Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

QuoVadis Research @ FEI

FEI

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 metodík, použitie existujúcej metodológie na nový problém, čím prispievajú k rozširovaniu základných vedeckých poznatkov v danej vednej disciplíne.

Obsah

On sharp oscillation results for second-order half-linear differential equations with deviating arguments

Irena Jadlovská

Department of Mathematics and Theoretical Informatics
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 040 00, Košice, Slovakia
Email: irena.jadlovska@tuke.sk

Blanka Baculíková

Department of Mathematics and Theoretical Informatics
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 040 00, Košice, Slovakia
Email: blanka.baculikova@tuke.sk

Abstract—In the paper, we deal with the oscillation problem for certain second-order functional differential equations. These equations enjoy the longest history, from both theoretical or applicational perspective. Our results concern so-called half-linear functional differential equations, which form an imaginary borderline between linear and nonlinear equations. The most valuable contribution of this work consists in offering unimprovable results for a large class of equations investigated in hundreds of works for more than 40 years. The basis of newly employed techniques mostly lies in establishing and subsequent application of iterative estimates relating a nonoscillatory solution with its derivatives, which, in the limit case, lead to the necessary and sufficient conditions for the oscillation of these equations. The resulting criteria have been verified on suitable comparison equations, using the computational methods.

Abstract—Článok sa venuje problému vyšetrovania oscilatorických vlastností riešení určitých funkcionálnych diferenciálnych rovníc druhého rádu. Tieto rovnice majú najdlhšiu históriu, z teoretického ako aj aplikačného hľadiska. Naše výsledky sa týkajú tzv. pololineárnych diferenciálnych rovníc, ktoré vytvárajú myslenú hranicu medzi lineárnymi a nelineárnymi rovnicami. Najväčší prínos tejto práce leží v poskytnutí nevylepšiteľných výsledkov pre veľkú triedu rovníc vyšetrovanú v stovkách prác v priebehu posledných viac než 40 rokov. Základom novorozvinutých techník je vytvorenie a následná aplikácia iteratívnych odhadov spájajúcich neoscilatorické riešenie s jeho deriváciami, ktoré v limitnom prípade vedie na nutnú a postačujúcu podmienku pre osciláciu týchto rovníc. Výsledné kritériá boli overené na vhodných porovnávacích typoch rovníc, s využitím výpočtových prostriedkov.

I. Asymptotic behavior and oscillatory phenomena

Many types of system exhibit oscillation as their fundamental behavior. Much physiology operates in an oscillatory manner including hormone oscillations, oscillatory gene expressions, and oscillation in muscles; oscillations are also observed in ecosystem populations, biochemistry, there are thermodynamic oscillations, oscillations in economic markets and so on [1], [2], [3]. The existence of the oscillatory phenomena is evident in each science eventually resulting in the acceptance of oscillation as a fundamental mode of behavior.

In 1836, when investigating the thermal conductivity, Jacques Charles Sturm posed oscillation problems of the second-order differential equation $x''(t) + q(t)x(t) = 0$ in the real domain, which initiated a whole new direction in the qualitative theory of differential equations. His innovative idea was to deduce oscillatory properties of unknown solutions of a given differential equation from known ones of an another, see [4]. Since Sturm's pioneering work, the oscillation theory, as an important branch of the qualitative theory, has become an indispensable mathematical tool for many sciences and high technologies.

The essence of the oscillation theory lies in establishing conditions for the existence of oscillatory (nonoscillatory) solutions and/or the convergence to zero, in studying the laws of distribution of the zeros, in obtaining the lower bounds for the distance between consecutive zeros, in studying the number of zeros in a given interval, as well as in examining the relationship between the oscillatory properties of solutions and corresponding oscillatory processes in systems of diversified physical nature.

In the paper, we deal with the oscillation problem for certain second-order functional differential equations. These equations enjoy the longest history, from both theoretical or applicational perspective. Our results concern so-called half-linear functional differential equations, which form an imaginary borderline between linear and nonlinear equations. Its most valuable contribution lies in offering unimprovable results for a large class of equations investigated in hundreds of works for more than 40 years.

The structure of the paper consists of the following sections:

1) Classification of studied equations consisting of definitions of the equations along with the assumptions posed on coefficients and deviating arguments.
2) The current-state-of-the art, closely related to the subject and subsequent author's contribution. Herein, the reader is familiarized with some of the recent results, which

served as a primary motivation for doing our research on this subject. Most importantly, various occurring open problems related are outlined.

