predchádzajúcom grafe je možné vidieť pozitívne výsledky, ktoré nám zabezpečilo obohatenie pôvodnej obrazovej dátovej sady COCO o 28.22% údajov triedy objektu *dog* (z pôvodných 4385 obrazov z triedy objektu *dog* na 6109 (1724 vygenerovaných obrazov)).

Časť experimentov bola zameraná na porovnanie pôvodnej a obohatenej obrazovej dátovej sady (Tabuľka I a Obrázok 5) a časť na generovaní obrazov (Tabuľka II a Obrázok 6).

Generovanie obrazov závisí od kvality spracovávanej videosekvencie a konfiguračných nastavení, ktoré sú späté s konkrétnou situáciou v konkrétnom prostredí. Počet vygenerovaných obrazov závisí od premennej *frameSaveRate*, ktorej minimálna hodnota pri 30 FPS videosekvencii je 0.02 na základe vzťahu:

$$\text{round}(\text{round}(\text{FPS}) * \text{frameSaveRate}) > 0.5 \quad (10)$$

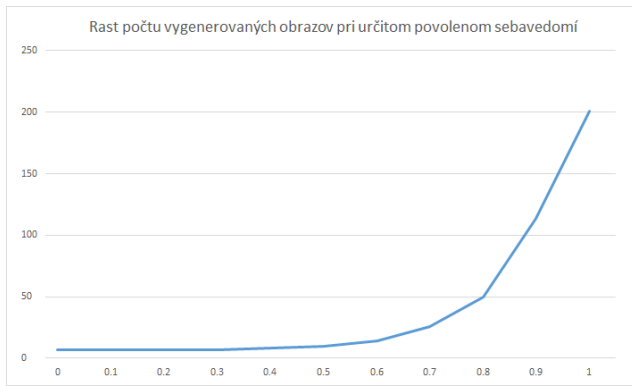
a teda pri približne 7 sekundovej videosekvencii boli vygenerované nasledujúce údaje:

TABUĽKA II GENEROVANIE OBRAZOV PRI 30 FPS A 7 SEKUNDOVEJ VIDEOSEKVENCII

Povolené sebavedomie parameter: <i>allowedConfidence</i>	Počet vygenerovaných obrazov	% z možných vygenerovaných obrazov
1	201	100%
0.9	114	56.72%
0.8	50	24.86%
0.7	26	12.94%
0.6	14	6.96%
0.5	10	4.96%
0.4	8	3.98%
0.3	7	3.48%
0.2	7	3.48%
0.1	7	3.48%
0	7	3.48%

TABUĽKA I ZHRNUTIE ROZDIELOV V DETEKcii NA PÔVODNEJ DÁTOVEJ SADE COCO A OBOHATENEJ DÁTOVEJ SADE COCO VYTVORENOU METÓDOU

Posudzovaný Obrázok	Detekcia na pôvodnej dátovej sade COCO	Detekcia na obohatenej dátovej sade
Obrázok 7	správna - dog - 71%	správna - dog - 97%
Obrázok 8	správna - dog - 55%	správna - dog - 93%
Obrázok 9	žiadna - X - 0%	správna - dog - 95%
Obrázok 10	žiadna - X - 0%	správna - dog - 100%
Obrázok 11	nesprávna - bear - 65%	správna - dog - 99%
Obrázok 12	nesprávna - sheep - 92%	správna - dog - 97%
Obrázok 13	správna - dog - 63%	správna - dog - 99%
Obrázok 14	žiadna - X - 0%	správna - dog - 98%
Obrázok 15	žiadna - X - 0%	správna - dog - 99%
Obrázok 16	žiadna - X - 0%	správna - dog - 94%
Obrázok 17	nesprávna - bear - 77%	správna - dog - 94%
Obrázok 18	nesprávna - bear - 86%	správna - dog - 99%
Obrázok 19	nesprávna - bear - 96%	správna - dog - 66%
Obrázok 20	žiadna - X - 0%	správna - dog - 74%

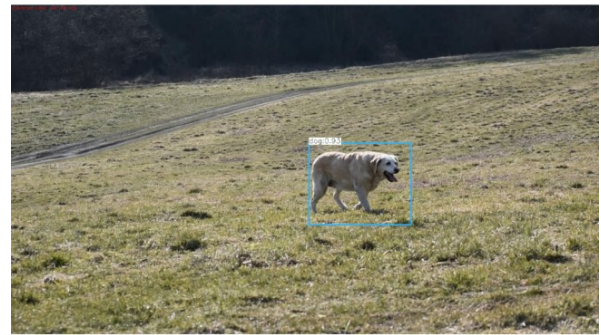


Obrázok 6. Porovnanie generovania obrazov pri určitých povolených sebavedomiach

S vyšším povoleným sebavedomým teda rastie aj počet vygenerovaných obrazov. Údaje vyjadrujú aj to, že so znižujúcim sa povoleným sebavedomým je počet vygenerovaných obrazov menší a teda aj tie vygenerované obrazy sú vzácnejšie, keďže súčasný stav detekcie objektov na pôvodnej obrazovej dátovej sade dosahuje pri pravdepodobnosti určenia triedy objektu od 0% - 90% 56.72% a pri pravdepodobnosti od 0% - 50% len 4.96%.

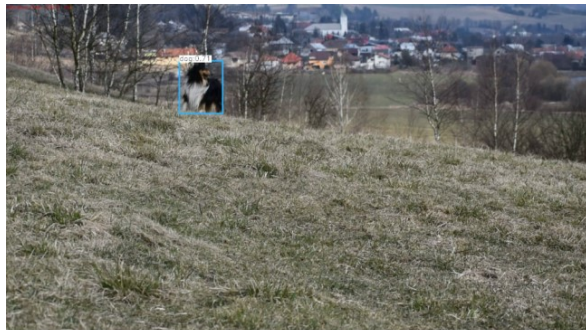


a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

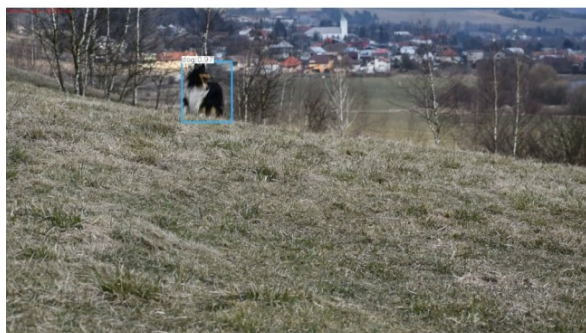
Obrázok 8. Príklad zlepšenia pravdepodobnosti detekcie z a) 55% na b) 93%



a) pôvodná dátová sada COCO



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO



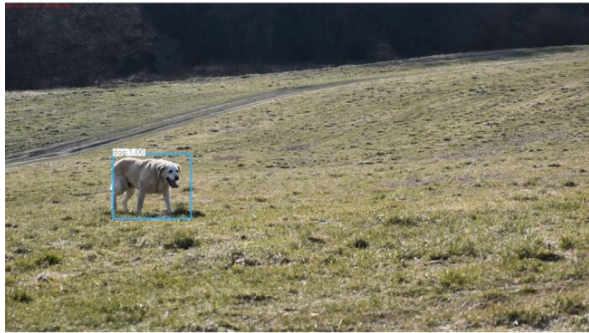
b) obohatená dátová sada COCO

Obrázok 7. Príklad zlepšenia pravdepodobnosti detekcie z a) 71% na b) 97%

Obrázok 9. Príklad, kde pôvodné COCO nezaznamenalo žiadnu triedu objektu



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 10. Príklad, kde pôvodné COCO nezaznamenalo žiadnu triedu objektu

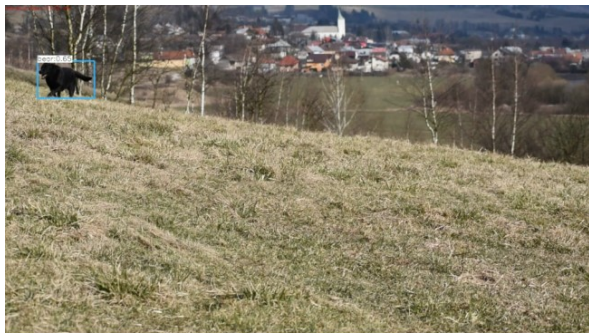


a) pôvodná dátová sada COCO

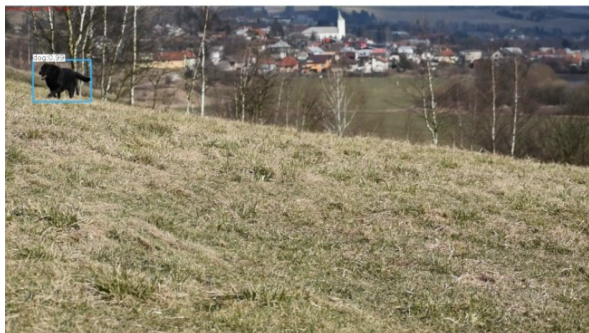


b) obohatená dátová sada COCO

Obrázok 12. Príklad a) nesprávnej detekcie a b) správnej detekcie



a) pôvodná dátová sada COCO

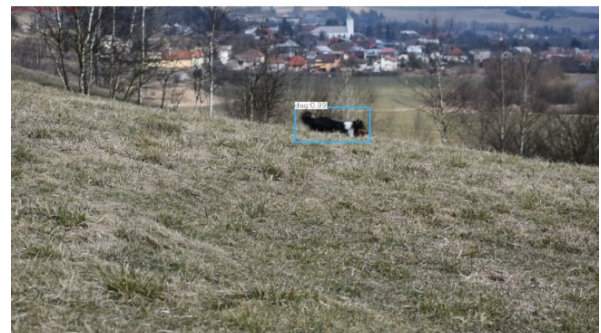


b) obohatená dátová sada COCO

Obrázok 11. Príklad a) nesprávnej detekcie a b) správnej detekcie



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 13. Zlepšenie detekcie z a) 63% na b) 99% pri prekrytom objekte



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 14. Detekcia odseknutého objektu a) žiadna, b) správna

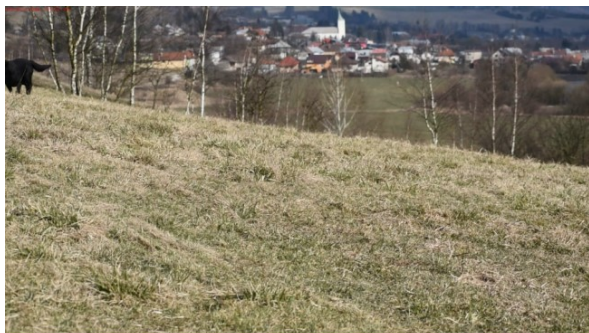


a) pôvodná dátová sada COCO

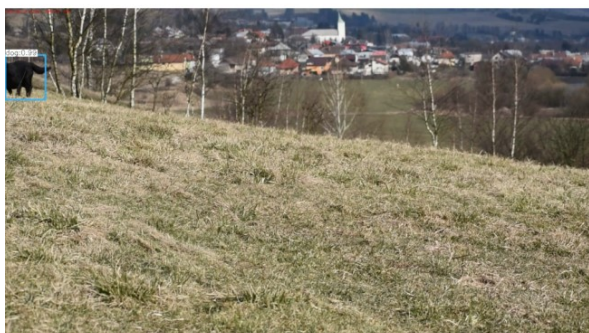


b) obohatená dátová sada COCO

Obrázok 16. Detekcia odseknutého objektu a) žiadna, b) správna



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 15. Detekcia odseknutého objektu a) žiadna, b) správna



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 17. Príklad a) nesprávnej b) správnej detekcie odseknutého objektu



a) pôvodná dátová sada COCO

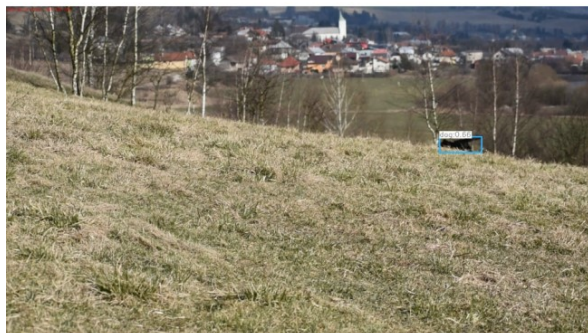


b) obohatená dátová sada COCO

Obrázok 18. Príklad a) nesprávnej b) správnej detekcie prekrytého objektu



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 19. a) nesprávna b) správna detekcia vyše 50% prekrytého objektu



a) pôvodná dátová sada COCO



b) obohatená dátová sada COCO

Obrázok 20. a) žiadna b) správna detekcia vyše 50% odseknutého objektu

ZÁVER

Predložená práca sa na základe dosiahnutých analyzovaných záverov a výskumu zaoberá návrhnutím a implementovaním novej metódy pre automatické generovanie dát potrebných pre tréningový proces neurónových sietí na vytvorenie nových modelov z obrazových dátových sád, ktoré budú následne použité pri detegovaní objektov detektormi. Keďže tvorba úplne novej dátovej sady je časovo náročná operácia, predkladaná metóda umožňuje takéto dátové sady obohacovať o nové údaje a tak zlepšovať ich finálny efekt pri použití na detegovanie objektov.

Na to, aby sme dosiahli požadovaný výskum, bolo potrebné v práci analyzovať informatické oblasti, ktoré sa zaoberajú klasifikáciou obrazu a detekciou objektov.

Najprv boli analyzované umelé neurónové siete, ktoré využívajú metodológie strojového učenia a počítačového videnia na získavanie dát z digitálnych obrazových informácií a videosekvencií.

Neurónové siete sú podoblast'ou strojového učenia, ktoré je v dnešnej dobe považované za najvhodnejšiu metódu pri spracovávaní obrovského objemu dát z dátových sád, a na základe ktorej neurónové siete dosahujú vysokú presnosť pri rozpoznávaní vzorov v neznámych prostrediach.

Učenie takejto neurónovej siete sa spolieha na dátové sady, z ktorých čerpá dáta, a na základe ktorých potom vie pracovať s novými, dovtedy pre ňu neznámymi, dátami.

Predpokladáme, že nami vytvorená metóda bude mať veľký význam pri redukcii času a dokonca aj financií pri tvorbe nových obrazových dátových sád určených na tréningové neurónových sietí, čím sa vytvoria nové modely, ktorých následné použitie je práve pri klasifikácii obrazov a detekcii objektov.

Na základe experimentov, hlavnými výhodami tohto riešenia sú:

- zlepšenie sebavedomia pri správnom rozpoznaní triedy objektu v oboch dátových sádach,
- správne rozpoznanie triedy objektu na obohatenej dátovej sade oproti jej pôvodnej verzii, ktorá nezaznamenala žiadne rozpoznanie,
- správne rozpoznanie triedy objektu na obohatenej dátovej sade oproti nesprávne rozpoznannej triede objektu na pôvodnej verzii,
- implementovaná metóda si poradila aj s odseknutými objektami a rozpoznaná ich správne,
- správne rozpoznanie je aj v situáciách, kde vyše 50% rozpoznávaného objektu je odseknutých alebo prekrytých.

Navrhnutú a implementovanú metódu by sme v budúcnosti chceli použiť aj v ďalších výskumoch a aplikovať jej výsledky v detekcii, ktorá by sa odohrávala v reálnom čase.

Úspešnosť tohto výskumu prináša nasledovné:

- obohatenie pôvodnej (súčasnej) obrazovej dátovej sady na dosahovanie lepších výsledkov pri detekcii objektov,
- zníženie času pri tvorbe obrazovej dátovej sady pre špecifickú doménu,
- automatizáciu spracovávania dát potrebných na tréningový proces neurónovej siete a následnému vytvoreniu modelu na detegovanie objektov,
- keďže sa jedná o novú, vylepšenú metódu tvorby obrazových dátových sád, tak na základe nej bude možné vytvoriť nové prístupy a techniky, ktoré túto oblasť posunú ešte ďalej.

V budúcnosti by sa tento proces mohol automatizovať natoľko, žeby bolo k nemu vytvorené príslušné API a komunita používateľov by prispievala do špecifických domén vlastné vygenerované dáta, kde by sa takto zozbierali, a v konečnom dôsledku by sa tak vytvorila gigantická obrazová dátová sada, slúžiaca na vytváranie konkrétnych tréningových modelov na následnú detekciu objektov v doménovo-špecifickej oblasti.

Na záver by som dodal, že tvorba obrazových dátových sád je serióznym problémom súčasnosti, keďže sa využívajú prevažne manuálne techniky v označovaní a lokalizovaní objektov na obrazoch. Touto metódou bola docieľená automatizácia generovania potrebných dát pre vytváranie nových obrazových dátových sád, kde tieto manuálne, práčne a časovo náročné úlohy sú eliminované. Tento prístup umožňuje zľahčovať aj prácu výskumníkom v rôznych oblastiach, kde

momentálne využívajú limitované obrazové dátové sady tak, ako je to napríklad spomenuté aj v tejto publikácii [41] v poľnohospodárskom sektore.

POĎAKOVANIE

Táto práca bola podporená KEGA projektom 002TUKE-4/2021 Implementácia moderných metód a foriem výučby v oblasti kyberbezpečnosti k požiadavkám praxe.

LITERATÚRA

- [1] A. Krizhevsky, I. Sutskever, G. E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks." In: Advances in Neural Information Processing Systems (NIPS), Vol. 25, 2012, pp. 1097-1105.
- [2] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R. R. Salakhutdinov "Improving neural networks by preventing co-adaptation of feature detectors." Cornell University, 2012, arXiv:1207.0580
- [3] Y. Lecun, Y. Bengio, G. E. Hinton "Deep learning." In: Nature, 2015, pp. 436-444.
- [4] S. Albawi, T. A. Mohammed, S. Al-Zawi "Understanding of a convolutional neural network." In: 2017 International Conference on Engineering and Technology (ICET), Turkey, 2017, pp. 264-270.
- [5] V. Nair, G. E. Hinton "Rectified Linear Units Improve Restricted Boltzmann Machines." In: Proceedings of the 27th International Conference on Machine Learning (ICML), 2010, pp. 807-814.
- [6] N. Kalchbrenner, E. Grefenstette, P. Blunson "A Convolutional Neural Network for Modelling Sentences." In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, USA, 2014.
- [7] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen "Recent advances in convolutional neural networks." In: Pattern Recognition, Elsevier, Vol. 77, 2018, pp. 354-377.
- [8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei "ImageNet Large Scale Visual Recognition Challenge." In: International Journal of Computer Vision, Vol. 115, No. 3, 2015, pp. 211 - 252.
- [9] D. Lu, Q. Weng "A survey of image classification methods and techniques for improving classification performance." In: International Journal of Remote Sensing, Taylor and Francis, Vol. 28, No. 5, 2007, pp. 823-870.
- [10] R. Szeliski "Computer Vision: Algorithms and Applications." Washington, USA, Springer, 2010, ISBN 978-1-84882-934-3, pp. 3-10s.
- [11] M. Korytkowski, L. Rutkowski, R. Scherer "Fast image classification by boosting fuzzy classifiers." In: Information Sciences, Elsevier, 2016, Vol. 327, pp. 175-182.
- [12] M. D. Zeiler, R. Fergus "Visualizing and Understanding Convolutional Networks." In: European Conference on Computer Vision, Springer, 2014, pp. 818-833s.
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich "Going Deeper with Convolutions." In: Computer Vision and Pattern Recognition (CVPR), IEEE Xplore, 2014, pp. 1-9.
- [14] K. Simonyan, A. Zisserman "Very Deep Convolutional Networks For Large-Scale Image Recognition." In: International Conference on Learning Representations, Computer Vision and Pattern Recognition, 2015.
- [15] K. He, X. Zhang, S. Ren, J. Sun "Deep Residual Learning for Image Recognition." In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Xplore, 2015, pp. 770-778.
- [16] R. Girshick, J. Donahue, T. Darrell, J. Malik "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, Tech report (v5)." In: The IEEE Conference on Computer Vision and Pattern Recognition, IEEE Xplore, 2014, pp. 580-587.
- [17] R. Girshick "Fast R-CNN." In: The IEEE International Conference on Computer Vision (ICCV), IEEE Xplore, 2015, pp. 1440-1448.