3) Problem statement, where a clear formulation of the open problems is given.
4) Methods and tools which we plan to apply are briefly described. Most importantly, discussion about the contribution to the subject is made, where the main results are formulated.

II. Formulation

Consider the second-order half-linear delay differential equation (DDE)

$$\left(r(t)\left(y'(t)\right)^{\alpha}\right)' + q(t)y^{\alpha}(\tau(t)) = 0, \quad t \geq t_0, \qquad (1)$$

and the dual advanced differential equation (ADE)

$$\left(r(t)\left(y'(t)\right)^{\alpha}\right)' + q(t)y^{\alpha}(\sigma(t)) = 0, \quad t \geq t_0, \qquad (2)$$

where $\alpha > 0$ is a quotient of odd positive integers, $r, q \in \mathcal{C}([t_0, \infty), (0, \infty))$, and $\tau, \sigma \in \mathcal{C}^1([t_0, \infty), \mathbb{R})$ are such that

$$\tau(t) \leq t \quad \text{for} \quad t \geq t_0, \quad \tau'(t) \geq 0, \quad \text{and} \quad \lim_{t \to \infty} \tau(t) = \infty;$$

and

$$\sigma(t) \geq t \quad \text{for} \quad t \geq t_0 \quad \text{and} \quad \sigma'(t) \geq 0,$$

respectively. Set $\tau^{-1}(t) = \inf\{\tau(s) : s \geq t\}$. Under the solution of equation (1) we mean a function $y \in \mathcal{C}([t_0, \infty), \mathbb{R})$, which has the property $r(y')^{\alpha} \in \mathcal{C}^1([\tau^{-1}(t_0), \infty), \mathbb{R})$ and satisfies (1) on $[\tau^{-1}(t_0), \infty)$. Analogous definition is adopted for (2). We consider only those solutions of (1) (or (2)) which exist on some half-line $[\tau^{-1}(t_0), \infty)$ and satisfy the condition $\sup\{|x(t)| : T \leq t < \infty\} > 0$ for any $T \geq \tau^{-1}(t_0)$.

Following the classical terminology for linear disconjugate equations due to Polya and Trench [5], we shall say that the equation (1) is in *canonical* form if

$$R(t, t_0) := \int_{t_0}^t r^{-1/\alpha}(s)\mathrm{d}s \to \infty \quad \text{as} \quad t \to \infty \qquad (3)$$

holds, which obviously allows one to take $r(t) = 1$. Conversely, we say that (1) is in *noncanonical* form if

$$\pi(t_0) := \int_{t_0}^{\infty} r^{-1/\alpha}(s)\mathrm{d}s < \infty \qquad (4)$$

holds. For simplicity, we omit the second argument in R when $t_x = t_0$, i.e., we use $R(t)$ instead of $R(t, t_0)$.

The oscillatory behavior is considered in the usual sense, that is, *a solution is called oscillatory if is has infinitely many zeros and nonoscillatory otherwise.* Rather important remark here is that the property of oscillation or nonoscillation is simply the behavior in the neighborhood of the infinite points. Thus initial data are not obviously involved in the resulting criteria. *The equation itself is oscillatory if all its solutions are oscillatory as well.*

III. The current state-of-the art

With regard to the immense ammount of obtained results, the oscillation theory for second-order ordinary differential equations is considered as well-developed, full and concise, turning more towards nonlinear and functional differential equations.

The first oscillation results for differential equations with deviating argument were obtained in the seminal paper by Fite [6] in 1921, who pointed out to various qualitative phenomena exclusively caused by the argument deviation. Since then, a great deal of the effort has been made by many researchers in order to advance the knowledge further. For the summary of most essential contributions, we refer the reader to the monographs by Agarwal et al. [7], [8], [9], [10], Došlý and Řehák [11] and Győri and Ladas [12]. Despite the intense research interest during the last fifty years, the development on the subject is still far from complete.

The equations (1) and (2) considered in our work may be regarded as natural generalizations of, e.g., the Sturm-Liouville differential equation

$$(r(t)y'(t))' + q(t)y(t) = 0,$$

or its half-linear extension

$$\left(r(t)\left(y'(t)\right)^{\alpha}\right)' + q(t)y^{\alpha}(t) = 0,$$

and the linear delay differential equation

$$y''(t) + q(t)y(\tau(t)) = 0 \qquad (5)$$

or the advanced differential equation

$$y''(t) + q(t)y(\sigma(t)) = 0. \qquad (6)$$

Although the new results which we present in Section V are formulated for general nonlinear cases, they are new even for the linear ones (5) and (6).