- [18] S. Ren, K. He, R. Girshick, J. Sun "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6, 2016, pp. 1137 – 1149.
- [19] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C-Y. Fu, A. Berg "SSD: Single Shot MultiBox Detector." In: European Conference on Computer Vision (ECCV), Springer, 2016, pp. 21-37.
- [20] J. Redmon, S. Divvala, R. Girshick, A. Farhadi "You Only Look Once: Unified, Real-Time Object Detection." In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Xplore, 2016, pp. 779-788.
- [21] S. Yuheng, Y. Hao "Image Segmentation Techniques Overview." In: 2017 Asia Modelling Symposium (AMS), Malaysia, 2017, pp. 103-107.
- [22] V. Badrinarayanan, A. Kendall, R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 12, 2017, pp. 2481-2495.
- [23] P. Jaccard "The Distribution Of The Flora In The Alpine Zone." In: New Phytologist, Vol. 11, No. 2, 1912, pp. 37-50.
- [24] X. Wang, T. Kong, C. Shen, Y. Jiang, L. Li "SOLO: Segmenting Objects by Locations." In: European Conference on Computer Vision - ECCV, 2020, pp. 649-665.
- [25] I. Piper, G. Citerio, I. Chambers, C. Contant, P. Enblad, H. Fiddes, T. Howells, K. Kiening, P. Nilsson, Y. H. Yau "The BrainIT group: concept and core dataset definition." In: Acta Neurochirurgica, Vol. 145, 2003, pp. 615–629.
- [26] M. Marinho, D. Arruda, F. Wanderley, A. Lins "A Systematic Approach of Dataset Definition for a Supervised Machine Learning Using NFR Framework." In: 2018 11th International Conference on the Quality of Information and Communications Technology (QUATIC), 2018, pp. 110-118.
- [27] M. Defferrard, K. Benzi, P. Vandergheynst, X. Bresson "FMA: A Dataset For Music Analysis." In: 18th International Society for Music Information Retrieval Conference, Suzhou, China, 2017.
- [28] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, L. Zitnick "Microsoft COCO: Common Objects in Context." In: European Conference on Computer Vision (ECCV), 2014, pp. 740-755.
- [29] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, L. Fei-Fei "ImageNet: A large-scale hierarchical image database." In: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255.
- [30] G. A. Miller "WordNet: A Lexical Database for English." In: Communications of the ACM, Vol. 38, No. 11, 1995, pp. 39-41.
- [31] C. Fellbaum "WordNet: An Electronic Lexical Database." Cambridge, MA, MIT Press, 1998.
- [32] M. Everingham, L. v. Gul, C. Williams, J. Winn, A. Zisserman "The Pascal Visual Object Classes (VOC) Challenge." In: International Journal of Computer Vision, Springer, Vol. 88, No. 2, 2010, pp. 303-338.
- [33] A. Krizhevsky "CIFAR-10 Dataset." University of Toronto, [online], [cit. 1.3.2021], Dostupné na internete: <https://www.cs.toronto.edu/~kriz/cifar.html>
- [34] J. Snoek, H. Larochelle, R. P. Adams "Practical Bayesian Optimization Of Machine Learning Algorithms." In: Advances in Neural Information Processing Systems 25 (NIPS), 2012.
- [35] S. Shahinfar, P. Meek, G. Falzon "How many images do I need? Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring." In: Ecological Informatics, Vol. 57, 2020.
- [36] M. Štancel, D. Bural "Algoritmus pre automatické generovanie obrazových dátových sád." 2020.
- [37] P. J. Reddie "Darknet." [online], [cit. 4.3.2021], Dostupné na internete: <https://pjreddie.com/darknet/>
- [38] OpenCV "Real-time Optimized Computer Vision Library." [online], [cit. 4.3.2021], Dostupné na internete: <https://opencv.org/>
- [39] GitHub "AlexeyAB, Darknet on Windows." [online], [cit. 4.3.2021], Dostupné na internete: <https://github.com/AlexeyAB/darknet>
- [40] Microsoft COCO "COCO: Common Objects in Context." [online], [cit. 4.3.2021], Dostupné na internete: <https://cocodataset.org/#download>
- [41] J. Garcia, A. Barbedo "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification." In: Computers and Electronics in Agriculture. Elsevier, Vol. 153, 2018, pp. 46-53.

Lung Ultrasound Analysis with Deep Learning

Miroslav Jascur

Department of Cybernetics and Artificial Intelligence
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 04001 Košice, Slovakia
Email: miroslav.jascur@tuke.sk

Marek Bundzel

Department of Cybernetics and Artificial Intelligence
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 04001 Košice, Slovakia
Email: marek.bundzel@tuke.sk

Abstract—Particular post thoracic surgery complications are standardly monitored using methods that employ ionizing radiation. Following a clinical trial of novel lung ultrasound (LUS) examination procedure that can replace X-raying, a need to automatize the diagnosis procedure has arisen. Deep learning was used as a powerful tool for LUS analysis. We present a novel deep learning method - Motion Mode (M-Mode) classification - to detect lung sliding motion absence in LUS. M-Mode classification leverages semantic segmentation to select a 2D slice across the temporal dimension of the video recording. This 2D slice is an input to a convolutional neural network (CNN). The output of the CNN indicates the presence or absence of lung sliding in the given time slot. We aggregate the partial predictions over the entire video recording to determine if the subject has developed post-surgery complications. For comparison, we implemented two other deep learning methods: Lung sliding 2D classification and Lung sliding 3D classification. M-Mode classification achieves an accuracy of 88.18% and aggregated accuracy of 95.83%. Lung sliding 3D classification outperformed M-Mode classification with an accuracy of 95.19%; the aggregated accuracy was again 95.83%. We concluded that all implemented methods are applicable to detect the lung sliding absence in LUS video recordings.

Index Terms—Deep learning, convolutional neural network, lung ultrasound, lung sliding, pneumothorax.

Abstract—Komplikácie po operáciách hrudníka sa štandardne monitorujú metódami využívajúcimi ionizujúce žiarenie. Po vykonaní klinickej skúšky pri novom postupe vyšetrenia pľúc ultrazvukom, ktorý môže nahradiť röntgenové žiarenie, vznikla potreba automatizovať diagnostický postup. Hlboké učenie sa použilo ako výkonný nástroj na analýzu ultrazvuku pľúc. Predstavujeme novú metódu hlbokého učenia - klasifikáciu M-Mode snímok - na detekciu absencie posuvného pohybu pľúc v ultrazvukovom zobrazení. Klasifikácia M-Mode využíva sémantickú segmentáciu na výber 2D rezu cez časový rozmer záznamu videa. Tento 2D rez je vstupom do konvolučnej neurónovej siete (KNS). Výstup KNS indikuje prítomnosť alebo neprítomnosť kĺzania pľúc v danom časovom úseku. Čiastkové predpovede agregujeme v rámci celého videozáznamu, aby sme určili, či sa u subjektu vyvinuli pooperačné komplikácie. Na porovnanie sme implementovali ďalšie dve metódy hlbokého učenia: klasifikáciu 2D ultrazvuku a 3D klasifikáciu videa. Klasifikácia M-Mode dosahuje presnosť 88,18% a agregovanú presnosť 95,83%. Klasifikácia 3D kĺzania pľúc prekonala klasifikáciu M-Mode s presnosťou 95,19%; agregovaná presnosť bola opäť 95,83%. Dospeli sme k záveru, že všetky implementované metódy sú použiteľné na detekciu absencie kĺzania pľúc v ultrazvukových záznamoch.

Index Terms—Hlboké učenie, konvolučná neurónová sieť, ul-

trazvuk pľúc, kĺzanie pľúc, pneumotorax.

The standard imaging method of monitoring post-surgery complications such as pneumothorax and pleural effusion after thoracic surgery is a chest X-ray. Previous clinical trials performed by the co-authors of this paper have shown that lung ultrasound (LUS) can be successfully used instead of X-ray [1]. The advantages of LUS include avoiding unnecessary radiation exposure, reduced need for in-hospital transfers of patients with inserted chest drains (a flexible plastic tube that is inserted through the chest wall and into the pleural cavity) for X-ray examination, and lowering the economic costs. Using LUS in this context is innovative and addresses the drawbacks of X-raying. On the other hand, interpretation of the LUS imagery in this novel application requires re-training and educating human experts. Objections to the subjectivity of LUS examination with possible forensic impact could be raised. We aim to objectivize the LUS interpretation utilizing artificial intelligence, specifically by deep learning (DL).

Feature extraction from ultrasound (US) heavily relies on the radiologist's expertise, and this may be difficult to share, especially if a novel diagnostic protocol is being implemented, such as in our case. Automated computer vision methods and machine learning algorithms that can distill information from US images or videos have been introduced over the past decades. DL methods can transform raw image data to a high level of abstraction and automate feature extraction from the US. They can be applied to image analysis and aid in decision support systems. To our best knowledge, DL has not been applied to analyze LUS in the context of the detection of post thoracic surgery complications.

LUS is a powerful medical imaging method for the examination of the lung and pleura. It aids in the diagnostics and management of pneumothorax, pleural effusion, acute dyspnoea, pulmonary edema, pulmonary embolism, pneumonia, interstitial processes, and for monitoring of patients supported by mechanical ventilation [2]. The Bedside Lung Ultrasound in Emergency (BLUE) protocol is a standard scanning protocol for immediate acute respiratory failure diagnosis. BLUE protocol produces correct diagnoses in 90% of the cases [3]. While examining the patient, the physician infers nodes in the diagnostic decision tree, and the first node/decision that the physician makes is a qualitative estimate of lung sliding

presence. Lung sliding, described in [4], is the apparent back and forth movement of the visceral and parietal pleura as the patient breathes.

We present a novel method for the detection of lung sliding presence - Motion Mode (M-Mode) classification, Fig. 1. We compare the M-Mode classifier with two other approaches: 2D convolutional neural network (CNN) and 3D CNN. All classifiers form a part of the same architecture, Fig. 1. First, the classifier takes a subset of the frames from the source video. Second, one of the classifiers makes a binary prediction of whether lung sliding is present or absent. Third, this process is repeated for each subset of frames available. Finally, the partial predictions are algorithmically aggregated, and the decision if lung sliding is present or absent is made. Our results show that all of these models are suitable for the detection of lung sliding motion in LUS. We have also produced a LUS dataset containing recordings of patients after thoracic surgery. We have labeled the recordings for binary classification of lung sliding presence and tissue detection. We present here a segmentation mask production method based on selecting bounding boxes and simple image processing. We have experimentally analyzed each classifier's performance, the optimal number of input frames, and the performance of the proposed architecture.

I. STATE OF THE ART IN ALGORITHMIC US ANALYSIS

A. Ultrasound Analysis with Deep Learning

US is one of the most commonly used medical imaging techniques in clinical practice. Over the decades, it has been demonstrated that it has several significant advantages over other medical imaging modalities such as X-ray, magnetic resonance imaging, and computed tomography. US does not employ ionizing radiation, it is portable, accessible, and cost-effective [5]. Furthermore, it remains one of the best modalities for studying in-vivo motion. On the other hand, the quality of US images is compromised by low resolution, image artifacts, and noise. The quality also highly depends on the operator's expertise and skills and the capabilities of the device used [5].

During the last decade, DL-based methods have been primarily applied in 4 different computer vision tasks on US imagery: classification, detection, segmentation, and registration [6]. From a clinical point of view, DL was extracting information for biometric measurements, computer-aided diagnosis, image-guided interventions, or therapy. These methods were employed to analyze different anatomical structures, including lung, liver, breast, prostate, fetus, spine, heart, kidney, and muscle [6].

In US, DL classification methods are used to determine the presence of anatomical or pathological features [6]. Many machine learning approaches work with pre-extracted, hand-crafted features on which classifiers such as decision trees, Naive Bayes classifiers, or K-nearest neighbors are applied. The hand-crafted features are often not invariant to occlusion, illumination, morphological variation, and rotation. Convolutional Neural Networks (CNNs) are presently among the most advanced and most widely used artificial intelligence

methods. CNNs almost completely perform the feature extraction process and, by adding fully connected layers, end-to-end training for classification is enabled. The problems addressed include classification of tumors and lesions in brain and liver [7], thyroid nodules [8] and examination of fetal biometry [9]. All these applications use a certain type of DL architecture to extract image features. While the body of research [10], [11] employs transfer learning from natural images to compensate for the limited availability of labeled US datasets. Some approaches directly address the lack of data samples by using specialized architectures, for example, deep polynomial networks [12], or by data augmentation [13]. CNNs have been applied for US video classification [11], [14], [15]. Certain applications, like in our case, lung sliding detection, require an analysis of movement patterns. If 3D CNNs are to be used for this purpose, a large 3D labeled dataset and significant computational capabilities are required.

Semantic image segmentation is the task of assigning a class label to every pixel in an image. To perform quantitative analysis of pathological and non-pathological tissues, we often need to locate them in a US image. Fully convolutional neural networks such as U-Nets [16] are state-of-the-art in biomedical image processing. For the segmentation of rigid structures, fine-tuned CNNs paired with an effective post-processing method are currently the best-performing approaches [17], [18]. For the segmentation of non-rigid anatomical structures, prior information is often introduced to the process. This information can be provided through active contours [19], deformable templates [20] or image registration.

B. Lung Ultrasound Analysis and Deep Learning

In intensive care, two protocols are based on LUS: the BLUE protocol for diagnostics of acute respiratory failure and the Fluid Administration Limited by Lung Sonography (FALLS) protocol for the management of acute circulatory failure [21]. LUS aids in the diagnosis of lung parenchymal and pleural diseases in bedside procedures. It enables diaphragmatic and lung recruitment assessment. Further, it can be applied to the early detection and management of respiratory complications under mechanical ventilation or after thoracic surgery.

Each of these applications requires expert knowledge of LUS patterns. The protocols mentioned above use patterns like lung sliding, and artifacts, specifically A- and B-lines [22]. A-lines appear as horizontal hyperechoic lines beneath and parallel to the pleura. On the other hand, B-lines appear as vertical comet-tail artifacts arising from the pleural line. B-lines correlate with different pathologies. Therefore, an important body of research is dedicated to the detection of B-lines. Van Sloun et al. [23], [24] present a method for B-line detection based on weakly-supervised CNN and B-line localization through class activation mapping.

Roy et al. [25] describe a DL architecture for automatic analysis of LUS. Specifically, the system contains three modules: frame scoring with Spatial Transformer Network and CNN, video score aggregation, and frame-based semantic

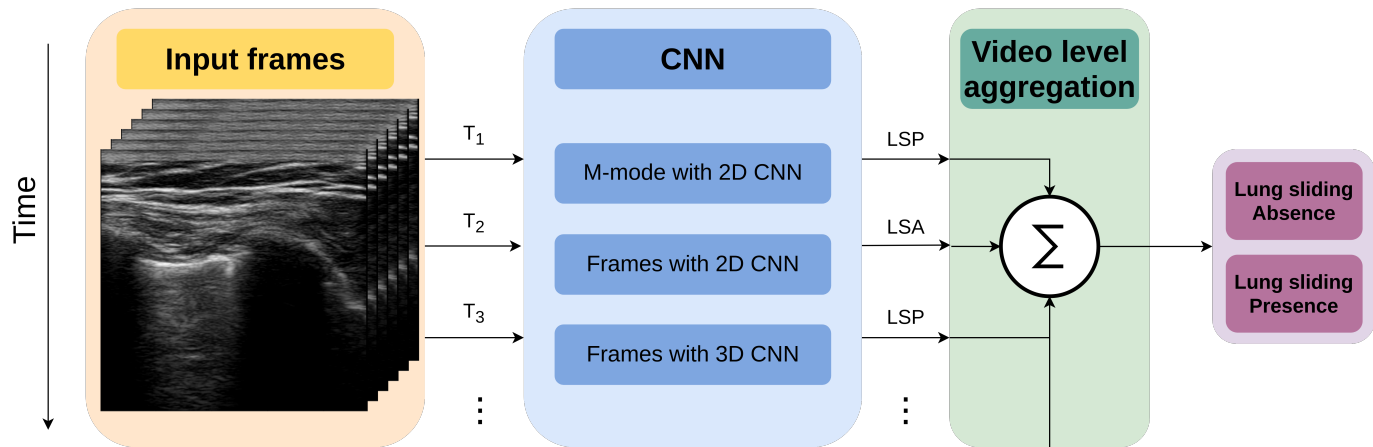


Fig. 1. Lung sliding detection architecture. Sequences of frames are inputted into CNN. One of the CNNs makes a partial prediction. Partial predictions are aggregated for the whole video, and the final prediction, whether lung sliding is present or absent, is made. ($T_1 \dots T_n$ - frame time intervals, LSA - lung sliding absence, LSP - lung sliding presence).

segmentation with a fully convolutional network. With this architecture, the authors achieved accurate prediction and localization of COVID-19 infection biomarkers. Similarly, Born et al. [26] present a CNN that classifies a LUS image into three classes - COVID-19, bacterial pneumonia, and healthy lung images. The CNN used achieved 89% accuracy per frame and 92% accuracy per video file and had a sensitivity of 0.96 and a specificity of 0.79 in COVID-19 detection. [27] proposes a frame-based classification with CNN that correctly distinguishes COVID-19 LUS videos from healthy and bacterial pneumonia data with a sensitivity of 0.90 ± 0.08 and a specificity of 0.96 ± 0.04 .

Correa et al. [28] used hand-crafted LUS features to detect pneumonia. From the rectangular section intersecting pleura, their algorithm extracts four hand-crafted features. Feature is average brightness distribution of 25, 50, 75 and 100 of the section. The authors trained a fully connected neural network for classification, that was able to correctly identify pneumonia infiltrates with 90.9% sensitivity and 100% specificity.

II. DATASET DESCRIPTION

In collaboration with physicians from the Department of Thoracic Surgery at the Jessenius Faculty of Medicine in Martin, we prepared a LUS video dataset. The dataset contains recordings of patients after thoracic surgery. Pneumothorax and pleural effusion are common postoperative complications in this context. Our dataset contains examinations for pneumothorax; during this diagnostic approach, the patient is laid on his/her back. The physician scans the anterior and lateral sides of the thorax using a linear probe. The probe is placed between two ribs longitudinally to the body. Both lungs are examined. The lung not affected by the surgical procedure serves for comparison to the physician, who is looking for three signs: lung sliding, B-lines, and lung point.

The videos were recorded in B-Mode (Brightness mode displays the acoustic impedance of a two-dimensional cross-

section of tissue.). The video files were divided by physicians into two classes:

- 1) **Lung Sliding Presence** - this class represents healthy lung sliding movement, 28 recordings.
- 2) **Lung Sliding Absence** - this class represents lack of lung sliding movement, 20 recordings.

The total number of the available examinations is 545, taken from 297 subjects, out of which physicians categorized the above 48 video files at this time. All the examinations were made with Sonoscape S2 Portable Ultrasound Machine with the linear probe Sonoscape L741 working at a 5.0 to 10.0 MHz frequency. The recordings are captured at resolutions of either 640x480 or 854x480 pixels. We have cropped the frames to a 480x480 pixels square image containing the middle section of the frames by cutting off black margins. Together, we have generated 17,338 frames. Out of 17 338 frames, 11,469 frames have lung sliding present, and 5869 frames have lung sliding absent. These we used in DL training, validation, and testing.

A. Segmentation Labels

A highly trained expert must label LUS video files, but the procedure is not time consuming. On the other hand, labeling a 2D dataset for basic semantic segmentation in LUS is more time expensive but requires only low expertise. We have educated a group of 80 volunteers, who have drawn rectangular bounding boxes around the objects of interest using labelling tool [29]. Labeling was the first step in the production of semantic segmentation masks.

We randomly selected 6400 frames from 182 videos, and each volunteer labeled 80 frames. The objects of interest were: *pleura*, *lung*, *ribs* and *tissue*. Class *pleura* represents both the parietal and visceral pleura. In the LUS image, it is a shimmering white line connecting the lung and the chest wall. Class *lung* is the area under the pleura between the anechoic shadows of the ribs. Class *ribs* represents the area of anechoic shadow. This artifact appears because the sound waves are

well reflected by the hard material of the ribs. Finally, the class *tissue* represents the chest wall comprised of the skin, fat, and muscle tissue.

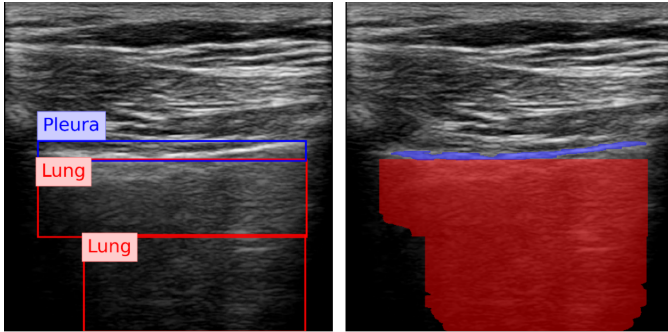


Fig. 2. Creation of semantic masks for pleura and lung from labeled data. On the left: bounding boxes labeled by volunteers. On the right: semantic masks generated from the bounding boxes.

We produced binary semantic segmentation masks for the objects of interest by applying several computer vision algorithms to the areas delimited by the bounding boxes. We assumed that the objects of interest differ from other objects in the bounding box by brightness. To create the segmentation mask for the *lung* class, we applied intensity thresholding in the bounding box with a threshold value determined by Otsu's method. The initial result was refined by morphological dilation, and the holes in the mask were filled. Finally, the mask was returned to the coordinate system of the image. We repeated this process for each bounding box in an image. In the end, we kept only the largest contiguous region because some of the images contained multiple separate regions of the lung. We generated *pleura* masks using the same process.

III. METHODS

We compare three methods for the classification of lung sliding motion:

- 1) **M-Mode classification** - classifies lung sliding motion with 2D CNN on a 2D data slice selected across the temporal resolution of the video.
- 2) **Lung sliding 2D classification** - classifies lung sliding patterns from either a single frame or a series of frames with 2D CNN.
- 3) **Lung sliding 3D classification** - classifies lung sliding patterns from series of frames with 3D CNN.

A. *M-mode classification*

M-mode classification (see Fig. 3) is a modification of the traditional approach of pneumothorax detection. Physicians use this view of the data to confirm their qualitative estimate. It automates two tasks performed by a physician - slice selection and lung sliding presence estimation. First, our architecture takes the first frame from the series, and the U-Net predicts the segmentation mask for the selected tissue of interest (see Section III-A1 - the pleura or the lung). Second, it uses the x-coordinate of the centroid of the predicted tissue mask as the x-position of a one-pixel wide column selected in the consecutive

frames to form a 2D temporal slice. This slice represents the CNN's input, and it is categorized into two categories - lung sliding present or lung sliding absent.

1) *Lung and pleura segmentation*: For the segmentation of the lung and pleura we used fully convolutional architecture - U-Net [16], Fig. 3. Our model contains three encoder and three decoder blocks. Each block uses a 2D convolutional layer, 2D batch normalization, and a rectified linear unit layer in this order, twice, followed by either 2D max-pooling in the case of the encoder block or 2D transposed convolution in the case of the decoder block. The first two encoder blocks connect to a deeper layer and skip-connect to the corresponding decoder layer with the same resolution. The third and deepest encoder block connects directly to the deepest decoder block. Decoder blocks take the skip connected output from the encoder block and the output from the decoder's deeper layer and produce the input to the higher resolution decoder block. Finally, the prediction is made using a softmax layer. Loss is calculated as weighted pixelwise cross-entropy. The weight is determined as a class distribution from a random subsample of the dataset. In the pleura segmentation, decreasing the weight of pleura segments by 80% improved the segmentation results.

The input image resolution is 480×480 pixels. We used four types of data augmentation: vertical and horizontal flipping and random rotation, translation, and scaling. We trained the neural networks for 30 epochs with early stopping that occurred when after ten epochs, the validation loss did not drop below the lowest previously recorded value. We randomly divided the dataset into three parts: 50% for training and 25% each, for validation and testing.

To improve the resulting segmentation, we apply post-processing both for the pleura and for the lung:

- The connected regions in the prediction mask are detected.
- The largest connected region is kept, the remaining are discarded.
- The holes in the largest connected region are filled.

2) *M-mode classification*: M-mode image is a 2D slice across the temporal dimension in the ultrasound video, Fig. 3. For the classification of this type of image, we constructed a custom 2D CNN. Input dimensions are $1 \times H \times N$, where H is the image height, and N is the number of frames. The model contains three convolutional blocks. Each of the 2D convolutional layers is followed by a Rectified Linear Unit (ReLU) and a 2D max pooling. The features from the last convolutional layer are flattened by global average pooling [30]. The kernel size of average pooling is dynamically adjusted to produce 64 neurons. The height and width of the average pooling kernel are the same as the output height and width of the last convolutional layer. Average pooling features are fully connected to the network output. Finally, the binary prediction is made using a softmax layer. The loss is calculated as a weighted cross-entropy, where the weights represent the class distribution.

All the presented models were trained for a maximum of 100 epochs with early stopping when the validation loss did

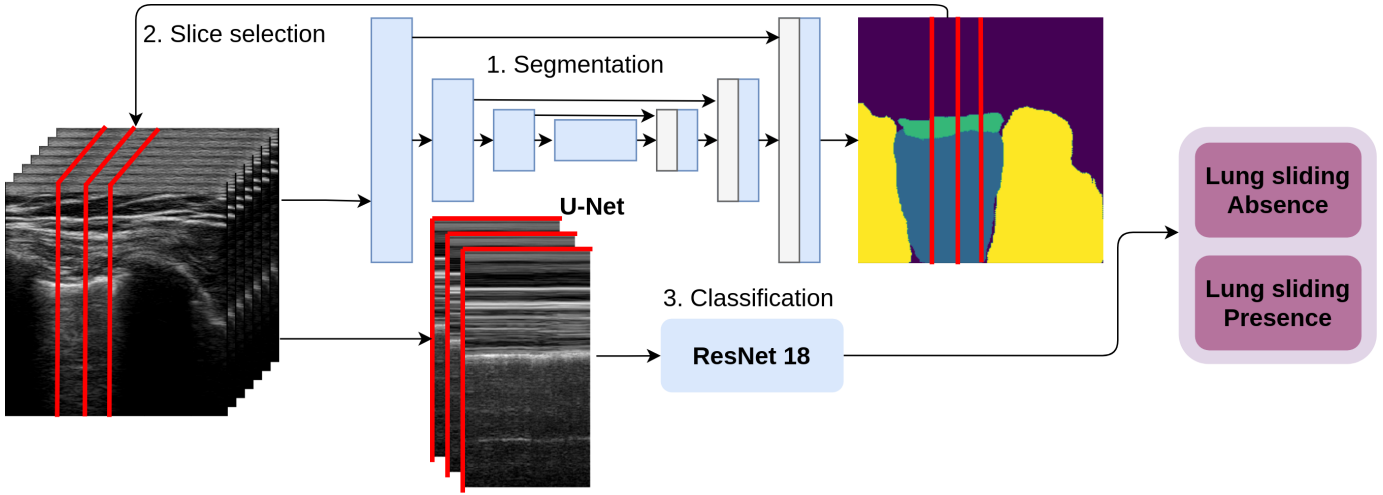


Fig. 3. The architecture of the automated M-Mode classification. (1. Segmentation of the tissue from a single LUS frame. 2. Selection of a temporal slice based on the tissue centroid found after semantic segmentation. 3. Classification of the slice with a custom 2D CNN.)

not improve in 10 consecutive epochs. We used the Adam optimization algorithm with the learning rate set to 0.001. The dataset was split into three equally sized subsets for training, validation, and testing. For data augmentation, we used random erasing [31] and random affine transformations (rotation, translation, and scaling). The probability of the transforms was equally set to 0.5. Rotation ranged between $[-10, 10]$ degrees, horizontal translation between $[-0.1, 0.1]$ of the image height, vertical translation between $[-0.1, 0.1]$ of the image width and size scales ranged from 0.9 to 1.1 of the original size. Random erasing removed rectangular areas (pixels values set to 0) between $[0.05, 0.33]$ of the image area.

B. 2D and 3D Lung Sliding Classification

We used a custom 2D CNN described in Sec. III-A2 for lung sliding 2D classification. The input size was $N \times H \times W$, where N is the number of frames, H is the height of the video, and W is the width of the video. The images were resized to 128×128 pixels to match the resolution of 3D CNNs. The training strategy was the same as by M-mode classification.

We used a 3D version of SqueezeNet [32], [33] for lung sliding 3D classification. 3D-SqueezeNet begins with a convolution layer, followed by max-pooling, 8 Fire modules (Fire module is a building block for convolutional neural networks, notably used as part of SqueezeNet), and ending with a final convolutional layer. The input had the same shape as for 2D image classification, and images were resized to 128×128 pixels to save the computational resources. Feature maps were downsampled by max-pooling after every other Fire block. The 3D version of Fire block accepts an input with the size of $F \times N \times H \times W$, where F is the number of feature maps, N is the number of frames, H is the height of the video, and W is the width of the video. First, the Fire block applies $1 \times 1 \times 1$ convolution with batch normalization (BN) and ReLUs. Second, on one half of the feature maps, it applies $1 \times 1 \times 1$ convolution with BN, and on the other half, it applies

$3 \times 3 \times 3$ convolution with BN. Third, it concatenates both halves of the feature maps. Finally, it sums the result with the block’s input signal and activates the signal with ReLUs. 3D SqueezeNet was trained for 30 epochs, using stochastic gradient descent starting with learning rate 0.01. After ten epochs, the training rate was decreased to 0.001.

For both architectures, augmentation was a combination of random vertical flipping, random erasing [31] and random rotation, translation, and scaling, see Fig. 4. The probability of all these transforms was set to 0.5. The random rotation ranges between $[-30, 30]$ degrees; horizontal translation ranges between $[-0.1, 0.1]$ of the image height and vertical translation ranges between $[-0.1, 0.1]$ of image width and image size scaling ranges from 0.75 to 1.25 of the original size. Random erasing parameters are set to the same values as described in Sec. III-A2.

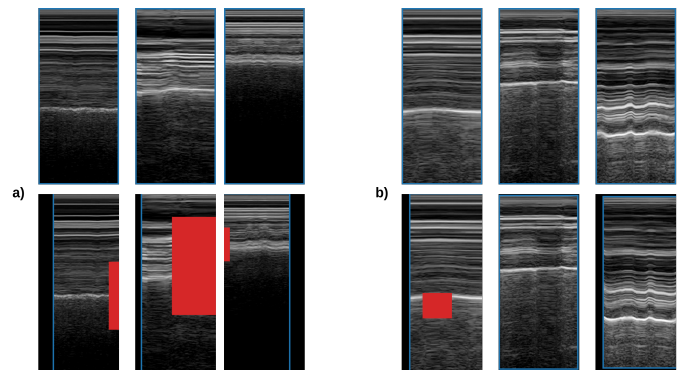


Fig. 4. Dataset augmentation for M-Mode classification, Lung sliding 2D and 3D classification. Green outlines the original image. Red represents randomly erased areas.

IV. EVALUATION AND AGGREGATION

We use accuracy, precision, recall, and F1-score to evaluate both semantic segmentation and classification CNNs. We

use the terms sensitivity and specificity instead of precision and recall for the classification tasks. Further, we calculate balanced accuracy for the classification. We add Intersection over Union (IoU) metric to evaluate semantic segmentation results and use the term Dice coefficient instead of F1-score.

In binary segmentation, the tissue (lung/pleura) is our target class. Therefore, accuracy represents the ratio of the correctly classified pixels to all pixels in the mask. Precision measures the proportion of the correctly classified tissue pixels to all pixels classified as tissue. Recall measures the proportion of correctly classified background pixels to all pixels classified as background. The dice coefficient is a harmonic mean of precision and recall. IoU measures the overlap between the correctly identified tissue pixels and the area of the union of all pixels identified as tissue pixels and misclassified background pixels.

In classification, lung sliding absence is our target class. Accuracy represents the ratio of correctly classified samples with lung sliding presence and with lung sliding absence to all samples. Balanced accuracy is the average accuracy per class to counteract the class imbalance. Sensitivity measures the proportion of correctly classified samples with lung sliding absence to all samples classified with lung sliding absence. On the other hand, specificity measures the proportion of correctly classified samples with lung sliding presence to all samples classified with lung sliding presence. F1-score is a harmonic mean of sensitivity and specificity.

In addition to these commonly used measures, it is desirable to evaluate aggregated performance over video files, since each video file is a recording taken with a single subject. From the medical point of view it is not only important to categorize samples but also the subjects (healthy, not healthy) and to evaluate the ability of the classifiers to categorize the subjects. Let \hat{y}_t be the binary prediction of the classifier for a sequence of frames (a sample) starting at time t . We aggregate this prediction by averaging:

$$y_{agg}(\hat{y}_t) = \frac{1}{N_{samples}} \sum_{t=0}^{N_{samples}-1} \hat{y}_t \quad (1)$$

where $N_{samples}$ is the number of samples from the specific video file. Let y_t be the ground truth for the binary classification for the sample starting at time t . Then the aggregated video accuracy is:

$$acc_{video}(\hat{y}_t, y_t) = \frac{1}{N_{samples}} \sum_{t=0}^{N_{samples}-1} \hat{y}_t = y_t \quad (2)$$

where $N_{samples}$ is the number of samples from the specific video file. This measure is the ratio of correctly classified samples from the video file to all samples from the given video. To measure the performance of a classifier on a set of testing videos we introduce Thresholdedsubsu Aggregated Accuracy $taa_{threshold}$, that represents the ratio of the number of video files with the aggregated video accuracy over the given threshold to the number of all files. For example, taa_{50} is defined as:

$$taa_{50} = \frac{1}{N_{videos}} \sum_{t=0}^{N_{videos}-1} acc_{video}(\hat{y}_t, y_t) > 0.5 \quad (3)$$

where N_{videos} is the total number of video files. taa_{50} represents the ratio of video files that have in aggregate more than 50% of samples correctly classified. Similarly taa_{75} and taa_{90} are defined. These measures require that more than 75% and 90% of the samples be classified correctly for the video file to be counted in.

V. EXPERIMENTAL RESULTS

This section is structured into two subsections. First, we describe the experiments with M-Mode classification and then report 2D and 3D classification results. In Sec. V-A, we first review the semantic segmentation results, then the M-mode classification results. In Sec. V-A2, we compare the performance of CNNs with various numbers of frames used on the input. Further, each classifier is evaluated with and without video aggregation. In addition, we contrast the performance of the M-mode slice selection methods based on two tissues - pleura and lung tissue. We review the classifiers' performance with various numbers of input frames in 2D and 3D classification of lung sliding presence. We report the performance for each classifier with and without video aggregation.

A. M-Mode Classification

1) *Segmentation results:* Tab. I shows the performance metrics of U-Nets' binary segmentation of the pleura and the lung on the test dataset. We use accuracy, recall, precision, IoU, and Dice coefficient. In segmentation, improving IoU metrics is the criterion.

TABLE I
U-NET PERFORMANCE CHARACTERISTICS FOR THE SEGMENTATION OF PLEURA AND LUNG (POST.- INDICATES IF POST-PROCESSING ACCORDING TO SEC. III-A1 WAS APPLIED. ACC. - ACCURACY, IOU - INTERSECTION OVER UNION, DICE - DICE COEFFICIENT.)

Tissue	Post.	Acc.	Precision	Recall	IoU	Dice
Pleura	No	0.9816	0.7580	0.5985	0.5025	0.6689
Pleura	Yes	0.9863	0.7497	0.7093	0.5735	0.7289
Lung	No	0.9198	0.9390	0.7644	0.7283	0.8428
Lung	Yes	0.9292	0.9369	0.7919	0.7518	0.8584

Without post-processing, the pleura segmentation was 96.81% accurate, but the IoU score was only 50.25%. By applying post-processing described in Section III-A1, the resulting accuracy improved by 0.47% and IoU by 7.1%. The post-processing improvement was at the cost of decreased precision caused by the removal of true positive pixels.

Lung segmentation with U-Net was more successful than pleura segmentation with an IoU higher by 32% without post-processing. The overall accuracy is lower with or without post-processing because the lung samples are pixel-wise more balanced than the pleura samples. With post-processing, every metric improved except for precision that decreased by 0.21%. The best-performing classifier achieved an accuracy of 92.92% and IoU 75.18%.

2) *M-Mode Classification Results:* M-Mode classification results are presented in Tab. II. Tab. II contains from left to right: the number of input frames, name of tissue for slice selection, accuracy, balanced accuracy, F1-score, sensitivity, and specificity. Each row shows the performance metrics for one of eight classifiers. Four of the classifiers use the pleura tissue to select a temporal slice, and the other four use the lung tissue. We segment the tissue using binary U-Nets with the postprocessing from Sec V-A1. The input of the CNNs ranges from 32 frames to 256 frames. The difference between the two CNNs that use the same tissue for slice selection is the number of input frames. Therefore, the two CNNs inputting 32-frames trained for different tissues use the same architecture.

TABLE II
PERFORMANCE OF M-MODE CLASSIFICATION WITH CUSTOM 2D CNN. M-MODE SLICE IS SELECTED FROM EITHER PLEURA OR LUNG TISSUE. (N - NUMBER OF INPUT FRAMES, ACC. - ACCURACY, BAL. ACC. - BALANCED ACCURACY, F1 - F1-SCORE).

N	Tissue	Acc	Bal. Acc.	F1	Sensitivity	Specificity
32	pleura	0.8327	0.8467	0.7812	0.6964	0.8895
64	pleura	0.8818	0.8945	0.8327	0.7544	0.9292
128	pleura	0.8806	0.8838	0.8175	0.7546	0.8919
256	pleura	0.8864	0.8950	0.7890	0.6958	0.9110
32	lung	0.7994	0.7997	0.7242	0.6613	0.8005
64	lung	0.8196	0.8189	0.7423	0.6800	0.8171
128	lung	0.8527	0.8529	0.7777	0.7143	0.8535
256	lung	0.7839	0.7928	0.6394	0.5283	0.8098

The best-performing CNN with lung slice selection takes 128 frames as input and achieves accuracy 85.27%, with a balanced accuracy of 85.29%. This CNN also achieves the highest F1 score - 77.77%, sensitivity - 71.43%, and specificity- 85.35%. However, on average, CNNs that use the lung tissue for slice selection perform worse than those that use the pleura tissue in every metric. The CNNs with pleura slice selection achieve accuracy and balanced accuracy from 88.06 to 89.5% with one exception, that is the CNN with 32 input frames and accuracy of 83.27%. The 64-frame CNN reaches the highest F1-score - 83.27% and the highest specificity 92.92%.

TABLE III
AGGREGATE PERFORMANCE OF M-MODE CLASSIFICATION WITH CUSTOM 2D CNN WITH $taa_{threshold}$. THE SLICE IS POSITIONED BASED ON EITHER PLEURA OR LUNG TISSUE.

N	Tissue	taa_{50}	taa_{75}	taa_{90}
32	pleura	0.9583	0.8333	0.6042
64	pleura	0.9583	0.8542	0.5417
128	pleura	0.8667	0.7333	0.5111
256	pleura	0.7931	0.6552	0.4138
32	lung	0.9375	0.6875	0.1667
64	lung	0.9375	0.6875	0.2500
128	lung	0.9111	0.7556	0.4889
256	lung	0.6207	0.4828	0.1034

In Tab. III we compare the thresholded video accuracy of the 8 classifiers from Tab. II. CNNs with slice selection based on pleura tissue outperforms CNNs with lung tissue slice selection in most metrics with regard to the thresholded accuracy. The highest value of taa_{50} for slice selection based on the

lung is 93.75%, while slice selection with pleura achieves an accuracy of 95.83%. These highest taa_{50} are with 32 and 64-frame inputs in both slice selection methods. Quantifier taa_{75} peaks at 85.42% at 64-frames and taa_{90} peaks at 60.43% at 32 frames with pleura tissue for slice selection. Furthermore, 32-frame CNN with pleura slice selection achieves taa_{75} 2.09% worse than the best 64-frame CNN with pleura slice selection. For the CNNs with lung tissue slice selection the best taa_{75} is 75.56% and taa_{90} is 48.89% at 128 input frames.

B. 2D and 3D Classification Results

The values of the performance metrics of 2D CNN described in Sec. III-B are shown in Tab. IV. Each row represents an experiment with a different number of input frames. In the case of 2D CNN we vary the number of frames from 1 to 32. Accuracy and balanced accuracy increases with an increasing number of input frames and peaks at 85.18 and 85.92% at 8 frames. The 8-frame 2D CNN also achieves the best F1-score - 79.83% and specificity 88.12%, while 4-frame 2D CNN reaches the best sensitivity at 76.06%.

TABLE IV
PERFORMANCE OF LUNG SLIDING 2D CLASSIFICATION WITH CUSTOM 2D CNN (N - NUMBER OF INPUT FRAMES, ACC. - ACCURACY, BAL. ACC. - BALANCED ACCURACY, F1 - F1-SCORE).

N	Acc.	Bal. Acc.	F1	Sensitivity	Specificity
1	0.7050	0.6780	0.5782	0.5642	0.5929
2	0.7482	0.7680	0.6922	0.5935	0.8303
4	0.8447	0.8308	0.7741	0.7606	0.7880
8	0.8518	0.8592	0.7983	0.7297	0.8812
16	0.8387	0.8257	0.7667	0.7487	0.7856
32	0.7992	0.7826	0.7097	0.6885	0.7322

The thresholded aggregated accuracies of the 2D CNNs are displayed in Tab. V. The 2D CNN with 1 input frame reaches taa_{50} of 87.5%. 2D CNNs with 2 input frames and 4 input frames achieve the highest value of taa_{50} at 93.75%. The taa_{75} quantifier peaks at 85.42% with 2-frame 2D CNN and quantifier taa_{90} peaks at 1 and 4 frames at 62.5%. The increasing number of frames for this 2D CNN does not improve any three thresholded aggregated accuracy.

TABLE V
AGGREGATE PERFORMANCE OF LUNG SLIDING 2D CLASSIFICATION WITH CUSTOM 2D CNN WITH $taa_{threshold}$.

N	taa_{50}	taa_{75}	taa_{90}
1	0.8750	0.7708	0.6250
2	0.9375	0.8542	0.5833
4	0.9375	0.7292	0.6250
8	0.7292	0.4375	0.2500
16	0.8936	0.6170	0.4468
32	0.8913	0.5652	0.2609

Performance metrics of 3D SqueezeNet from Sec. III-B are shown in Tab. VI. Each row represents one of the 4 3D SqueezeNets with an increasing number of frames from 4 to 32 frames. 3D SqueezeNet achieves an accuracy from 85.45% to 96.12% and a balanced accuracy from 88.11% to 95.19%. 3D SqueezeNet with a 32-frame input outperforms all other

classifiers in all metrics. It achieves an F1-score, sensitivity and specificity of 93.16%, 88.08% and 98.85%, respectively.