Most often, if an oscillation criterion or knowledge on asymptotic properties is obtained; it is appropriate to examine, in some way, *the sharpness* of this result. This is usually done by testing it on some particular cases of the equation studied, whose solution space is known. For this purpose, the so-called Euler equations are widely used.

Euler-type differential equations and their various generalizations have always played an important role in the oscillation theory since its origins. They are commonly used to test the strength of general criteria derived by different methods. The optimal scenario is when the obtained criterion gives a sharp result for the Euler-type equation; or at least it is closer to it for a given set of parameters, compared to another one. In the oscillation terminology, we simply examine how close we are to the *critical oscillation constant*[1] of a particular Euler differential equation.

---

[1]Together with the oscillation of (1) (or (2)), the term *conditional oscillation* is sometimes used for the situation, when there is a constant $\Gamma_0 > 0$ such that (1) (or (2)) with $q(t)$ replaced by $\Gamma q(t)$ is oscillatory for $\Gamma > \Gamma_0$ and nonoscillatory for $\Gamma < \Gamma_0$. The constant Γ_0 is termed *a critical oscillation constant*.

In the sequel, we will briefly survey some important results reported in the literature, serving as the main motivation of our research. Despite the fact that we are mainly interested by qualitative properties caused by the argument deviation; the first part is devoted to a short introduction to ordinary differential equations. Depending on the notion of canonical and noncanonical equations, we will devide our survey of known results for (1) into two parts, separately for (1) in canonical and noncanonical form. Since the theory of advanced equations of the form (2) and its particular cases is in general less developed, we devote to it the third part.

A. A note on ordinary differential equations

Having a natural importance as arising in the mechanical modelling and a variety of physical problems involved [13], the ordinary second-order differential equation

$$y''(t) + q(t)y(t) = 0 \tag{7}$$

(which is a special case of (1) or (2) with $\alpha = 1$, $r(t) = 1$, and $\tau(t) = \sigma(t) = t$) is one of the classical objects in the qualitative theory of linear differential equations. As already mentioned, this problem began as early as in 1836 by the work of Sturm [4] with the innovative idea to deduce oscillatory properties of unknown solutions of a given differential equation from known ones of an another, and was continued in 1893 by A. Kneser [14]. Essential contribution to the subject was made by E. Hille, A. Wintner, P. Hartman, W. Leighton, Z. Nehari and others (see, e.g., the monographs by R. Bellman [15], P. Hartman [16] and I. T. Kiguradze and T. A. Chanturia [17]).

Perhaps the most famous Euler differential equation

$$y''(t) + \frac{q_0}{t^2}y(t) = 0, \quad q_0 > 0, \tag{8}$$

with the so-called critical oscillation constant $\Gamma_0 = 1/4$, plays an important role in the oscillation theory of (7). Such equation can be regarded as a good comparative equation in the sense that (8) is oscillatory if and only if

$$q_0 > \frac{1}{4} \tag{9}$$

and consequently, (7) is oscillatory if

$$\liminf_{t \to \infty} t^2 q(t) > \frac{1}{4} \tag{10}$$

and nonoscillatory (meaning that all solutions are nonoscillatory) if

$$\limsup_{t \to \infty} t^2 q(t) < \frac{1}{4}.$$

This fact was firstly observed by A. Kneser [14] in 1893, by using Sturmian methods and (8) to show that (7) is oscillatory. It is useful to note that a corresponding result for the equation

$$(r(t)y'(t))' + q(t)y(t) = 0 \tag{11}$$

(which is a special case of (1) or (2) with $\alpha = 1$ and $\tau(t) = \sigma(t) = t$) can be obtained by the simple change of variables

$$s = R(t) \quad \text{and} \quad y(t) = u(s),$$

or

$$s = \frac{1}{\pi(t)} \quad \text{and} \quad y(t) = \frac{u(s)}{s},$$

depending on whether (11) is in canonical or noncanonical form, respectively. Then (11) in canonical form is oscillatory if