TABLE VI

PERFORMANCE OF LUNG SLIDING 3D CLASSIFICATION WITH 3D SQUEEZE NET (N - NUMBER OF INPUT FRAMES, ACC. - ACCURACY, BAL. ACC. - BALANCED ACCURACY, F1 - F1-SCORE).

N	Acc.	Bal. Acc.	F1	Sensitivity	Specificity
4	0.9170	0.9207	0.8835	0.8397	0.9320
8	0.9331	0.9468	0.9090	0.8409	0.9891
16	0.8906	0.8993	0.8501	0.7859	0.9258
32	0.9519	0.9612	0.9316	0.8808	0.9885

Aggregated thresholded accuracies of 3D SqueezeNet are presented in Tab. VI. Quantifier taa_{50} ranges from 93.75% to 95.83% for all frame counts. The highest value of taa_{50} at 95.83% is obtained with 8 and 32-frame versions of the 3D SqueezeNets. The other two quantifiers - taa_{75} and taa_{90} peak at 91.67% and 77.08% respectively. These peaks are at 32-frames.

TABLE VII

AGGREGATE PERFORMANCE OF LUNG SLIDING 3D CLASSIFICATION WITH 3D SQUEEZE NET WITH $taa_{threshold}$.

N	taa_{50}	taa_{75}	taa_{90}
4	0.9375	0.8542	0.7083
8	0.9583	0.8750	0.7292
16	0.9375	0.7708	0.5625
32	0.9583	0.9167	0.7708

C. Cumulative Results

We present the best-performing classifier from each category in Tab. VIII. We select the classifier based on both, taa_{50} and taa_{75} . Tab. VIII contains a column describing the average number of frames used for video aggregation. For a complete picture, we add an F1-score, which describes the classifiers' performance without video aggregation.

TABLE VIII

PERFORMANCE OF THE BEST OVERALL CLASSIFIERS FROM EACH CATEGORY (F1 - F1-SCORE).

Name	N	$\bar{N}_{samples}$	F1	taa_{50}	taa_{75}
M-Mode 2D CNN	64	73	0.8327	0.9583	0.8542
2D CNN	2	120	0.6922	0.9375	0.8542
3D SqueezeNet	32	97	0.9316	0.9583	0.9167

Overall, the best-performing model is 3D SqueezeNet with 32 frames on the input, followed by M-Mode 2D CNN with 64 frames on the input and 2D CNN with 2 frames on the input. The largest gap between classifiers comes from F1-score. 2D CNN achieves an F1-score of 69.22% compared to 83.27% for M-Mode 2D CNN and 93.16% for 3D SqueezeNet. While this is a significant difference, it does not translate to differences in taa_{50} . On average, 2D CNN makes the aggregate prediction from 23 more samples than 3D SqueezeNet and from 47 more samples than M-Mode CNN. M-Mode 2D CNN and 2D CNN have the same taa_{75} score at 85.45%. 3D SqueezeNet has the highest taa_{75} value at 91.67%.

VI. DISCUSSION

Our architecture selects a temporal slice for classification based on the tissue segments from either pleura or lung. The U-Net for lung segmentation outperforms the U-Net for pleura segmentation in almost every measure, see Tab. I. These performance differences result from different tissue to background ratios and the possible mislabeling during the process of semi-automated production of semantic segmentation masks. The lung tissue is constrained by the pleura and the anechoic shadows of the ribs, but the contours of the pleura tissue are more fuzzy. However, CNNs that use the lung tissue for slice selection perform on average worse than CNNs that use the pleura tissue, see Tab. II. Therefore, the segmented pleura tissue is more suitable for slice selection.

Another interesting finding is that increasing the number of the input frames above a specific number usually decreases taa in M-Mode classification, see Tab. III. Our aggregated accuracy metrics are influenced by two factors - the classifier's accuracy and the number of samples for aggregation. Due to the nature of our taa metrics, misclassification of a sample in a video with fewer samples is more costly than such misclassification in a video with more samples. The number of samples is tied to the number of frames that the classifier takes as input. By increasing the number of votes, we increase the probability of correct classification. Therefore, the classifier with a smaller number of frames at the input will be favored.

Lung sliding is a dynamic phenomenon. However, our 2D CNN with 1 frame input without any specific fine-tuning was able to achieve a balanced accuracy of 67.8% and a taa_{50} of 87.5%. Therefore, spatial patterns such as motion blurring of the pleura, A-line, and B-line artifacts contain sufficient information for lung sliding classification. Further, 2D CNNs with more than 1 frame as an input achieve a balanced accuracy of up to 85.92%.

The dataset used in this study consists of LUS samples taken with a single device from patients after thoracic surgery in a single institution. While we do not think that this invalidates the presented research results, the dataset pertains to this bias. It is reasonable to assume that the classifiers would perform slightly differently if the samples would be taken with US devices of varying imaging quality.

The 3D SqueezeNet with 32-frame input achieves the best performance in both aggregation and classification metrics, see Tab. VI and Tab. VII. Further, it outperforms both M-Mode and image classification with 2D CNNs, see Tab. VIII. The performance gain in the thresholded aggregated accuracy is not as significant as the performance gain in classification metrics. We experienced convergence stability issues with this network.

VII. CONCLUSIONS AND FUTURE RESEARCH

CNNs are suitable for lung sliding detection from LUS videos. We presented three architectures for the detection of lung sliding - M-Mode, Lung sliding 2D, and Lung sliding 3D classification. One of the architecture modules is a CNN that accepts as the input either a sequence of video frames or a 2D slice from this sequence and classifies lung motion into two

classes - lung sliding presence and lung sliding absence. M-Mode, 2D image, and 3D CNNs achieve a balanced accuracy of 89.45%, 76.8%, and 96.12%, respectively. Finally, we aggregated the predictions for each video file and achieved a thresholded aggregated accuracy taa_{50} of 95.83%, 93.75%, and 95.83%, respectively. All three aforementioned methods achieved a performance applicable in clinical practice. The differences between the methods are more on the side of possible explainability of the predictions. The three methods utilize black-box classifiers - CNNs - that implicitly do not provide any explanation. As a part of future research, we plan to visualize the input data with highlighted areas that contributed to the CNN's decision the most and create a more elaborate decision support system. Furthermore, explainability methods for CNNs such as [34],[35],[36] could provide valuable insight to reinforce or contradict the decision of the physician.

Another key difference between the methods is in the data and the type of convolution that each method applies. While lung sliding classification with 2D CNN and 3D CNN uses the same type of input data - a sequence of frames - with 3D convolution, we do not work with spatial patterns but rather with spatio-temporal patterns. While, on average, 3D CNN performs better than 2D CNN, the complexity of visualizing and explaining a 3D CNN prediction outweighs its relatively small performance gain. On the other hand, the automated M-Mode classification uses the temporal patterns directly. In clinical practice, physicians usually confirm their assumption on pneumothorax with a similar view on the data. This 2D temporal slice forms a so-called seashore or barcode sign. In our opinion, physicians' familiarity with these patterns and their simplicity marks the way for future application in clinical practice. Further, automatic selection of the M-Mode slice simplifies pneumothorax detection and relieves a physician from manual slice selection.

Our dataset contains only two classes. Lung sliding is either present or absent. Detection and localization of A-line and B-line artifacts may provide more insight into the correlation between the lung sliding motion and these artifacts.

REFERENCES

- [1] M. Malík, A. Dzian, M. Skaličanová, L. Hamada, K. Zelenák, and M. Grendár, "Chest ultrasound can reduce the use of x-ray in post-operative care after thoracic surgery," *Ann. Thorac. Surg.*, 2020.
- [2] P. Mayo, R. Copetti, D. Feller-Kopman, G. Mathis, E. Maury, S. Mongodi, F. Mojoli, G. Volpicelli, and M. Zanobetti, "Thoracic ultrasonography: a narrative review," *Intensive Care Med.*, pp. 1–12, 2019.
- [3] D. A. Lichtenstein and G. A. Meziere, "Relevance of lung ultrasound in the diagnosis of acute respiratory failure*: the blue protocol," *Chest*, vol. 134, no. 1, pp. 117–125, 2008.
- [4] D. A. Lichtenstein and Y. Menu, "A bedside ultrasound sign ruling out pneumothorax in the critically ill: lung sliding," *Chest*, vol. 108, no. 5, pp. 1345–1348, 1995.
- [5] J. Virmani, R. Agarwal *et al.*, "Deep feature extraction and classification of breast ultrasound images," *Multimedia Tools and Applications*, vol. 79, no. 37, pp. 27 257–27 292, 2020.
- [6] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. X. Li, D. Ni, and T. Wang, "Deep learning in medical ultrasound analysis: a review," *Engineering*, vol. 5, no. 2, pp. 261–275, 2019.
- [7] H. Kutlu and E. Avci, "A novel method for classifying liver and brain tumors using convolutional neural networks, discrete wavelet transform and long short-term memory networks," *Sensors*, vol. 19, no. 9, p. 1992, 2019.
- [8] Q. Guan, Y. Wang, J. Du, Y. Qin, H. Lu, J. Xiang, and F. Wang, "Deep learning based classification of ultrasound images for thyroid nodules: a large scale of pilot study," *Annals of translational medicine*, vol. 7, no. 7, 2019.
- [9] X. P. Burgos-Artizzu, D. Coronado-Gutiérrez, B. Valenzuela-Alcaraz, E. Bonet-Carne, E. Eixarch, F. Crispi, and E. Gratacós, "Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.
- [10] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, and B. Hu, "Liver fibrosis classification based on transfer learning and fcnet for ultrasound images," *Ieee Access*, vol. 5, pp. 5804–5810, 2017.
- [11] Y. Gao, M. A. Maraci, and J. A. Noble, "Describing ultrasound video content using deep convolutional neural networks," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016, pp. 787–790.
- [12] J. Shi, S. Zhou, X. Liu, Q. Zhang, M. Lu, and T. Wang, "Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset," *Neurocomputing*, vol. 194, pp. 87–94, 2016.
- [13] A. R. Jamieson, K. Drukker, and M. L. Giger, "Breast image feature learning with adaptive deconvolutional networks," in *Medical Imaging 2012: Computer-Aided Diagnosis*, vol. 8315. International Society for Optics and Photonics, 2012, p. 831506.
- [14] W. Huang, C. P. Bridge, J. A. Noble, and A. Zisserman, "Temporal heart-net: towards human-level automatic analysis of fetal cardiac screening video," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 341–349.
- [15] J. Shin, N. Tajbakhsh, R. T. Hurst, C. B. Kendall, and J. Liang, "Automating carotid intima-media thickness video interpretation with convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2526–2535.
- [16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, Munich, Germany, 5-9 October, 2015, pp. 234–241.
- [17] Y. Zhang, M. T. Ying, L. Yang, A. T. Ahuja, and D. Z. Chen, "Coarse-to-fine stacked fully convolutional nets for lymph node segmentation in ultrasound images," in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2016, pp. 443–448.
- [18] Y. Li, R. Xu, J. Ohya, and H. Iwata, "Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 1485–1488.
- [19] O. Bernard, B. Touil, A. Gelas, R. Prost, and D. Friboulet, "A rbf-based multiphase level set method for segmentation in echocardiography using the statistics of the radiofrequency signal," in *2007 IEEE International Conference on Image Processing*, vol. 3. IEEE, 2007, pp. III–157.
- [20] G. Jacob, J. A. Noble, C. Behrenbruch, A. D. Kelion, and A. P. Banning, "A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography," *IEEE transactions on medical imaging*, vol. 21, no. 3, pp. 226–238, 2002.
- [21] D. A. Lichtenstein, "Blue-protocol and falls-protocol: two applications of lung ultrasound in the critically ill," *Chest*, vol. 147, no. 6, pp. 1659–1670, 2015.
- [22] D. A. Lichtenstein, G. A. Meziere, J.-F. Lagoueyte, P. Biderman, I. Goldstein, and A. Gepner, "A-lines and b-lines: lung ultrasound as a bedside tool for predicting pulmonary artery occlusion pressure in the critically ill," *Chest*, vol. 136, no. 4, pp. 1014–1020, 2009.
- [23] R. J. van Sloun and L. Demi, "Localizing b-lines in lung ultrasonography by weakly supervised deep learning, in-vivo results," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 957–964, 2019.
- [24] R. van Sloun and L. Demi, "Deep learning for automated detection of b-lines in lung ultrasonography," *The Journal of the Acoustical Society of America*, vol. 144, no. 3, pp. 1668–1668, 2018.
- [25] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli *et al.*, "Deep learning for

- classification and localization of covid-19 markers in point-of-care lung ultrasound,” *IEEE Transactions on Medical Imaging*, 2020.
- [26] J. Born, G. Brändle, M. Cossio, M. Disdier, J. Goulet, J. Roulin, and N. Wiedemann, “Pocovid-net: automatic detection of covid-19 from a new lung ultrasound imaging dataset (pocus),” *arXiv preprint arXiv:2004.12084*, 2020.
- [27] J. Born, N. Wiedemann, M. Cossio, C. Buhre, G. Brändle, K. Leidermann, A. Aujayeb, M. Moor, B. Rieck, and K. Borgwardt, “Accelerating detection of lung pathologies with explainable ultrasound image analysis,” *Applied Sciences*, vol. 11, no. 2, p. 672, 2021.
- [28] M. Correa, M. Zimic, F. Barrientos, R. Barrientos, A. Román-Gonzalez, M. J. Pajuelo, C. Anticona, H. Mayta, A. Alva, L. Solis-Vasquez *et al.*, “Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition,” *PLoS One*, vol. 13, no. 12, p. e0206410, 2018.
- [29] D. Tzatalin, “Labelimg,” <https://github.com/tzatalin/labelImg>. Accessed 14 June 2021, 2020.
- [30] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.
- [31] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random erasing data augmentation,” in *AAAI*, 2020, pp. 13 001–13 008.
- [32] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size,” *arXiv preprint arXiv:1602.07360*, 2016.
- [33] O. Köpüklü, N. Kose, A. Gunduz, and G. Rigoll, “Resource efficient 3d convolutional neural networks,” in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, 2019, pp. 1910–1919.
- [34] A. Shrikumar, P. Greenside, and A. Kundaje, “Learning important features through propagating activation differences,” in *International Conference on Machine Learning*. PMLR, Sydney, Australia, 6–11 August 2017, pp. 3145–3153.
- [35] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-cam: Visual explanations from deep networks via gradient-based localization,” in *Proceedings of the IEEE international conference on computer vision*, Venice, Italy, 22–29 October 2017, pp. 618–626.
- [36] M. Sundararajan, A. Taly, and Q. Yan, “Axiomatic attribution for deep networks,” in *International Conference on Machine Learning*. PMLR, Sydney, Australia, 6–11 August 2017, pp. 3319–3328.

Výskum komunikačných architektur v prostredí internetu vecí

Rastislav Petija

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovenská republika
Email: rastislav.petija@tuke.sk

František Jakab

Katedra počítačov a informatiky
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Košice, Slovenská republika
Email: frantisek.jakab@tuke.sk

Abstrakt—Tento článok sa venuje návrhu, implementácií a experimentálnemu overeniu architektúry pre zber, transport a detekciu anomálií v prostredí internetu vecí. Pre účely transportu informácií o dátových tokoch boli využité protokoly IPFIX, pre bežné neobmedzené prostredie a TinyIPFIX pre obmedzené prostredie. Detekcia anomálií bola realizovaná využitím algoritmov strojového učenia. Hlavným prínosom práce je vznik univerzálneho riešenia, ktoré je schopné prepojiť monitorovanie prostredia internetu vecí s bežnými sieťami. Nosná časť práce bola venovaná tvorbe IPFIX/TinyIPFIX kolektora, ktorý zároveň predstavuje bod v sieti, na ktorom bude dochádzať k detekcii anomálií. Vďaka procesu mediácie, ktorý je podporovaný na kolektore je možné transparentne prenášať dáta medzi bežným sieťovým prostredím a prostredím internetu vecí. Experiment potvrdil funkčnosť prenosovej architektúry spolu s každým jej komponentom a zároveň univerzálnosť pri detekcii bez ohľadu na pozorované prostredie. Vďaka implementácií protokolov IPFIX/TinyIPFIX na zariadenie LoRa brány je možné tento systém využiť zároveň aj pre zber a spracovanie dát z UAV zariadení.

Kľúčové slová—internet vecí, dátové toky, detekcia anomálií, IPFIX, TinyIPFIX

Abstract—This article deals with the design, implementation and experimental verification of an architecture for the collection, transport and detection of anomalies in the Internet of Things environment. For the purpose of transporting information about data flows, the IPFIX protocols were used, for a common unrestricted environment and TinyIPFIX protocol for a restricted environment. Anomaly detection was performed using machine learning algorithms. The main benefit of this work is the creation of a universal system that is able to connect the monitoring of the Internet of Things environment with common networks. The main part of the work was devoted to the creation of an IPFIX / TinyIPFIX collector, which also represents a point in the network where anomalies will be detected. Thanks to the mediation process that is supported on the collector, it is possible to transparently transfer data between the common network environment and the Internet of Things environment. The experiment confirmed the functionality of the transmission architecture together with each of its components and at the same time the versatility of detection regardless of the observed environment. Thanks to the implementation of IPFIX / TinyIPFIX protocols on the LoRa gateway device, this system can also be used for data collection and processing from UAV devices.

Keywords—internet of things, data flows, anomaly detection, IPFIX, TinyIPFIX

I. ÚVOD

Koncept internetu vecí (IoT) predstavuje rýchlo sa rozvíjajúcu oblasť informatiky. Aktuálnym trendom je pripájať čoraz väčšie množstvo zariadení k počítačovým sieťam, aby s nimi bolo možné vzdialene komunikovať. V počiatkoch vývoja tohto smeru sa hlavná pozornosť výskumu venovala spôsobom transportu dát rôznymi prostrediami a tvorbou inteligentných zariadení. Menšia pozornosť bola venovaná monitorovaniu tohto prostredia a zabezpečeniu kvality služieb pri transporte dát. Cieľom tejto práce bude navrhnúť, implementovať a experimentálne overiť architektúru a systém pre zber a transport informácií o dátových tokoch s ich následným spracovaním, kde primárnym cieľom bude detekcia anomálií. Daný systém bude univerzálny, založený na otvorených štandardoch, kde dôraz bude kladený na prepojenie monitorovania prostredia internetu vecí s monitorovacími nástrojmi využívanými v bežných počítačových sieťach. Pre transport dát o dátových tokoch budú využívané protokoly, v IoT prostredí TinyIPFIX a v prostredí bežných sietí protokol IPFIX. V čase písania tejto práce nebola dostupná implementácia protokolu TinyIPFIX. Keďže sa ale tento protokol na základe analýzy javil ako vhodný, ktorý umožní značné šetrenie spotreby šírky pásma, čo predstavuje pre obmedzené IoT zariadenia veľkú výhodu, tak v rámci tejto práce bude realizovaná implementácia tohto protokolu pre IoT zariadenia. Splnenie stanovených cieľov vyžadovalo tiež vytvorenie exportéra, kolektora a mediátora. Tieto komponenty umožnia transparentný prenos údajov z IoT prostredia do prostredia bežných sietí. Podobne ako v predchádzajúcom prípade, všetky tieto komponenty budú vytvorené v rámci tejto práce, na základe príslušných štandardov. Detekcia anomálií bude implementovaná priamo na zariadení, kde bude spustený kolektor, kde okamžite po prijatí nových dát bude na týchto dátach realizovaná analýza, ktorej výsledkom bude označenie daného dátového toku značkou pre konkrétnu anomáliu.

II. SÚVISIACE PRÁCE

Medzi aktuálne problémy, ktorým sa začínajú venovať výskumy v súvislosti s IoT prostredím je možné zaradiť problematiku klasifikácie dátovej prevádzky, detekciu útokov a hodnotenie kvality IoT prostredia.

A. Klasifikácia dátovej prevádzky v IoT prostredí

Ku klasifikácií je možné pristupovať štyrmi hlavnými prístupmi. Prvým je portovo založená klasifikácia, ktorá využíva mapovanie medzi použitými portami transportnej vrstvy a príslušnými aplikáciami. Tento prístup v súčasnosti nie je vhodný vzhľadom k tomu, že väčšina aplikácií dynamicky mení čísla portov, ako uvádzajú autori v [1]. Ďalším presnejším typom klasifikácie je obsahovo založená klasifikácia, počas ktorej dochádza k hĺbkovej kontrole celého prijatého paketu [2]. Táto metóda je ale pamäťovo a výpočtovo najnáročnejšia. Tretím prístupom je klasifikácia na základe

správania sa danej komunikácie, ktorá najčastejšie využíva rôzne typy strojového učenia [3] pre identifikáciu daného typu prevádzky. Posledným prístupom, ktorý bude využitý aj v rámci tejto práce je štatisticky založený prístup. Podľa autorov v [4] je práve tento prístup najvhodnejší pre IoT prostredie vzhľadom k tomu, že pracuje len s štatistickými údajmi bez potreby hlbokého prehládavania paketov. Publikácia [5] uvádza, že dáta pochádzajúce z IoT prostredia majú väčšiu komplexitu ako dáta z bežných sietí. Klasifikáciou útokov sa zaoberali v publikácii [6], kde bolo využité hlboké učenie využívajúce konvolučné siete.

B. Detekcia útokov v IoT prostredí

Rozšírenie bežných sietí o IoT prostredie prináša nové možnosti pre útočníkov a realizáciu útokov. Preto sa množstvo výskumov venuje práve tomuto smeru. Autori v [7] uvádzajú priamy súvis medzi vývojom inteligentných zariadení a vznikajúcimi bezpečnostnými incidentmi. Vo všeobecnosti je možné detekčné algoritmy rozdeliť podľa ich fungovania na algoritmy založené na detekcii hraničných hodnôt, klasifikačných pravidiel a hľadajúcich podobnosti. Výskum [8] poukazuje na to, že IoT prostredie môže byť často zdrojom pre DDoS útoky. Ďalšou hrozbou, na ktorú poukazujú v [9] je možná strata súkromia, keďže IoT zariadenia majú často za úlohu monitorovať prostredie, v ktorom sa pohybujú ľudia. Zabezpečením systémov sa venujú tiež v prácach [10] a [11]. Jedným z prístupov, ako odhaliť nedôveryhodné zariadenia je využitie štatistickej klasifikácie skúmania paketov v spolupráci so strojovým učением [12]. Detekciu rôznych typov útokov sa tiež zaoberali autori v [13], pričom podobne ako v množstve ďalších výskumov na to používali rôzne metódy strojového učenia. Medzi veľmi efektívne metódy detekcie útokov sa zaradila analýza dátových tokov, ktorá bola využitá vo výskume [14], kde bolo na základe informácií získaných protokolmi NetFlow a IPFIX, ktoré boli neskôr spracované modelmi strojového učenia, odhalené ťaženie kryptomien.

C. Hodnotenie kvality IoT prostredia

Ako uvádzajú autori v [15] architektúru pre vyhodnocovanie spoľahlivosti je veľmi náročné implementovať do IoT prostredia vzhľadom k množstvu obmedzení [16] a rôznorodosti použitých protokolov [17]. Hlavným odporúčaním pri analýze dát je implementácia detekčných systémov, čo najbližšie zdroju [18]. V rámci tejto práce bude detekcia realizovaná na základe dát získaných protokolmi IPFIX, ktorý je definovaný v štandarde RFC 7011 [19] a TinyIPFIX definovanom v RFC 8272 [20]. Definovanie univerzálneho prístupu je náročné preto, lebo každé prostredie má iné nároky na kvalitu. V rámci tejto práce bude realizovaný experiment, ktorý bude mať za úlohu detegovať anomálie v LoRa prostredí na základe atribútov prenosu ako RSSI, SNR, veľkosti prenášaných dát a doby prenosu. Rozborom dôležitých polí a procesov pri IPFIX a TinyIPFIX komunikáciách sa venovali autori v [21]. Jednou z architektúr pre monitorovanie IoT infraštruktúr sa zaoberali autori v publikácii [22].

Na základe vyššie spomenutých poznatkov a nedostatkov v existujúcich riešeniach sa budú nasledujúce kapitoly zaoberať návrhom, implementáciou a overením systému pre zber a transport kvalitatívnych parametrov z IoT prostredia a ich následnou analýzou, ktorá umožní detekciu anomálií.

III. UNIVERZÁLNA KOMUNIKAČNÁ A MONITOROVACIA ARCHITEKTÚRA

Pre splnenie vyššie uvedených cieľov bolo najskôr potrebné navrhnuť architektúru, ktorá by umožnila prepojenie monitorovania bežných sietí s IoT prostredím. Zároveň by mala umožňovať transparentný prenos dát medzi prostrediami. Tento proces bude zabezpečený mediálnym mechanizmom. Celková architektúra prostredia, do ktorého bude daný systém nasadený je znázornená na Obr. 1.

Ako je možné vidieť, tak komponenty, u ktorých sa predpokladá vzájomná interakcia sú rôzne druhy senzorov alebo akčných členov,

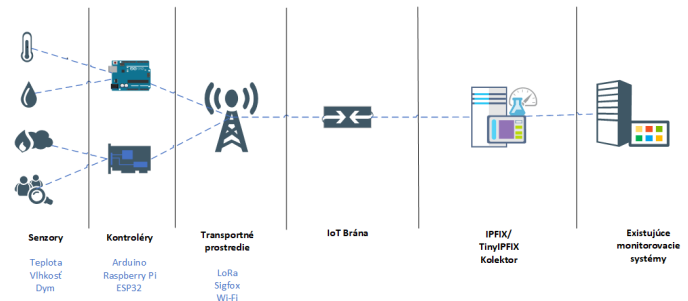


Fig. 1. Celková architektúra systému

ktoré budú mať za úlohu získavať údaje z prostredia, prípadne v prostredí niečo meniť. Tieto koncové komponenty sa pripájajú k nejakému druhu zariadení, ktoré budú schopné sprostredkovať komunikáciu medzi senzorom / akčnou jednotkou a riadiacim systémom. Medzi tento typ zariadení je možné zaradiť Arduino, Raspberry Pi alebo ESP32. Tieto zariadenia pretransformujú prijaté dáta do podoby, ktorú bude možné preniesť na väčšiu vzdialenosť. Technológia, s ktorými sa najčastejšie zvykne pracovať sú LoRa, Sigfox a Wi-Fi. Prijem týchto dát je realizovaný na zariadení, ktoré nazývame IoT bránou, ktorá predstavuje rozhranie medzi bežnou sieťou a IoT prostredím. Prijaté dáta budú následne spracované hlavným prvkom nášho systému a to IPFIX/TinyIPFIX kolektorom, ktorý bude schopný spracovať a interpretovať dáta o dátových tokoch. Na tomto istom mieste bude možné zároveň vykonať detekciu anomálií využitím strojového učenia. V prípade potreby detailnejšej analýzy bude tento kolektor pracovať aj ako exportér a využitím mediálneho mechanizmu bude možné dáta transparentne preniesť na existujúce monitorovacie zariadenia.

IV. NÁVRH KOMPONENTOV KOMUNIKAČNÉHO SYSTÉMU

Keďže na zber a transport údajov o dátových tokoch budú využívané protokoly TinyIPFIX a IPFIX, tak práve procesy súvisiace s týmito protokolmi bolo potrebné navrhnuť a implementovať pre IoT prostredie. Konkrétne sa jedná o exportér, kolektor a mediátor.

A. IPFIX/TinyIPFIX exportér

Prvým komponentom, ktorý je potrebný na odosielanie dát z obmedzených IoT zariadení je exportér. V našom prípade, aby bola zabezpečená univerzálnosť a rozšíriteľnosť, tak exportovací proces bol vytvorený na základe štandardov. Stavový diagram IPFIX a TinyIPFIX exportéra znázorňuje Obr. 2. Univerzálnosť exportéra je zabezpečená vďaka konfiguračnému súboru, ktorý je možné použiť na nastavenie veľkého množstva funkcií. Po spustení aplikácie exportéra dôjde k inicializácii a načítaniu konfiguračného súboru. Následne na základe údajov z konfigurácie sa otvorí súbor s dátami, ktoré je potrebné preniesť. Tieto dáta je najskôr potrebné rozparsovať a pretransformovať do podoby pre IPFIX/TinyIPFIX protokol. Ďalej dôjde ku kontrole, či sa jedná o prvé spustenie, alebo opakované. V prípade prvého spustenia je potrebné odoslať IPFIX/TinyIPFIX šablónu, ktorá bude definovať význam neskôr prenášaných dát. Vďaka tomuto prístupu je možné šetriť šírku pásma a zároveň prenášať rôzne typy dát, nie len dáta charakterizujúce dátové toky. V prípade, že kolektor už spustený bol, tak sa overí, či uplynul časovač na aktualizáciu šablóny. Ak ubehol, tak sa odošle znova šablóna. Ak neubehol, tak sa odošlú samotné dáta. To, či budú dáta alebo šablóna odosielané v IPFIX alebo TinyIPFIX formáte určuje konfiguračný súbor. Po odoslaní dát sa overí úspešnosť procesu. V prípade, že došlo k chybe, tak bude vygenerovaná chybová správa. V opačnom prípade sa na chvíľu zastaví činnosť exportéra a následne sa dostávame na začiatok procesu čítaním ďalšieho záznamu, ktorý chceme odoslať.

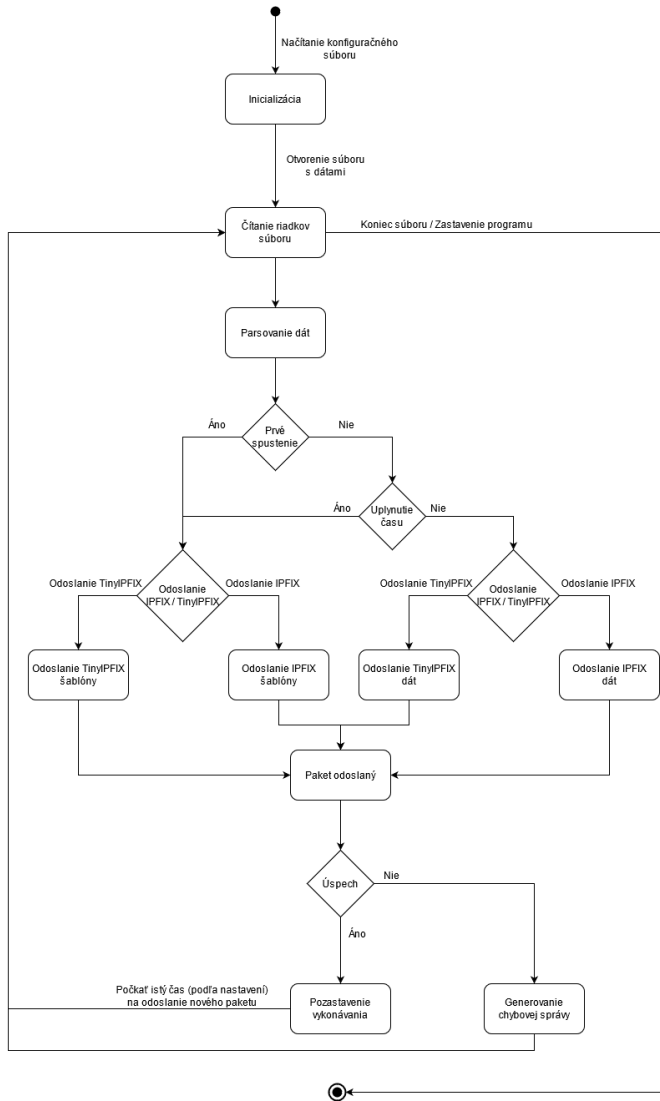


Fig. 2. Stavový diagram exportéra

Exportér je možné vypnúť v akomkoľvek bode vykonávania. Aplikácia pre proces exportéra bola na základe návrhu vytvorená pre obmedzené IoT zariadenia, v našom prípade ESP32 platformu, ale aj pre bežné linuxové systémy, aby bolo možné daný systém nasaďiť kdekoľvek.

B. IPFIX/TinyIPFIX kolektor

Kolektor predstavuje pre túto prácu hlavný prvok komunikačnej architektúry. Úlohou kolektora je prijať dáta v IPFIX alebo TinyIPFIX formáte a uložiť ich na určené miesto, prípadne ich transparentne preposlať ďalej, či už v pôvodnej alebo konvertovanej podobe v IPFIX formáte. Stavový diagram navrhnutého kolektora je možné vidieť na nasledujúcom Obr. 3. Podobne ako exportér, tak aj kolektor pracuje na základe konfiguračného súboru, v ktorom je možné nastavovať množstvo funkcionalít. Po spustení kolektora začne príslušný proces prijímať UDP správy, ktoré sú používané na transport dát protokolmi IPFIX a TinyIPFIX. Po prijatí dát dochádza k ich prečítaniu a detekcii, či prijaté dáta boli vo formáte protokolu IPFIX alebo TinyIPFIX. V oboch prípadoch sa následne overí, či daný kolektor nemá plniť zároveň úlohu mediátora. Ak áno, tak v prípade prijatia IPFIX dát sú tieto dáta transparentne, bez zmeny odoslané. V prípade

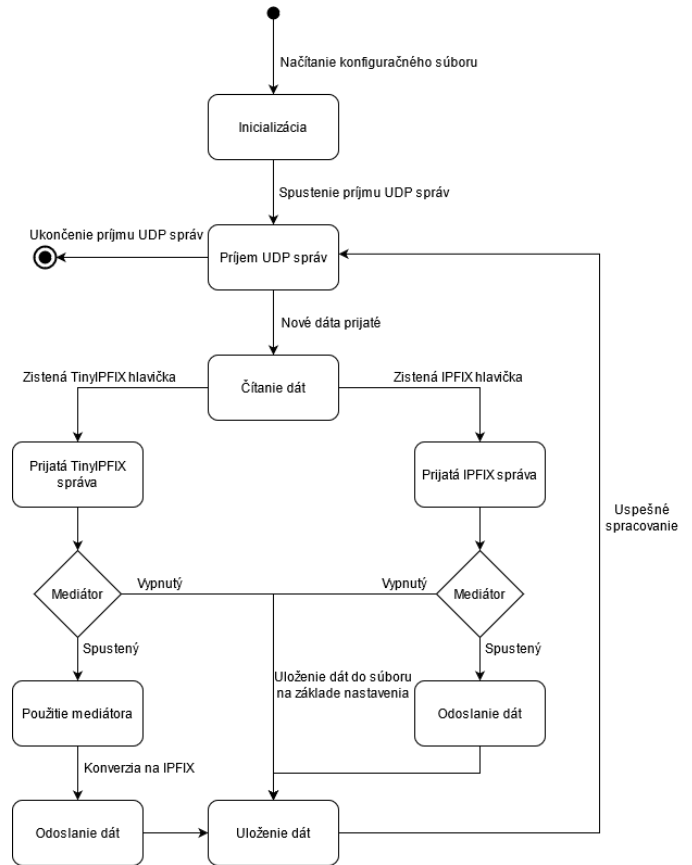


Fig. 3. Stavový diagram kolektora

protokolu TinyIPFIX mediálny mechanizmus zabezpečí konverziu dát do IPFIX formátu a rovnako ich odošle. Po odoslaní dát sú prijaté dáta zároveň uložené na príslušné miesto do pamäte. K uloženiu dát dochádza aj v prípade, ak mediálny proces nebol spustený. Po uložení dát sa proces kolektora opakuje čakaním na ďalšie dáta.

Navrhnutý kolektor je rovnako ako exportér implementovaný formou aplikácie ako pre obmedzené IoT prostredie, tak aj pre bežné linuxové systémy. Zdôrazniť treba to, že v čase písania tejto práce neexistovala voľne dostupná implementácia spomenutých komponentov pre TinyIPFIX protokol.

C. IPFIX/TinyIPFIX mediátor

Posledným procesom, ktorý je potrebné implementovať je mediálny proces. Tento proces nie je dôležitý pre základnú funkcionálnu transportnej infraštruktúry, ale pre rozšírenie funkcionality systému. Mediácia umožní prepojiť monitorovanie viacerých rôznych prostredí a zároveň umožní koexistenciu a vzájomnú kompatibilitu protokolov IPFIX a TinyIPFIX, čím dovoľí prepojiť náš systém s už existujúcimi implementáciami. Stavový diagram mediálneho procesu je znázornený na Obr. 4. Proces mediácie je súčasťou kolektora. Začína sa parsovaním dát, kde je potrebné získať informácie z prijatých hlavičiek a rozhodnúť o aký formát dát sa jedná. Ak boli prijaté dáta v IPFIX formáte, tak dochádza k transparentnému odoslaniu dát. V prípade prijatia dát v TinyIPFIX formáte začína proces parsovania dát a vytvorenia TinyIPFIX hlavičky. Následne je potrebné zistiť aké dáta boli prijaté. Ak boli prijaté dáta, tak sa vytvorí hlavička dátovej sady, ku ktorej sa pridajú dáta. Ak bola prijatá šablóna, tak sa vytvorí hlavička šablóny a následne aj celá šablónová sada. Nakoniec sa vytvorí IPFIX paket, ktorý je odoslaný na cieľové zariadenie.

Spustenie procesu mediácie a definovanie cieľovej IPv4 adresy, na

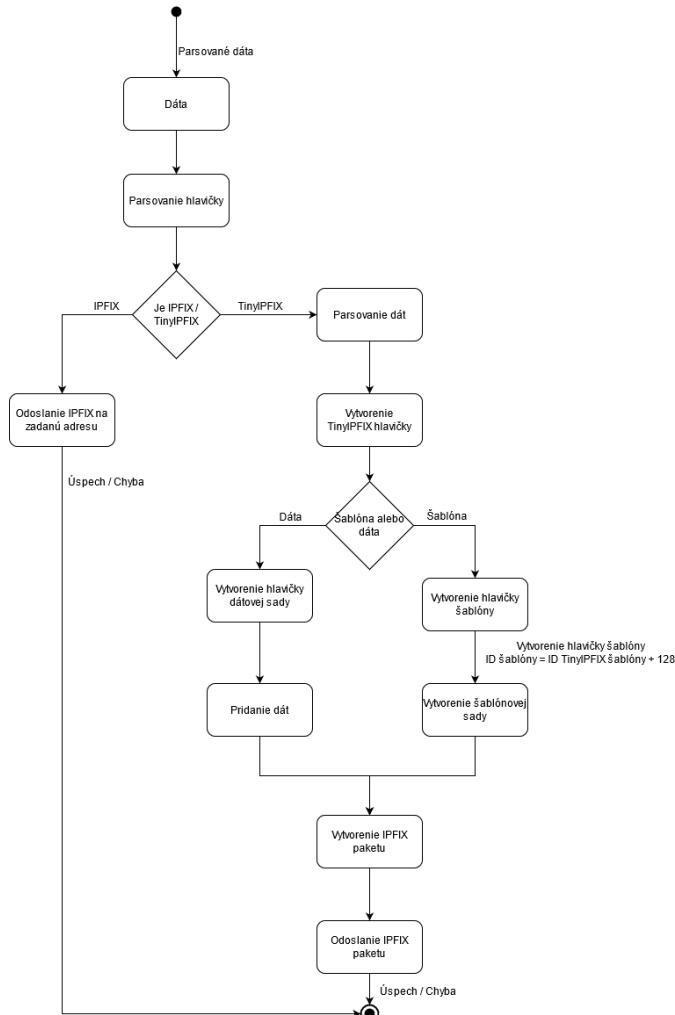


Fig. 4. Stavový diagram mediátora

ktorý je potrebné dáta odosielať sa realizuje využitím konfiguračného súboru. Rovnako je v konfiguračnom súbore potrebné definovať ID pozorovanej domény, aby bol splnený štandard. Vzhľadom k tomu, že informácie o dátových tokoch nie je potrebné odosielať späť do IoT prostredia, tak nie je definovaný proces mediácie z IPFIX formátu na TinyIPFIX formát.

V. IMPLEMENTÁCIA SYSTÉMU

Hlavnou úlohou monitorovacieho systému je detekcia anomálií, pričom stavový diagram je znázornený na Obr. 5.

Celkový systém pozostáva z kolektora a detekčnej časti. Oba procesy budú spustené na tom istom zariadení, kde kolektor má za úlohu prijímať dáta a ukladať ich do súboru. Detekčný systém ich má následne prečítať a označiť príslušnou značkou anomálie. Detekcia anomálií bude pracovať na princípe strojového učenia. Proces detekcie začína načítaním konfiguračného súboru. Následne je načítaná dátová sada, ktorá sa najskôr spracováva, následne sa vykoná fáza vzorkovania, aby sa zabránilo práci s nevyváženou dátovou sadou. Takto upravené dáta sa následne využijú na tréning modelov strojového učenia, kde súčasťou tejto prípravnej fázy je zároveň otestovanie úspešnosti detekcie v danom prostredí. Natrénovaný model je následne využitý na robenie predikcie pre novo prijaté dáta, čoho výsledkom je označenie prijatých dát značkou identifikujúcou príslušnú anomáliu. Tento systém bol navrhnutý univerzálne, kde

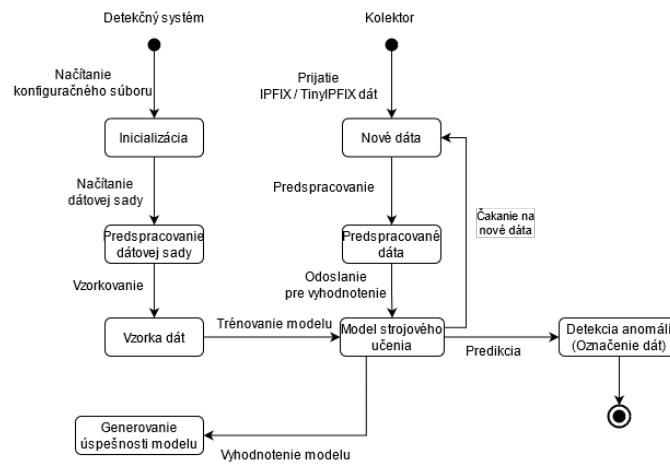


Fig. 5. Stavový diagram detekčného systému

jediný predpoklad je, aby boli dostupné dáta pre tréning, teda dáta, ktoré obsahujú označené dáta príslušného prostredia.