$$\liminf_{t \to \infty} r(t)R^2(t)q(t) > \frac{1}{4}$$

and nonoscillatory if

$$\limsup_{t \to \infty} r(t)R^2(t)q(t) < \frac{1}{4},$$

and similarly, (11) in noncanonical form is oscillatory if

$$\liminf_{t \to \infty} r(t)\pi^2(t)q(t) > \frac{1}{4}$$

and nonoscillatory if

$$\limsup_{t \to \infty} r(t)\pi^2(t)q(t) < \frac{1}{4},$$

In particular, many oscillation criteria have been found which involve the behavior of the integral of the coefficient q. For the sake of later comparison, we recall the famous Hille criterion [18] from 1948 upon which every solution of (7) oscillates if

$$\limsup_{t \to \infty} t \int_t^\infty q(s)\mathrm{d}s > 1$$

or

$$\liminf_{t \to \infty} t \int_t^\infty q(s)\mathrm{d}s > \frac{1}{4}. \tag{12}$$

Finally, we consider the half-linear ordinary differential equation

$$\left(r(t)\left(y'(t)\right)^\alpha\right)' + q(t)y^\alpha(t) = 0 \tag{13}$$

(which is a special case of (1) or (2) with $\tau(t) = \sigma(t) = t$).

Equation (13) has been deeply studied in the literature, as an important particular case of the so-called *Emden-Fowler* differential equation

$$\left(r(t)\left(y'(t)\right)^\alpha\right)' + q(t)y^\beta(t) = 0, \tag{14}$$

arising in the study of gas dynamics and fluid mechanics, relativistic mechanics, nuclear physics and in study of chemically reacting systems, see Wong [19].

The used terminology *half-linear* reflects the fact that the solution space is homogeneous (i.e., a constant multiple of every solution of (13) is also a solution of this equation), but not generally additive (the sum of two solutions in general fails to be a solution of (13)). The first papers focused to the second-order half-linear differential equation (13) are [20], [21]. Most of the results on half-linear differential equations known up to 2005 are sorted and summarized in the comprehensive book [11]. Since a lot of results from the linear oscillation theory for (7) are extendable almost verbatim to the half-linear case (13), it is reasonable to suppose that the critical oscillation constant can be found for the corresponding Euler-type half-linear equations as well. This hypothesis has been shown to be true for the couple of equations

$$\left(r(t)\left(y'(t)\right)^\alpha\right)' + \frac{q_0}{r^{1/\alpha}(t)R^{\alpha+1}(t)}y^\alpha(t) = 0, \quad q_0 > 0. \tag{15}$$

and

$$\left(r(t)\left(y'(t)\right)^{\alpha}\right)' + \frac{q_0}{r^{1/\alpha}(t)\pi^{\alpha+1}(t)}y^{\alpha}(t) = 0, \quad q_0 > 0. \quad (16)$$

which are associated to (13) in canonical and noncanonical case, respectively. It is well-known that (15) as well as (16) are oscillatory if and only if the associated characteristic equation

$$c(m) := \alpha m^{\alpha}(1 - m) = q_0$$

has no real roots what happens if

$$q_0 > \max\{c(m) : 0 < m < 1\}$$
$$= c\left(\frac{\alpha}{\alpha + 1}\right) = \left(\frac{\alpha}{\alpha + 1}\right)^{\alpha+1}, \quad (17)$$

cf. [7, Remark 3.7.1] or [11]. If condition (17) fails, then (15) has a nonoscillatory solution $R^m(t)$ or, (16) has a nonoscillatory solution $\pi^m(t)$. Therefore, $\Gamma_0 = (\alpha/(\alpha+1))^{\alpha+1}$ is a critical oscillation constant for both Euler equations (15) and (16). As an immediate consequence of the Sturmian comparison theorem and the above result, we get the following version of the classical Kneser oscillation and nonoscillation criterion for the canonical and noncanonical equation (13).

Theorem A: Let (3) hold. Suppose that

$$\liminf_{t\to\infty} r^{1/\alpha}(t)R^{\alpha+1}(t)q(t) > \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}.$$

Then (13) is oscillatory. If

$$\limsup_{t\to\infty} r^{1/\alpha}(t)R^{\alpha+1}(t)q(t) < \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1},$$

then (13) is nonscillatory.

Theorem B: Let (4) hold. Suppose that

$$\liminf_{t\to\infty} r^{1/\alpha}(t)\pi^{\alpha+1}(t)q(t) > \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}.$$

Then (13) is oscillatory. If

$$\limsup_{t\to\infty} r^{1/\alpha}(t)\pi^{\alpha+1}(t)q(t) < \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1},$$

then (13) is nonscillatory.

B. Review of known results for half-linear DDEs (1) *in canonical form*

The simplest linear second-order delay differential equation

$$y''(t) + q(t)y(\tau(t)) = 0 \quad (18)$$

(which is a special case of (1) with $\alpha = 1$ and $r(t) = 1$) has a long history. Most of the interest on this subject has been reflected by extensive references in monographs of Agarwal [7], [8], [9], Erbe et al. [22], Gyori and Ladas [12], and Hale [13].

In general, two factors mainly contribute to the qualitative behavior of (18): *the delay argument* and *the second-order nature* of the studied equation. Most often, their role has been taken separately in the literature; that is, one can select works providing conditions which ensure that (18) behaves

like ordinary differential equation and also such works which reveal a different behavior of delay equation. For example, despite that fact that the Sturmian theory (e.g. separation and comparison theorems) fails to extend to (18) due to delay argument (that is, the oscillatory solutions may coexist with nonoscillatory ones), it is possible to find criteria under which all solutions of (18) are oscillatory.

One of the basic techniques in oscillation theory is to acquire criteria by comparing the given differential equation with first-order delay differential equations or inequalities, whose oscillatory behavior is known in advance. The first results in this direction for second-order delay differential equations were given by Koplatadze [23] in 1986 and Wei [24] in 1988, who proved that the equation (18) is oscillatory if

$$L := \limsup_{t\to\infty} \int_{\tau(t)}^{t} \tau(s)q(s)\mathrm{d}s > 1 \quad (19)$$

or

$$k := \liminf_{t\to\infty} \int_{\tau(t)}^{t} \tau(s)q(s)\mathrm{d}s > \frac{1}{\mathrm{e}}. \quad (20)$$

There is an obvious gap between conditions (19)–(20) if $k < L$.

In 2000, Koplataze et al. [25] presented following oscillation criteria for (18) which improve (19)–(20), namely,

$$\limsup_{t\to\infty} \int_{\tau(t)}^{t} \left(\tau(s) + \int_{t_0}^{\tau(s)} \xi\tau(\xi)q(\xi)\mathrm{d}\xi\right) q(s)\mathrm{d}s > 1$$

or

$$\liminf_{t\to\infty} \int_{\tau(t)}^{t} \left(\tau(s) + \int_{t_0}^{\tau(s)} \xi\tau(\xi)q(\xi)\mathrm{d}\xi\right) q(s)\mathrm{d}s > \frac{1}{\mathrm{e}}. \quad (21)$$

In many works, however, authors aim to establish criteria which ensure that the studied delay equation behaves much like ordinary differential equation.

In 1971, Wong [26] generalized the famous Hille's condition (12) for the ordinary differential equation (7) and proved that if $\tau(t) \geq \alpha t$ and

$$\liminf_{t\to\infty} t\int_{t}^{\infty} q(s)\mathrm{d}s > \frac{1}{4\alpha}, \quad 0 < \alpha < 1,$$

then every solution of (18) is oscillatory. In 1973, Erbe [27] generalized the above condition to

$$\liminf_{t\to\infty} t\int_{t}^{\infty} q(s)\frac{\tau(s)}{s}\mathrm{d}s > \frac{1}{4}, \quad (22)$$

without any additional restriction of τ, at cost of the requirement that τ is monotone increasing.

To compare the strength of the above results due to Koplatadze and Erbe; we consider the linear Euler-type equation with proportional delay, namely,

$$y''(t) + \frac{q_0}{t^2}y(kt) = 0, \quad 0 < k \leq 1, \quad q_0 > 0. \quad (23)$$

Then conditions (20), (21), and (22) reduce to

$$kq_0 \ln\frac{1}{k} > \frac{1}{\mathrm{e}},$$

$$(k + k^2 q_0)q_0 \ln \frac{1}{k} > \frac{1}{e},$$

and

$$kq_0 > \frac{1}{4},$$

respectively. It can be observed that criteria (20) and (21) obtained from comparison with first-order delay differential equations, are more efficient for large delays, while Hille-type criterion (22) is more efficient when $\tau(t)$ approaches t. It is useful to notice that the critical oscillation constant can be obtained for (23). By transforming (23) into a constant-coefficient-constant-delay differential equation, Kulenovic [28] showed that (23) is oscillatory if and only if the assosiated characteristic equation

$$c(m) := m(1-m)k^{-m} = q_0$$

has no real root what happens if

$$q_0 > \max\{c(m) : 0 < m < 1\} = c(m_{\max}), =: \Gamma_0,$$

where

$$m_{\max} = \frac{-\sqrt{r^2 + 4} + r - 2}{2r}, \quad k = \exp(-r).$$

There has been a constant interest on how to obtain sharper results for (III-B).