VI. EXPERIMENTÁLNE OVERENIE SYSTÉMU

Overenie systému bolo realizované v dvoch úrovniach. Najskôr bolo overené transportné prostredie v podobe overenia implementácie protokolov IPFIX a TinyIPFIX. Následne bola overená úspešnosť samotného procesu identifikácie anomálií.

A. Overenie úspešnosti implementácie IPFIX a TinyIPFIX protokolov

Prvé overenie spočívalo v odoslaní dát našim exportérom a následným odchytením dát sieťovým nástrojom Wiresharkom. Ako je možné vidieť na Obr. 6, tak tento nástroj rozoznal nami odoslané dáta, čo je jasným dôkazom dodržania všetkých pravidiel pri implementácii protokolov na základe štandardov. Týmto spôsobom sa ale nedala

```
Cisco NetFlow/IPFIX
Version: 10
Length: 36
> Timestamp: Feb 28, 2021 13:49:15.000000000 Stredoeurópsky čas (normálny)
FlowSequence: 0
Observation Domain Id: 1322
Set 1 [id=2] (Data Template): 1022
  FlowSet Id: Data Template (v10 [IPFIX]) (2)
  FlowSet Length: 20
  Template (Id = 1022, Count = 2)
    Template Id: 1022
    Field Count: 2
    Field (1/2): IP_SRC_ADDR
      0... .. = Pen provided: No
      .000 0000 0000 1000 = Type: IP_SRC_ADDR (8)
      Length: 4
    Field (2/2): 228 [pen: Example Enterprise Number for Documentation Use]
      1... .. = Pen provided: Yes
      .000 0000 1110 0100 = Type: 228 [pen: Example Enterprise Number for Documentation Use]
      Length: 5
      PEN: Example Enterprise Number for Documentation Use (32473)
```

Fig. 6. Výstup z nástroja Wireshark

overiť platnosť protokolu TinyIPFIX. Pre jeho overenie bol využitý mediálny mechanizmus, kde dáta boli pôvodne odoslané TinyIPFIX exportérom na univerzálny kolektor. Prijaté dáta boli manuálne skontrolované na kolektore, kde bola pozorovaná 100% zhoda. Scenár pre overenie mediácie je znázornený na Obr. 7.

Prijaté dáta boli následne cez mediálny proces odoslané na externý IPFIX kolektor, ktorý bol schopný dekodovať a spracovať dáta, čo jasne potvrdilo funkčnosť všetkých troch komponentov nami implementovanej infraštruktúry.

Ďalším experimentálnym overením bolo porovnanie protokolov IPFIX a TinyIPFIX z pohľadu spotreby šírky pásma, kde podľa

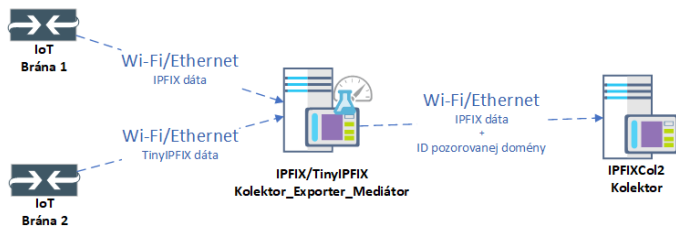


Fig. 7. Scenár pre overenie mechanizmu mediácie

predpokladať sa potvrdilo, že protokol TinyIPFIX je schopný šetriť šírku pásma a je teda vhodnejší pre obmedzené IoT prostredia. Prehľad daného porovnania je uvedený v Tab. I.

TABLE I
CELKOVÉ POROVNANIE SPOTREBY ŠÍRKY PÁSMO PRE PROTOKOLY IPFIX A TINYIPFIX

	IPFIX	TinyIPFIX
Hlavička	128b (16B)	min: 24b (3B) max: 40b (5B)
Hlavička sady	32b (4B)	16b (2B)
Hlavička šablóny	32b (4B)	15b
Špecifikátor poľa	64b (8B)	64b (8B)
Šablóna		
Celkovo	192b (24B)	min: 56b (7B) max: 72b (9B)
Úspora		min: 120b (15B) max: 136b (17B)
Dáta		
Celkovo	160b (20B)	min: 40b (5B) max: 56b (7B)
Úspora		min: 104b (13B) max: 120b (15B)

Ako je možné vidieť, tak šetrenie vzniká ako pri odosielaní šablón, tak aj pri posielaní dátových sád. Maximálne je možné pri šablóne ušetriť až 17B a pri dátovej sade až 15B.

B. Overenie úspešnosti detekčného mechanizmu

Reálne prostredie, v ktorom bol testovaný detekčný mechanizmus súviselo s LoRa technológiou. Scenár experimentu je znázornený na Obr. 8. Inteligentnými elektromermi namerané dáta boli odoslané na

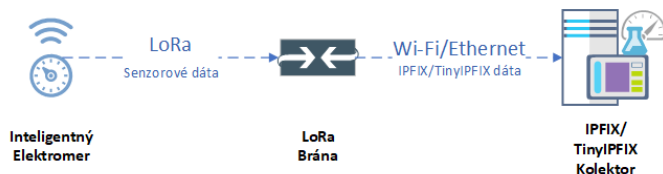


Fig. 8. Scenár experimentu pri LoRa technológií

LoRa bránu. Z tejto brány boli dáta odoslané exportovacím procesom vo formáte IPFIX/TinyIPFIX na náš univerzálny kolektor. Na tomto kolektore následne prebiehala analýza využitím strojového učenia. Vyskúšaných bolo viacero modelov strojového učenia, kde cieľom bolo odhaliť rôzne typy anomálií. Nasledujúca matica zobrazená na Obr. 9 znázorňuje úspešnosť dosiahnutej detekcie na novo prijatých dátach. Systém má ale schopnosť generovať aj informáciu o tom, ako dobre sa naučil štandardný dataset. Úspešnosť nášho tréningu je zobrazená na Obr. 10. Posledným výstupom daného

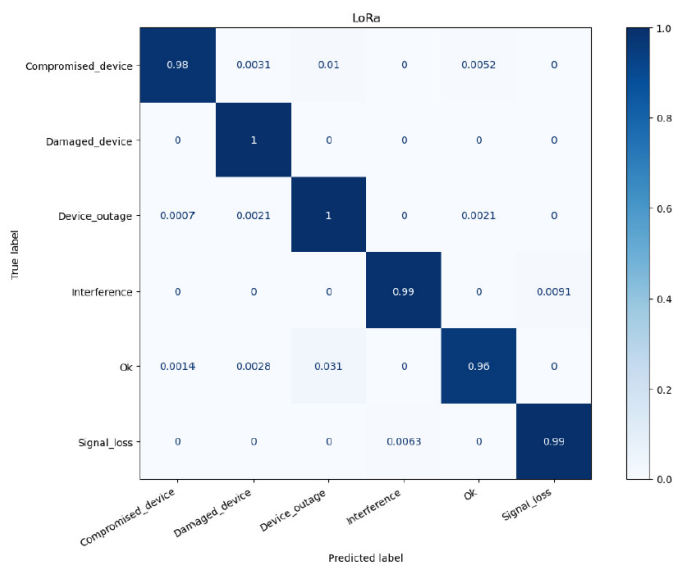


Fig. 9. Úspešnosť detekcie na LoRa dátach

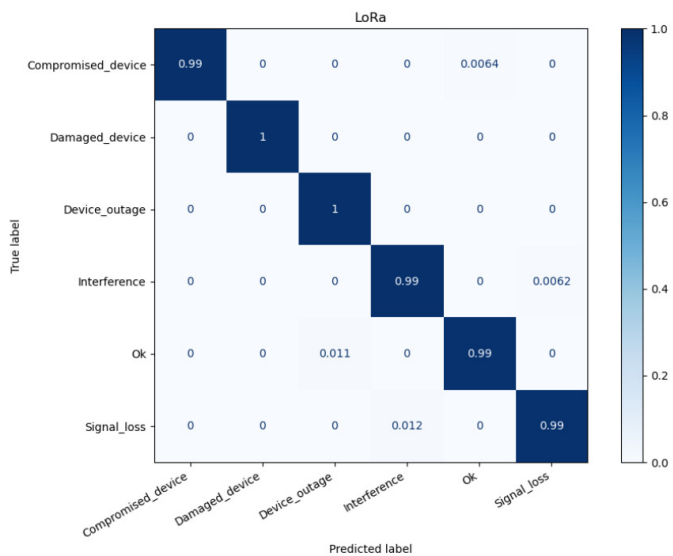


Fig. 10. Úspešnosť tréningu na LoRa dátach

detekčného systému pre bližší popis úspešnosti predikcie je nasledujúca Tab. II znázorňujúca rôzne metriky. Ako je možné vidieť, tak pri všetkých testoch boli dosiahnuté výsledky bližšie sa úspešnosti 100%. Spôsobené to bolo tým, že anomálie vnesené do daných dátových sád boli generované synteticky nami. V prípade nasadenia systému do reálneho prostredia by sa dali očakávať o niečo horšie výsledky. V rámci experimentov bolo použitých viacero modelov strojového učenia, no najlepšie výsledky boli dosiahnuté modelmi Random Forest, K-Nearest Neighbors a XGBoost.

VII. ZÁVER

Výsledkom tejto práce je navrhnutá, implementovaná a experimentálne overená architektúra spolu so systémom pre transport a zber kvalitatívnych informácií z IoT prostredia. Na základe zozbieraných dát je možné vykonávať detekciu anomálií, ktorá môže byť použitá na hodnotenie kvality IoT prostredia. Za účelom transportu a zberu dát boli implementované protokoly IPFIX a TinyIPFIX do

TABLE II
METRIKY PREDIKCIÍ NA TESTOVACÍCH DÁTACH

	Precision	Recall	F1-score	Support
Compromised_device	1.00	0.98	0.99	2858
Damaged_device	0.99	1.00	0.99	1429
Device_outage	0.95	1.00	0.97	1430
Interference	0.99	0.99	0.99	1430
OK	0.99	0.96	0.98	1430
Signal_loss	0.99	0.99	0.99	1430
Accuracy			0.99	10007
Macro avg	0.98	0.99	0.99	10007
Weighted avg	0.99	0.99	0.99	10007

IoT prostredia v podobe exportéra a kolektora. Detekcia anomálií prebieha využitím modelov strojového učenia. Za hlavný prínos sa dá považovať vytvorenie architektúry a univerzálneho systému, ktorý je možné používať neskôr pre množstvo ďalších výskumov. Príkladom jednej z aplikačných domény je aj možnosť nasadenia systému, vďaka jeho podpore pre LoRa prostredie, aj pre zber a analýzu dát z UAV zariadení. Aktuálny výskum je možné ďalej rozširovať napríklad návrhom prístupov ako využiť metódy strojového učenia založené na princípe učenia bez učiteľa, ktoré by mohli pomôcť odhaľovať novo vzniknuté anomálie, ktoré ešte neboli systémom pozorované.

POĎAKOVANIE

Táto publikácia bola podporená z operačného programu Integrovaná infraštruktúra v rámci projektu: Inteligentné operačné a spracovateľské systémy pre UAV, kód ITMS2014+: 313011V422 a spolufinancovaná Európskym fondom regionálneho rozvoja.

LITERATÚRA

- [1] KARAGIANNIS, T.; BROIDO, A.; BROWNLEE, N.; CLAFFY, K. C.; FALOUTSOS, M. Is P2P dying or just hiding? In: IEEE Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE, [b. r.]. Dostupné z doi: 10.1109/glocom.2004.1378239.
- [2] KHATER, Noora Al; OVERILL, Richard E. Network traffic classification techniques and challenges. In: 2015 Tenth International Conference on Digital Information Management (ICDIM). IEEE, 2015. Dostupné z doi: 10.1109/icdim.2015.7381869.
- [3] BERMOLAN, Paola; MELLIA, Marco; MEO, Michela; ROSSI, Dario; VALENTI, Silvio. Abacus: Accurate behavioral classification of P2P-TV traffic. *Computer Networks*. 2011, roc. 55, c. 6, s. 1394–1411. Dostupné z doi: 10.1016/j.comnet.2010.12.004.
- [4] ZHANG, Jun; CHEN, Xiao; XIANG, Yang; ZHOU, Wanlei; WU, Jie. Robust Network Traffic Classification. *IEEE/ACM Transactions on Networking*. 2015, roc. 23, c. 4, s. 1257–1270. Dostupné z doi: 10.1109/tnet.2014.2320577.
- [5] YAO, Haipeng; GAO, Pengcheng; WANG, Jingjing; ZHANG, Peiyang; JIANG, Chunxiao; HAN, Zhu. Capsule Network Assisted IoT Traffic Classification Mechanism for Smart Cities. *IEEE Internet of Things Journal*. 2019, roc. 6, c. 5, s. 7515–7525. Dostupné z doi: 10.1109/jiot.2019.2901348.
- [6] L-HAJAJ, Qasem Abu; ZEIN-SABATTO, Saleh. An Efficient Deep-Learning-Based Detection and Classification System for Cyber-Attacks in IoT Communication Networks. *Electronics*. 2020, roc. 9, c. 12, s. 2152. Dostupné z doi: 10.3390/electronics9122152.
- [7] SICARI, Sabrina; RIZZARDI, Alessandra; MIORANDI, Daniele; CAPIELLO, Cinzia; COEN-PORISINI, Alberto. A secure and quality-aware prototypical architecture for the Internet of Things. *Information Systems*. 2016, roc. 58, s. 43–55. Dostupné z doi: 10.1016/j.is.2016.02.003.
- [8] CVITIC, Ivan; PERAKOVIC, Dragan; PERIŠA, Marko; BOTICA, Mate. Smart Home IoT Traffic Characteristics as a Basis for DDoS Traffic Detection. In: Proceedings of the 3rd EAI International Conference on Management of Manufacturing Systems. EAI, 2018. Dostupné z doi: 10.4108/eai.6-11-2018.2279336.
- [9] SHAHID, Mustafizur R.; BLANC, Gregory; ZHANG, Zonghua; DEBAR, Herve. IoT Devices Recognition Through Network Traffic Analysis. In: 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018. Dostupné z doi: 10.1109/bigdata.2018.8622243.
- [10] ABBAS, Nadeem; ASIM, Muhammad; TARIQ, Noshina; BAKER, Thar; ABBAS, Sohail. A Mechanism for Securing IoT-enabled Applications at the Fog Layer. *Journal of Sensor and Actuator Networks*. 2019, roc. 8, c. 1, s. 16. Dostupné z doi: 10.3390/jsan8010016.
- [11] SONG, Tianyi; LI, Ruinian; MEI, Bo; YU, Jiguo; XING, Xiaoshuang; CHENG, Xiuzhen. A Privacy Preserving Communication Protocol for IoT Applications in Smart Homes. *IEEE Internet of Things Journal*. 2017, roc. 4, c. 6, s. 1844–1852. Dostupné z doi: 10.1109/jiot.2017.2707489.
- [12] SANTOS, Matias R. P.; ANDRADE, Rossana M. C.; GOMES, Danielo G.; CALLADO, Arthur C. An efficient approach for device identification and traffic classification in IoT ecosystems. In: 2018 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2018. Dostupné z doi: 10.1109/iscc.2018.8538630.
- [13] RASHID, Md Mamunur; KAMRUZZAMAN, Joarder; HASSAN, Mohammad Mehdi; IMAM, Tasadduq; GORDON, Steven. Cyberattacks Detection in IoT-Based Smart City Applications Using Machine Learning Techniques. *International Journal of Environmental Research and Public Health*. 2020, roc. 17, c. 24, s. 9347. Dostupné z doi: 10.3390/ijerph17249347.
- [14] MUNOZ, Jordi Zayuelas i; SUAREZ-VARELA, Jose; BARLETROS, Pere. Detecting cryptocurrency miners with NetFlow/IPFIX network measurements. In: 2019 IEEE International Symposium on Measurements & Networking (M&N). IEEE, 2019. Dostupné z doi: 10.1109/iwmm.2019.8804995.
- [15] CHEN, Juan; TIAN, Zhihong; CUI, Xiang; YIN, Lihua; WANG, Xianzhi. Trust architecture and reputation evaluation for internet of things. *Journal of Ambient Intelligence and Humanized Computing*. 2018, roc. 10, c. 8, s. 3099–3107. Dostupné z doi: 10.1007/s12652-018-0887-z.
- [16] SHEN, Jian; GUI, Ziyuan; JI, Sai; SHEN, Jun; TAN, Haowen; TANG, Yi. Cloud-aided lightweight certificateless authentication protocol with anonymity for wireless body area networks. *Journal of Network and Computer Applications*. 2018, roc. 106, s. 117–123. Dostupné z doi: 10.1016/j.jnca.2018.01.003.
- [17] LI, Jin; ZHANG, Yinghui; CHEN, Xiaofeng; XIANG, Yang. Secure attribute-based data sharing for resource-limited users in cloud computing. *Computers & Security*. 2018, roc. 72, s. 1–12. Dostupné z doi: 10.1016/j.cose.2017.08.007.
- [18] MOURADIAN, Carla; NABOULSI, Diala; YANGUI, Sami; GLITHO, Roch H.; MORROW, Monique J.; POLAKOS, Paul A. A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges. *IEEE Communications Surveys & Tutorials*. 2018, roc. 20, c. 1, s. 416–464. Dostupné z doi: 10.1109/comst.2017.2771153.
- [19] AITKEN, Paul; CLAISE, Benoît; TRAMMELL, Brian. Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information [RFC 7011]. RFC Editor, 2013. Request for Comments, c. 7011. Dostupné z doi: 10.17487/RFC7011.
- [20] SCHMITT, Corinna; STILLER, Burkhard; TRAMMELL, Brian. TinyIPFIX for Smart Meters in Constrained Networks [RFC 8272]. RFC Editor, 2017. Request for Comments, c. 8272. Dostupné z doi: 10.17487/RFC8272.
- [21] PETIJA, R.; GLEVANAK, M.; KUCAN, M.; FECILAK, P.; JAKAB, F. Experimental Implementation of TinyIPFIX Protocol for Arduino and Raspberry Pi Platform. In: 2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, 2020. Dostupné z doi: 10.1109/iceta51985.2020.9379188.
- [22] PETIJA, R.; FECILAK, P.; JAKAB, F.; MICHALKO, M. Critical analysis of Communication Protocols to Support the Quality of Services in IoT-based Infrastructures. In: 2019 17th International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, 2019. Dostupné z doi: 10.1109/iceta48886.2019.9039989.

Metódy hlbokého učenia v analytických úlohách a rozhodovaní

Viera Maslej Krešňáková

Katedra kybernetiky a umelej inteligencie
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Email: viera.maslej.kresnakova@tuke.sk

Peter Butka

Katedra kybernetiky a umelej inteligencie
Fakulta elektrotechniky a informatiky
Technická univerzita v Košiciach
Email: peter.butka@tuke.sk

Abstrakt—Tento príspevok prezentuje výsledky dizertačnej práce, v ktorej sme sa zamerali na dve hlavné oblasti a to podporu fyzikálneho výskumu a na spracovanie prirodzeného jazyka. Prinášame niekoľko výskumných riešení na báze metód hlbokého učenia v oblasti astrofyziky a fyziky atmosféry. Nakoľko na robustnosť modelov vplyva najmä variabilita dát, predstavené práce poskytujú niekoľko metód na zvýšenie robustnosti učenia. V druhej časti sa sústreďujeme na prepojenie technického a spoločenského aspektu. Nedávna hrozba pandémie ešte viac zdôraznila potrebu včasného odhaľovania antisociálneho správania sa v online priestore. Ako vhodný nástroj pre riešenie tejto úlohy sa ukázalo práve použitie hlbokých neurónových sietí. V rámci tejto témy predstavujeme niekoľko prác zaoberajúcich sa antisociálnym správaním, nie len v anglických ale aj v slovenských online komunitách.

Abstract— This paper presents the dissertation results, in which we focused on two main areas: the support of the scientific research processes in physics and natural language processing. We describe several research solutions based on deep learning methods in astrophysics and physics atmospheric physics. As the robustness of the models is mainly influenced by the variability of the data, the presented works provide several approaches to increase the robustness of learning. In the second part, we focus on the connection of technical and social aspects. The recent threat of a pandemic has further highlighted the need for early detection of antisocial behavior online. The use of deep neural networks has proven to be a suitable tool for solving this problem. Within this topic, we present several works dealing with antisocial behavior, not only in English but also in Slovak online communities.

I. ÚVOD

Metódy hlbokého učenia predstavujú oblasť úzko spätú s dátovou analytikou. Jednotlivé domény zhromažďujú obrovské množstvo dát, ktoré môžu obsahovať užitočné informácie o problémoch, ktoré je potrebné analyzovať. Algoritmy hlbokého učenia extrahujú vysoko-komplexné abstrakcie ako reprezentácie dát, prostredníctvom hierarchického procesu učenia, vďaka čomu majú široké uplatnenie v rôznych analytických úlohách. Veľmi dobre výsledky hlbokých neurónových sietí a veľké množstvo pribúdajúcich dát, dostávajú metódy hlbokého učenia do popredia v aplikáciách podporujúcich rozhodovanie.

V tomto príspevku sa venujeme metódam hlbokého učenia a ich aplikácii v rôznych analytických úlohách a rozhodovaní, s cieľom rozšíriť súčasný súbor poznatkov. Motiváciou je

vysoká úspešnosť a široké využitie metód hlbokého učenia v praxi. Zatiaľ čo v prvej časti prepájame technický aspekt s podporou výskumu v oblasti fyzikálnych vied, v druhej časti sa sústreďujeme na analýzu technického a spoločenského aspektu online priestoru. V časti podpory fyzikálneho výskumu sa venujeme najmä témam z oblasti astrofyziky a fyziky atmosféry, kde aplikujeme metódy hlbokého učenia na klasifikáciu, segmentáciu či detekciu objektov. Výsledné modely a navrhnuté postupy poskytujú fyzikom lepšie nástroje pre ďalší vedecký výskum. V rámci podpory spoločenského aspektu sme sa zamerali na odhaľovanie antisociálneho správania sa v online komunitách. Zatiaľ čo online priestor je miestom, kde používatelia čoraz viac komunikujú, nedávna hrozba infekcie COVID-19 ešte viac zdôraznila potrebu včasného, automatického odhaľovania antisociálneho správania. Ukázalo sa, že práve metódy založené na hlbokom učení sú vhodným prostriedkom na riešenie tohto problému.

II. VYUŽITIE METÓD HLBOKÉHO UČENIA NA RIEŠENIE ÚLOH NA PODPORU FYZIKÁLNEHO VÝSKUMU

V tejto kapitole predstavíme štyri publikácie ktoré vznikli v spolupráci s expertmi v oblasti rádioastronómie, kozmickej fyziky, astrofyziky a fyziky atmosféry. V každej práci sme riešili niekoľko problémov v súvislosti s dátami a ich prípravou, ktoré si priblížime nižšie. Túto kapitolu by sme zároveň mohli členiť podľa základného typu riešenej úlohy. Z nášho pohľadu, primárne postavenom na problémoch počítačového videnia, ide o tri úlohy a to klasifikácia, segmentácia a detekcia.

A. Klasifikácia

1) *Klasifikácia rádiových galaxií*: V spolupráci s Khadijou El Bouchefry z rádioastronomického observatória v Južnej Afrike (SARAO - South African Radio Astronomy Observatory) sme realizovali návrh výpočtového modelu na báze konvulčných neurónových sietí pre riešenie úlohy klasifikácie rádiových galaxií do ich jednotlivých typov. Nakoľko manuálne anotovanie dát je časovo veľmi náročné, a v tak špecifickej doméne ako je rádioastronómia často vyžaduje oko experta, pracovali sme s malým datasetom. Konkrétne sme mali k dispozícii celkovo 526 označených snímok z katalógu Faint Images of the Radio Sky at Twenty Centimeters (FIRST). Tieto snímky obsahujú galaxie 4 tried: Fanaroff-Riley I (FRI),

Tabuľka I
TABUĽKA ZOBRAZUJE PRIEMERNÉ HODNOTY EXPERIMENTOV
NATRÉNOVANÝCH NA ROZŠÍRENOM SÚBORE ÚDAJOV

Type	Prenosť	Návratnosť	F1 skóre	Počet
Test set 1 – úspešnosť: 0.95				
BENT	0.96	0.97	0.97	21
COMPT	0.98	1.00	0.99	17
FRI	0.93	0.94	0.93	25
FRII	0.96	0.94	0.95	44
avg	0.96	0.96	0.96	total: 107
Test set 2 – úspešnosť: 0.99				
BENT	0.99	0.99	0.99	84
COMPT	1.00	1.00	1.00	67
FRI	0.99	0.99	0.99	100
FRII	0.99	0.99	0.99	172
avg	0.99	0.99	0.99	total: 423

Fanaroff-Riley II (FRII), bent-tailed (BENT) a kompaktné rádiové galaxie (COMPT). Prvotné experimenty s jednoduchou architektúrou sme odprezentovali na konferencii WIKT 2019 [1]. Rozšírený návrh s novým modelom a použitými augmentáciami bol úspešne realizovaný a publikovaný v [2] pod názvom *Morphological classification of compact and extended radio galaxies using convolutional neural networks and data augmentation techniques*, a to v karentovanom časopise *Monthly Notices of Royal Astronomical Society*.

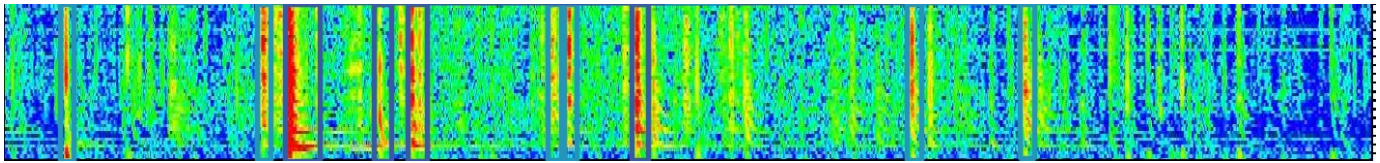
V tomto príspevku sme predstavili klasifikátor založený na konvolučných neurónových sieťach, ktoré môžu automaticky klasifikovať rádiové galaxie do štyroch morfológických tried. Náš klasifikátor, založený na kombinácii troch blokov konvulčnej neurónovej siete, za ktorými nasledovali dve plne prepojené vrstvy na konečnú klasifikáciu, priniesol výsledky porovnateľné s najmodernejšími klasifikátormi v doposiaľ publikovaných prácach. Jedným z podstatných aspektov bolo rozšírenie vstupných údajov augmentačnými technikami. Model sme natrénovali pomocou dátovej množiny rozšírenej o rotované obrázky a obrázky so zvýšeným jasom, vďaka ktorým bol model schopný lepšie sa naučiť morfológické vlastnosti, najmä v rozšírených rádiových galaxiách. Naučené váhy sme použili v druhom tréningu modelu, v ktorom sme rozšírili tréningový súbor údajov o ich vertikálne, horizontálne alebo vertikálne aj horizontálne preklopené kópie. Testovali sme dve nastavenia na základe existujúcich prác, a to rozdelenie pôvodných obrázkov pred augmentáciami (*test set 1*) a výber testovacej podmnožiny z pôvodných obrázkov s už aplikovanými rotáciami (*test set 2*). Obe nastavenia ukázali sľubné výsledky (viď Tabuľka I) a dosiahli lepšie výsledky v porovnaní s doposiaľ publikovanými prácami. Vďaka zmene veľkosti a relatívne malým vstupným obrázkom je architektúra aj výpočtovo efektívna a dá sa ľahko použiť na automatickú klasifikáciu.

2) *Klasifikácia svetelných kriviek zákrytových binárnych hviezd*: V posledných desaťročiach vzniklo samozrejme aj v optickej oblasti mnoho rôznych prehliadok oblohy (pozemské aj vesmírne), ktoré sa stali zdrojom značného množstva

nových dát. Tieto údaje obsahujú veľa informácií o binárnych hviezdnych systémoch, skrytých v ich svetelných krivkách, a to v podobe zákrytových premenných hviezd. V rámci tohto problému sme sa dostali k zaujímavej spolupráci s kolegami z odboru astrofyziky na Prírodovedeckej fakulte Univerzity Pavla Jozefa Šafárika v Košiciach. Naším cieľom bolo nájsť vhodný model pre klasifikáciu zákrytových premenných hviezd do definovaných tried na základe ich svetelných kriviek. Finálna práca bola úspešne publikovaná v [3] pod názvom *Automatic classification of eclipsing binary stars using deep learning methods*, a to v karentovanom časopise *Astronomy and Computing*. V tomto príspevku sa zameriavame sa na automatickú klasifikáciu zákrytových binárnych hviezdnych systémov pomocou metód hlbokého učenia. Náš klasifikátor poskytuje nástroj na kategorizáciu svetelných kriviek binárnych hviezd do dvoch tried: oddelených (detached) a dotkových (over-contacted). V tejto úlohe sme mali k dispozícii len 100 observačných svetelných kriviek, takže pre potreby zlepšenia klasifikácie bola variabilita dát dosiahnutá iba pomocou fyzikálneho modelu a ním generovaných umelých svetelných kriviek. Simulačný prístup je vo fyzike často používaný a táto práca preukázala, že dobre fyzikálne navrhnutý generátor je schopný poskytnúť potrebnú variabilitu podobne ako tradičné augmentačné techniky, ak nemáme k dispozícii dostatok observačných dát. Syntetické svetelné krivky, použité pre tréningovanie modelu, boli generované za pomoci nástroja ELISa [4] vyvinutého práve kolegami z UPJŠ. Na tréningovanie a validáciu modelu popri tréningu boli použité syntetické dáta, zatiaľ čo na testovanie sme použili observačné krivky. Najúspešnejší klasifikátor založený na obojsmernej LSTM sieti v kombinácii s konvulčnou neurónovou sieťou tréningovaný na umelo vytvorených svetelných krivkách dosiahol 98% presnosť pri klasifikácii observačných svetelných kriviek binárnych hviezd. V tomto prípade sme do triedy dotkových zaradili aj krivky z podtriedy tzv. polodotkových (semi-detached). Ak sme túto podtriedu nebrali do úvahy, klasifikátor dokázal zaradiť krivky do tried zo 100% presnosťou.

B. Detekcia

Ďalšia zaujímavá spolupráca nám umožnila skúmať možnosti detekcie objektov (udalostí) v spektrogramoch popisujúcich fenomén z oblasti fyziky atmosféry. V tomto prípade sa jednalo o spoluprácu s Ústavu experimentálnej fyziky Slovenskej akadémie vied v Košiciach (ÚEF SAV) a kolegov z Ústavu fyziky atmosféry Akadémie vied České republiky. V rámci aktuálneho stavu výskumu v tejto oblasti sú súčasné prístroje schopné zaznamenávať úder blesku na tisíce kilometrov pomocou meraní elektromagnetických impulzov vo frekvenčnom rozsahu nízkeho rádiového spektra až do 30 kHz. Následná analýza môže lokalizovať úder blesku a charakterizovať vlastnosti prostredia (atmosféry) medzi bleskom a detektorom. Techniky strojového učenia môžu výrazne zjednodušiť detekciu a analýzu týchto impulzov zaznamenaných vo forme spektrogramov. Jeden príklad takej dôležitej vlastnosti je určenie, či signál siaha pod 2 kHz alebo nie. Tento jav určuje, ako sa elektromagnetická vlna šíri od blesku



Obr. 1. Ukážka detekcie objektov v rámci spektrogramu

k detektoru. Môže sa šíriť iba v špecifickom režime s vlnovou dĺžkou asi 150 km, pričom výška medzi zemou a vrstvou D ionosféry je asi 90 km. V jednoduchosti môžeme povedať, že určením, či je detegovaná emisia sférickej hodnoty pod 2 kHz, môžeme určiť primárne rozloženie charakteristík šíriacej sa elektromagnetickej vlny. V našom prípade bol primárnym zdrojom spektrogramov analyzátor elektromagnetických vln ELMVAN-G [5], ktorý bol vyvinutý Ústavom fyziky atmosféry Akadémie vied Českej republiky pre satelitný projekt Resonance [6]. Každých 24 hodín merania sa získa takmer 2 600 snímok, čo je asi 950 000 snímok za 1 rok. Techniky strojového učenia by preto mohli poskytnúť užitočný nástroj na automatické zisťovanie bleskových udalostí v spektrogramoch s takým množstvom údajov. V prvom kroku bolo potrebné tieto data anotovať. Vytvorili sme preto crowdsourcingový anotačný projekt na platforme Zooniverse, vďaka ktorému sme získali 2 300 obrázkov, na ktorých bolo označených spolu 22 200 udalostí (3 200 tweekov a 19 000 sférikov).

V rámci spolupráce sme následne navrhli a realizovali spoľahlivú automatickú metódu na extrakciu požadovaných podrobností zo spektrogramov, ktoré môžu byť následne použité na ďalší výskum. Ide o jedinečný prístup, ktorý v tejto oblasti zatiaľ v podobnej forme nebol realizovaný. Výsledok a podrobný popis metódy, ktorý môže byť aj metodologickým príkladom pre realizáciu podobných úloh, bol spracovaný do podoby článku s názvom *Automatic detection of tweek atmospherics in radio spectrograms based on a deep learning approach*. Tento článok bol podaný do karentovaného časopisu *Earth and Space Science*. Predložená metóda je založená na hlbokom učení a následnom deterministickom prístupe. Vďaka detekčnej neurónovej sieti YOLOv5 dokážeme so zvolenou citlivosťou detegovať sférikky a tweeky zo spektrogramov (viď Obrázok 1). Pomocou deterministických výpočtov v ďalšej fáze určujeme presnú milisekundu výskytu udalosti a zároveň poskytujeme informácie o tom, či daná udalosť má emisie pod 2 kHz alebo nie. Extrahované detaily sú následne ukladané vo forme tabuľky na ďalšie štatistické vedecké analýzy. Prezentovaná metóda je veľmi rýchla a je vhodná aj na nasadenie v reálnom čase.

C. Segmentácia

Štruktúry slnečnej koróny sú hlavnými hnacími silami procesov vesmírneho počasia, ktoré môžu priamo alebo nepriamo ovplyvňovať Zem a často aj fungovanie dôležitých systémov (navigačné družice, telekomunikácie, infraštruktúry pre prenos elektrickej energie, atď.). Vďaka najnovším vesmírnym slnečným observatóriám so schopnosťami kontinuálneho získavania snímok s vysokým rozlíšením je možné

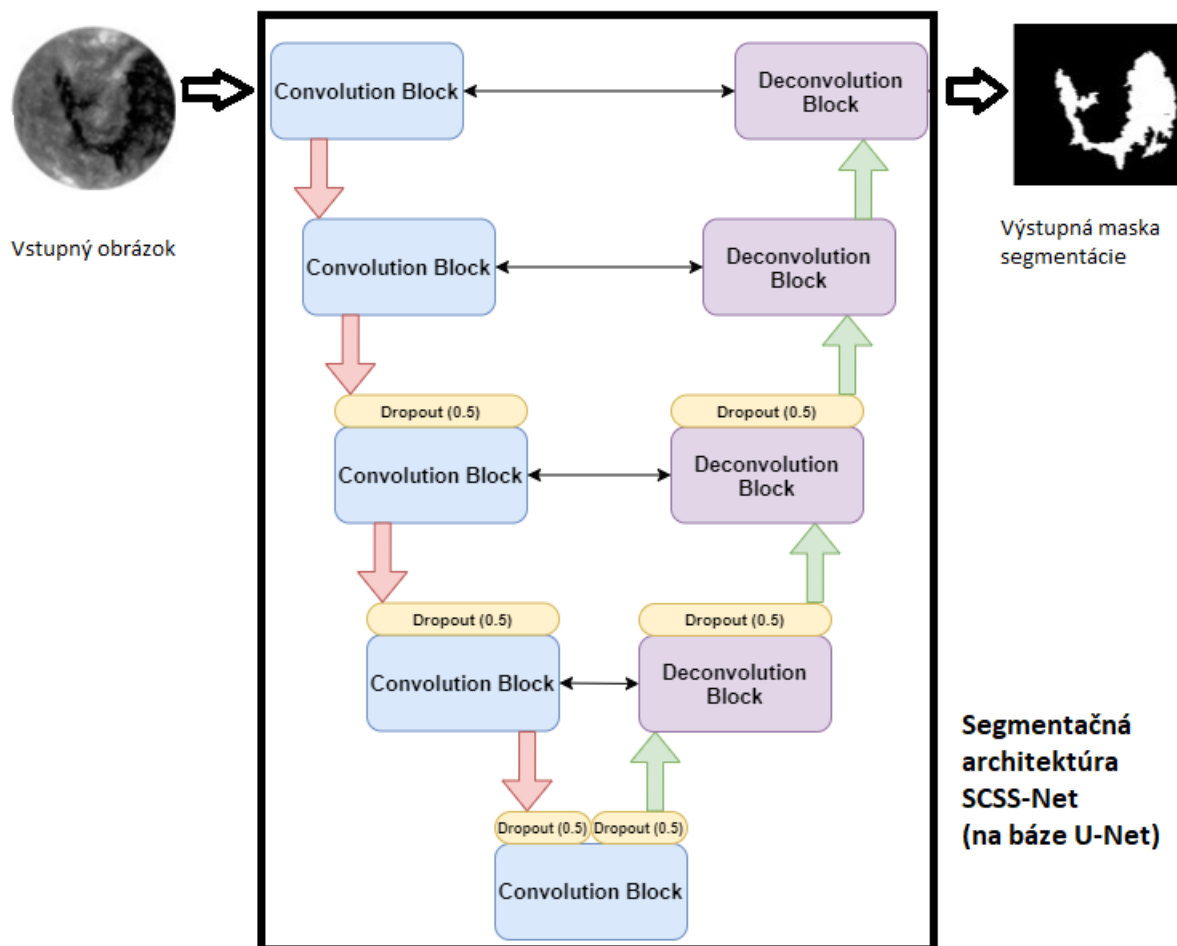
štruktúry v slnečnej koróne v priebehu rokov monitorovať s časovým rozlíšením minút. V práci [7] pod názvom *SCSS-Net: Solar Corona Structures Segmentation by Deep Learning*, ktorá bola publikovaná v karentovanom časopise *Monthly Notices of Royal Astronomical Society*, sme predstavili metódu automatickej segmentácie štruktúr slnečnej koróny pozorovaných v rámci elektromagnetického spektra v EUV (extrémne ultrafialové žiarenie). Ide o metódu založenú na hlbokom učení využívajúcom konvolučné neurónové siete pre segmentáciu vybraných štruktúr slnečnej koróny. Využili sme dostupné množiny anotovaných dát spolu s našim vlastným datasetom, ktorý sme získali na základe manuálnej anotácie cieľových štruktúr v platforme Zooniverse. Predstavený model *SCSS-Net* poskytuje segmentáciu koronálnych dier a aktívnych oblastí. Architektúra *SCSS-net*, ktorá je znázornená a popísaná na Obrázku 2 je inšpirovaná architektúrou U-Net [8], pôvodne vyvinutou na segmentáciu nádorov na mozgu. Predstavená metóda poskytuje univerzálny postup na identifikáciu štruktúr na slnečnej koróne¹.

Okrem návrhu modelu sme v uvedenom článku porovnávali viaceré existujúce anotačné datasety pre koronálne diery, pričom sme ukázali schopnosť nášho modelu prispôbiť sa rôznej kvalite anotácií a nájsť výstup lepšie pokrývajúci dané štruktúry. Niektoré anotačné sady totiž štruktúry do určitej miery podhodnocujú, iné práve naopak označia zbytočne väčšiu oblasť ako by bolo potrebné. Ako sa ukazuje, využitie viacerých anotácií a použitie augmentácií prispelo k schopnosti nášho modelu nájsť vo väčšine prípadov vhodný kompromis a v podstate tak automaticky potlačiť nepresnosti jednotlivých anotačných sád. Jedinečnosť tohto príspevku je aj v tom, že poskytuje prvýkrát štruktúrne jednotný model na báze hlbokého učenia priamo použitý nielen pre koronálne diery, ale súčasne aj pre segmentáciu aktívnych oblastí. Proces získavania anotácií, ich použitia pre tréning modelu a aplikácia modelu pre segmentáciu zároveň poskytujú určitý metodologický postup pre prípadné využitie podobného prístupu v segmentácii ďalších štruktúr v budúcnosti.

III. VYUŽITIE METÓD HLBOKÉHO UČENIA NA RIEŠENIE ÚLOH PRI SPRACOVANÍ TEXTU

Spracovanie prirodzeného jazyka sa zaoberá tvorbou výpočtových algoritmov na automatickú analýzu a reprezentáciu ľudského jazyka. Systémy založené na hlbokých neurónových sieťach prinášajú širokú škálu aplikácií, ako

¹Ukážka automatickej segmentácie koronálnych dier pomocou SCSS-Net: <https://www.youtube.com/watch?v=INGJu9yF19s>



Obr. 2. Architektúra modelu *SCSS-Net*. *SCSS-Net* je založená na princípe dekóder-dekóder. Príklad na tomto obrázku (vytvoreného na základe obrázku architektúry v [7]) predstavuje prípad segmentácie koronálnej diery. Kóder, ktorý predstavuje ľavú časť obrázku, sa skladá z piatich konvolučných blokov. Červené šípky označujú vzorkovacie vrstvy *max-pooling*. Dekóder je znázornený na pravej časti obrázku, obsahuje štyri de-konvolučné bloky. Základom de-konvolučných blokov sú dve transponované konvolučné vrstvy. Zelené šípky označujú vrstvy nad-vzorkovania. Čierne šípky označujú zrefazenie, ktoré spája príznakové mapy objektov na rovnakej úrovni. Táto operácia pomáha poskytnúť lokalizačné informácie z enkódera do dekódera. Výstupná klasifikačná vrstva pozostáva z konvolučnej vrstvy s aktivačnou funkciou sigmoid. Výstupom *SCSS-Net* je maska segmentácie.

vyhľadávacie nástroje, hlasových asistentov či schopnosť vykonávať komplexné úlohy súvisiace s prirodzeným jazykom, ako je strojový preklad a generovanie dialógov. V tejto kapitole predstavíme navrhnuté prístupy v oblasti detekcie antisociálneho správania a analýze emócií pri interakcii robot-človek.

A. Anti-sociálne správanie

V súčasnosti je celosvetová sieť Internet prostredím, v ktorom môžu jeho používatelia vytvárať a zdieľať informácie slobodne, takmer bez nejakých obmedzení. Drvivá väčšina používateľov Internetu sa v tomto prostredí správa zodpovedne, snaží sa ho udržiavať bezpečný a efektívny pre prácu alebo zábavu, no existuje skupina používateľov, ktorých správanie na Internete by sa dalo označiť ako anti-sociálne. Existuje viacero charakteristík anti-sociálneho správania sa, no vo všeobecnosti ho môžeme rozdeliť do dvoch hlavných skupín [9]. Prvá skupina zahŕňa šírenie dezinformácií, ktoré

môžu nadobúdať viaceré podoby. Patria tu napríklad hoaxy, falošné správy alebo recenzie. Druhá skupina sa týka hlavne reakcií používateľov, a to najmä v rámci manipulácie diskusií na Internete. Patrí tu kyberšikana a mnoho iných podobných foriem manipulácie. Obe skupiny anti-sociálneho správania predstavujú veľmi závažný problém, keďže ich následky môžu byť v skutočnom svete veľmi vážne. Ako príklad môžeme uviesť šírenie falošných správ, ktoré v súčasnosti je veľmi rozšírené. Falošné správy môžu ovplyvňovať názory ich čitateľov, a tak môžu mať ďalekosiahle dôsledky na ich rozhodovanie, napríklad aj počas volieb alebo referenda. Ako typický príklad môžu slúžiť prezidentské voľby v roku 2016 v Spojených štátoch amerických, počas ktorých sa skrz sociálnu sieť Facebook šírilo viac falošných správ ako tých skutočných [10]. Tento jav nadobudne ešte väčší impakt, ak si uvedomíme, že väčšina používateľov, na ktorých bol obsah cielený, považovala Facebook za primárny zdroj informácií

[11].

Z týchto dôvodov silnie potreba efektívne detegovať a eliminovať rôzne formy anti-sociálneho správania na Internete. V súčasnosti stále existuje tzv. manuálny spôsob odhaľovania a vysporiadania sa s takýmto správaním. Manuálne techniky zahŕňajú rôznych operátorov alebo moderátorov, ktorí sú zodpovední za vyhľadávanie, odhalenie a identifikáciu anti-sociálneho správania v daných on-line komunitách. Prirodzene je tento spôsob neefektívny. Jednak z časového hľadiska, kedy pri zvyšujúcom sa objeme informácií je ich spracovanie náročnejšie na kapacity moderátorov, a na druhej strane, takýto spôsob môže trpieť subjektivitou jednotlivých moderátorov [12]. Preto je dôležité zamerať sa na pokročilé metódy detekcie anti-sociálneho správania, najmä tie automatizovateľné. Tieto techniky sú postavené najmä na kombinácií rôznych disciplín, od dátovej analytiky, cez strojové učenie až po umelú inteligenciu.

V knižnej kapitole [13] pod názvom *Addressing False Information and Abusive Language in Digital Space Using Intelligent Approaches*, ktorá bola publikovaná v rámci knihy vydananej vo vydavateľstve Springer ako zborník pôvodne plánovaný ku konferencii *World Symposium on Digital Intelligence for Systems and Machines*, bol predstavený súhrn existujúcich výskumných prác, výziev a otvorených problémov. Predstavené prístupy sú založené na inteligentných metódach založených na znalostiach, hlavne na strojovom učení, spracovaní prirodzeného jazyka a poloautomatických prístupoch. Okrem detekcie nežiadúceho obsahu uvádzame aj menej preštudované prístupy k identifikácii používateľov, ktorí prispievajú alebo šíria takýto obsah.

V práci [14] pod názvom *Comparison of Deep Learning Models and Various Text Pre-Processing Techniques for the Toxic Comments Classification*, publikovanej v karentovanom časopise *Applied Sciences* je popísané podrobné porovnanie rôznych modelov hlbokého učenia používaných na identifikáciu toxických komentárov v internetových diskusiách, ako jednej z foriem antisociálneho správania objavujúceho sa na Internete. Naším hlavným cieľom bolo preskúmať vplyv predspracovania dát na výkon rôznych typov modelov. Pracovali sme s predpokladom, že použitie tradičných metód predspracovania môže viesť k strate charakteristických vlastností špecifických pre toxický obsah a porovnali sme niekoľko štandardných modelov neurónových sietí a populárnych transformers jazykových modelov. Naším cieľom bolo analyzovať vplyv rôznych techník prípravy dát a vektorových reprezentácií slov, a to vrátane štandardného TF-IDF váhovania, vopred natrénovaných *word embeddings* a taktiež preskúmať v súčasnosti populárne jazykové modely z rodiny *HuggingFace*. Experimenty sa uskutočňovali na súbore údajov zo súťaže Kaggle Toxic Comment Classification a najvýkonnejší model bol porovnaný s podobnými prístupmi pomocou štandardných metrick.

Schéma poronávaných experimentov je nasledovná:

Zatiaľ čo digitálny priestor je miestom, kde používatelia čoraz viac komunikujú, nedávna hrozba infekcie COVID-19 ešte viac zdôraznila potrebu efektívneho a dobre organizo-

vaného online prostredia. Nežiaduce správanie používateľov a obsah generovaný používateľmi v online prostredí (väčšinou na sociálnych sieťach) môžu mať rôzne podoby, asi najškodlivejšie je vytváranie a šírenie falošných správ.

Na konferencii CINTI 2019 bola prezentovaná práca [15] pod názvom *Deep learning methods for Fake News detection*. Cieľom tejto práce bola detekcia falošných správ už skôr spomínaných prezidentských volieb v US z roku 2016. Tieto články často obsahujú tzv. *clickbait* nadpis, teda akýsi senzačný text, ktorý by mal prilákať používateľov kliknúť na daný článok. Práve preto sme sa rozhodli porovnať úspešnosti modelov neurónových sietí ktoré boli trénované na celých textoch a len na nadpisoch týchto falošných správ.

V článku *Use of data augmentation techniques in detection of antisocial behavior using deep learning method*, sme sa zamerali na augmentácie dát pomocou techník EDA pri spracovaní textu v konkrétnej doméne, a to detekcia antisociálneho správania v online priestore. Príspevok sa zameriava na dva problémy, a to na odhaľovanie toxických komentárov a detekciu falošných správ. Ukázalo sa, že toxické komentáre obsahujú množstvo slangových slov, vulgarizmov, nespisovných slov či skratiek, ktoré nie sú v synonymickom slovníku, teda ich nie je možné nahradiť pri dvoch EDA technikách. Rovnako dôsledkom náhodne vymazania sme často prišli o slová, ktoré sú charakteristické pre daný typ toxicity. Preto tieto augmentačné techniky nie sú vhodným nástrojom na riešenie rozširovania dátovej množiny za účelom zlepšenia klasifikácie toxických komentárov. Na druhej strane pri odhaľovaní falošných správ, používanie EDA zvýšilo priemerné makro F1 hodnoty o 19 percent a návratnosť klasifikácie o 27 percent. Tento článok bol podaný do indexovaného časopisu *Journal of Information and Organizational Sciences*.

V ďalšej časti sme sa zamerali na odhaľovanie anti-sociálneho správania v slovenskom online priestore. V prvom rade sme sa zamerali na zber dát. Na získanie údajov z rôznych slovenských online webov sme použili platformu MonAnt [16]. Túto platformu sme použili na zhromažďovanie spravodajských článkov z relevantných webov, ako aj z nespoľahlivých zdrojov, ktoré často publikovali konšpiračný obsah. Konšpiračné weby sme vybrali na základe databázy www.konspiratori.sk. Databáza je výsledkom práce expertov, ktorí sa zameriavajú na monitorovanie falošných správ v rôznych médiách, pričom obsahuje rebríček mnohých online spravodajských portálov. Tento rebríček je usporiadaný na základe skóre, ktoré predstavuje pravdepodobnosť uverejnenia zavádzajúcich informácií a falošných správ na daných stránkach. Takéto označovanie berie do úvahy iba dôveryhodnosť celého zdroja (napr. webovej stránky alebo novín), nie však dôveryhodnosť jednotlivých správ. Takýto prístup môže viesť k nesprávnemu označeniu jednotlivých článkov, pretože mnoho konšpiračných webových stránok preberá aj relevantné správy. Dospeli sme k záveru, že najlepší spôsob, ako správne priradiť relevantnosť článkom, je manuálne znova označiť obsah. Na konferencii ICETA 2020 sme v práci [17] pod názvom *Annotated dataset for the fake news classification in Slovak language* predstavili takýto

anotačný projekt, ktorý fungoval na princípe crowdsourcingu.

V rámci analýzy falošných správ v slovenskom jazyku sme sa v prvej časti zamerali na detekciu fake news v politických a ekonomických správach. Na konferencii SAMI 2021 v práci [18] pod názvom *Fake news detection in Slovak language using deep learning techniques*, sme sa zamerali na automatickú detekciu práve takýchto falošných správ.

Nakoľko naše životy ovplyvnila pandémia COVID-19 a do popredia sa dostalo množstvo konšpiračných a falošných správ práve na túto tému, rozhodli sme sa zamerať aj na odhaľovanie týchto falošných správ. V práci *Covid-19 Fake News detection in Slovak online space using Deep Learning methods*, ktorá je podaná do impaktovaného časopisu *Acta Polytechnica Hungarica*, sme za pomoci platformy Monant zhromaždili články týkajúce sa pandémie COVID-19 za účelom automatickej detekcie falošných správ v slovenskom jazyku. Dáta sme manuálne anotovali a predstavili niekoľko modelov založených na hlbokom učení. Vďaka kombinácií obojsmernej LSTM siete a konvolučnej vrstve sme získali priemernú hodnotu macro F1 skóre 94% na nezávislej testovacej množine.

B. Analýza sentimentu – emócií

Okrem problematiky antisociálneho správania sme sa v rámci jednej z aplikácií spracovania textov zaoberali aj analýzou sentimentu v kontexte interakcie človek-robot. Výsledky tohto výskumu je možné nájsť v článku [19] pod názvom *Emotion Analysis in Human–Robot Interaction*, ktorý je publikovaný v karentovanom časopise *Electronics*. Do analýzy sentimentu radíme aj analýzu emócií, ktorá skúma aké emócie sú reprezentované v textoch. Cieľom analýzy emócií v interakcii človek-robot je vyhodnotiť emocionálny stav človeka a na tomto základe rozhodnúť, ako by mal robot prispôbiť svoje správanie k človeku. Existuje niekoľko prístupov a algoritmov na detekciu emócií v textových dátach. Mojm prínosom v tejto práci je využitie neurónových sietí. V tomto experimente sme trénovali dopredný klasifikátor neurónových sietí, aby sme porovnali výkonnosť prístupu neurónových sietí so štandardnými metódami strojového učenia používanými v predchádzajúcich experimentoch. Model doprednej neurónovej siete získal o niečo lepší výkon (pri zohľadnení priemerných metrík) v porovnaní so štandardnými modelmi strojového učenia. Na druhej strane, nedostatok dát na trénovanie spôsobil to, že pokročilejšie prístupy hlbokého učenia (ako sú modely CNN alebo LSTM) alebo pokročilejšie jazykové modely (napr. BERT) sa ukázali nevhodné na riešenie danej úlohy.

IV. ZÁVER

Predložená práca prináša niekoľko na pohľad rozdielných výskumom, ktoré však spája oblasť úzko spätá s dátovou analytikou a to hlboké učenie. Hlavnou motiváciou pri štúdiu metód hlbokého učenia sa stala multidisciplinárnosť a rozmanitosť použitia neurónových sietí. Ako bolo definované v úvode v každej z úloh na podporu fyzikálneho výskumu sme riešili špecifický problém vzhľadom k dátam. Zároveň sme tieto práce rozdelili podľa typu riešenej úlohy.

Pri morfolologickej klasifikácii rádiových galaxií sme prezentovali výskum založený na vhodnom parametrizovaní metód augmentácie dát pri trénovaní neurónových sietí. Na klasifikačnú úlohu sme sa zamerali aj z pohľadu spracovania sekvenčných dát. V tomto prípade sme predstavili špecifický výskum založený na trénovaní modelov na umelo vytvorených sekvenciách, ktoré reprezentujú svetelné krivky binárnych hviezd. Tieto krivky boli generované na základe fyzikálnych poznatkov. Neurónová sieť naučená na týchto syntetických dátach dokázala klasifikovať observačné binárne hviezdy so 100 percentnou úspešnosťou.

Metódy hlbokého učenia sme využili aj v úlohe zameranej na detekciu objekt v úlohe zameranej na podporu atmosférického výskumu. V tomto prípade sme disponovali s veľkým množstvom neoznačených dát, ktoré sme anotovali vďaka crowdsourcingovému projektu. Následne pomocou najmodernejšieho algoritmu na detekciu objektov YOLOv5 a následného deterministického prístupu, sme dokázali extrahovať informácie zo spektrogramov potrebné na následný fyzikálny výskum.

V poslednej časti, ktorá uzatvára prvú tézu, ktorá sa venuje podpore fyzikálneho výskumu, sme predstavili SCSS-Net, univerzálnu metodiku na segmentáciu štruktúr na slnečnej koróne. Táto metóda predstavuje robustný model vďaka použitým augmentáciám dát a regularizačným technikám v procese učenia.

Druhá časť prepája technický a sociálny aspekt. Neurónové siete zaznamenali veľký úspech aj v spracovaní prirodzeného jazyka. Zamerali sme sa na populárnu tému, a to odhaľovanie antisociálneho správania sa v online priestore. Spracovávali sme textové dokumenty nie len v anglickom ale aj v slovenskom jazyku. Opäť sme riešili niekoľko problémov vzhľadom k dátam. Porovnávali sme augmentačné techniky v dvoch špecifických úlohách a to odhaľovanie falošných správ a klasifikácia toxických komentárov. Zároveň prinášame prehľad vplyvu predspracovania dát a porovnanie rôznych metód na vytváranie slovných vektorov pri trénovaní nie len štandardných modelov neurónových sietí ale aj populárnych transformer jazykových modelov.

Vďaka manuálnym anotáciám dokumentov v slovenskom jazyku sme dokázali natrénovať modely, ktoré sú schopné detegovať politické a ekonomické falošné správy. Po príchode pandémie Covid-19 sme sa zamerali na detekciu dezinformácií spojených s touto témou.

V tejto práci sme sa venovali metódam hlbokého učenia v rôznych analytických úlohách a rozhodovaní, s cieľom rozšíriť súčasný rozbor poznatkov. Vďaka multidisciplinárnosti, ktorú ponúka táto oblasť, sme dokázali preniesť získané vedomosti do viacerých domén. V ďalšej práci, aj vďaka získanému európskemu projektu ESA, budeme aplikovať nadobudnuté vedomosti na ďalší výskum v oblasti astrofyziky. Na základe spolupráce s doménovými expertmi z Ústavu experimentálnej fyziky SAV, Ústavom fyziky atmosféry Českej akadémie vied, Ústavom fyzikálnych vied UJPŠ, Fakulty matematiky fyziky a informatiky UK, vedcami zo SARAO v Južnej Afrike, či inštitútom seizmológie University of Helsinki vo Fínsku, dokážeme prepojiť nami ponúkaný technický aspekt

zo znalosťami doménových expertov a podporiť tak fyzikálny výskumu v špecifických doménach astronómie či geografie. Z oblasti informatiky prinášame do týchto domén najmodernejšie automatické metódy založené na hlbokom učení. Poukazujeme tak na výhody využitia strojového učenia v týchto doménach.

PODĎAKOVANIE

Táto práca bola podporená APVV grantom – APVV-16-0213 a VEGA 1/0685/21.

LITERATÚRA

- [1] V. M. Krešňáková, K. H. Le Thanh, E. Pizur, and P. Butka, “Klasifikácia rádiových galaxií metódami hlbokého učenia,” *DATA A ZNALOSTI & WIKT 2019*, p. 35, 2019.
- [2] V. Maslej-Krešňáková, K. El Bouchefry, and P. Butka, “Morphological classification of compact and extended radio galaxies using convolutional neural networks and data augmentation techniques,” *Monthly Notices of the Royal Astronomical Society*, vol. 505, no. 1, pp. 1464–1475, 2021.
- [3] M. Čokina, V. Maslej-Krešňáková, P. Butka, and Š. Parimucha, “Automatic classification of eclipsing binary stars using deep learning methods,” *Astronomy and Computing*, p. 100488, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2213133721000421>
- [4] M. Fedurco, M. Čokina, and Š. Parimucha, “Light curve modelling of close eclipsing binaries,” *Contributions of the Astronomical Observatory Skalnaté Pleso*, vol. 50, no. 2, pp. 563–564, 2020.
- [5] O. Santolík and I. Kolmašová, “Unusual electromagnetic signatures of european north atlantic winter thunderstorms,” *Scientific reports*, vol. 7, no. 1, pp. 1–9, 2017.
- [6] M. Mogilevsky, L. Zelenyi, A. Demekhov, A. Petrukovich, and D. Shklyar, “Resonance project for studies of wave-particle interactions in the inner magnetosphere,” *Am. Geophys. Union*, vol. 199, pp. 117–126, 2012.
- [7] v. Mackovjak, M. Harman, V. Maslej-Krešňáková, and P. Butka, “Scsnet: Solar corona structures segmentation by deep learning,” *Monthly Notices of the Royal Astronomical Society*, 2021.
- [8] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [9] S. Kumar, J. Cheng, and J. Leskovec, “Antisocial Behavior on the Web: Characterization and Detection,” *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 947–950, 2017.
- [10] C. Budak, “What happened? The Spread of Fake News Publisher Content During the 2016 U.S. Presidential Election,” 2019, pp. 139–150.
- [11] A. Mitchell, J. Gottfried, and K. E. Matsa, “Facebook Top Source for Political News Among Millennials — Pew Research Center,” 2015.
- [12] I. Kayes, N. Kourtellis, D. Quercia, A. Iamnitchi, and F. Bonchi, “The Social World of Content Abusers in Community Question Answering,” 2016, pp. 570–580.
- [13] K. Machova, I. Srba, M. Sarnovský, J. Paralič, V. M. Kresnakova, A. Hreckova, M. Kompan, M. Simko, R. Blaho, D. Chuda *et al.*, “Addressing false information and abusive language in digital space using intelligent approaches,” in *World Symposium on Digital Intelligence for Systems and Machines*. Springer, 2020, pp. 3–32.
- [14] V. Maslej-Krešňáková, M. Sarnovský, P. Butka, and K. Machová, “Comparison of deep learning models and various text pre-processing techniques for the toxic comments classification,” *Applied Sciences*, vol. 10, no. 23, p. 8631, 2020.
- [15] V. M. Krešňáková, M. Sarnovský, and P. Butka, “Deep learning methods for fake news detection,” in *2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo)*. IEEE, 2019, pp. 000 143–000 148.
- [16] I. Srba, R. Móro, J. Simko, J. Sevech, D. Chudá, P. Návrát, and M. Bieleliková, “Monant : Universal and Extensible Platform for Monitoring , Detection and Mitigation of Antisocial Behaviour,” 2019.
- [17] M. Sarnovský, V. Maslej-Krešňáková, and N. Hrabovská, “Annotated dataset for the fake news classification in slovak language,” in *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 2020, pp. 574–579.
- [18] K. Ivancová, M. Sarnovský, and V. Maslej-Krešňáková, “Fake news detection in slovak language using deep learning techniques,” in *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMII)*. IEEE, 2021, pp. 000 255–000 260.
- [19] M. Szabóová, M. Sarnovský, V. Maslej Krešňáková, and K. Machová, “Emotion analysis in human-robot interaction,” *Electronics*, vol. 9, no. 11, p. 1761, 2020.

QuoVadis Research @ FEI

Vydavateľ: Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky

Sídlo: Letná 9, 042 00 Košice

Elektronická adresa: <http://quovadis.fei.tuke.sk>

Periodicita: 2-krát ročne

Založenie časopisu: September 2018.

Jazyk: V časopise sú články v anglickom a slovenskom jazyku.

ISSN: 2585-9587