In 1995, Kusano and Wang [29] used a variant of the Mahfoud's comparison principle [30] with the ordinary second-order differential equation

$$\left((x'(t))^\alpha\right)' + \frac{q(\tau^{-1}(t))}{\tau'\left(\tau^{-1}(t)\right)} x^\alpha(t) = 0$$

and proved that (1) is oscillatory if τ is nondecreasing and

$$\liminf_{t \to \infty} R^\alpha(\tau(t)) \int_t^\infty q(s) \mathrm{d}s > \frac{\alpha^\alpha}{(\alpha+1)^{\alpha+1}}. \quad (24)$$

If we consider the Euler-type half-linear extension of (23), namely,

$$\left((y'(t))^\alpha\right)' + \frac{q_0}{t^{\alpha+1}} y^\alpha(kt) = 0, \quad (25)$$

then motivated by the linear case, we may reveal after a direct substitution that (25) has a nonoscillatory solution $y(t) = t^m$, if

$$q_0 \leq \max\{c(m) : 0 < m < 1\}, \quad (26)$$

where

$$c(m) := \alpha m^\alpha (1-m) k^{-\alpha m}.$$

The "only if" part here is difficult to prove because the transformation to a constant-coefficient-constant-delay form as in [28] is impossible in the half-linear case. Note that condition (24) reduces to

$$q_0 k^\alpha > \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1} \quad (27)$$

for (25) to be oscillatory.

Most oscillation results for (1) existing in the literature use the Riccati transformation to reduce the second-order equation to a first-order Riccati inequality. In 2006, Sun and Meng [31] improved the oscillation result of Džurina and Stavroulakis [32] by employing the Riccati transformation

$$\omega(t) = R^\alpha(\tau(t)) \frac{r(t)(y'(t))^\alpha}{y^\alpha(\tau(t))},$$

which led to the following criterion for (1) to be oscillatory:

$$\int_{t_0}^\infty \left(R^\alpha(\tau(t))q(t) \right.$$
$$\left. - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1} \frac{\tau'(t)}{R(\tau(t))r^{1/\alpha}(\tau(t))} \right) \mathrm{d}t = \infty. \quad (28)$$

Erbe et al. [33] provided the following oscillation criterion for (1):

$$\limsup_{t \to \infty} \int_{t_0}^t \left(\psi(s)q(s) \left(\frac{\tau(s)}{s}\right)^\alpha \right.$$
$$\left. - \frac{(\psi'_+(s))^{\alpha+1} r(s)}{(\alpha+1)^{\alpha+1} \psi^\alpha(s)} \right) \mathrm{d}s = \infty, \quad (29)$$

where $\psi \in C^1([t_0, \infty), (0, \infty))$ and $\psi'_+(t) := \max\{0, \psi'(t)\}$, by considering the Riccati subsitution

$$\omega(t) = \psi(t) \frac{r(t)(y'(t))^\alpha}{y^\alpha(t)}. \quad (30)$$

Similar ideas as those above have been exploited and extended for (1) (and its various generalizations) in a number of papers, see, for instance, [34], [33], [35], [36], [32], [31], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46] and the references cited therein. The difference among them lies in choosing a particular Riccati-type substitution. In a consequent construction of a Riccati-type inequality, we need to eliminate the dependence on the quotient $x'(\tau(t))/x'(t)$ or $x(\tau(t))/x(t)$.

To the best of author's knowledge, all the above-mentioned papers provide at best (27) for the oscillation of (25). In [47], the authors suspected that "*the oscillation constant* (27) *is optimal and thus further refinements cannot be expected by decreasing q_0, but by studying various perturbations of the Euler type equation* (25)." As we will see, our results completely refute this claim.

C. Review of known results for half-linear DDEs (1) in noncanonical form

There is a significant difference in the structure of nonoscillatory (say positive) solutions between canonical and noncanonical equations. It is well known that the first derivative of any positive solution y of (1) is of one sign eventually, while (3) ensures that this solution is increasing eventually. On the other hand, when investigating noncanonical equations, both sign possibilities of the first derivative of any positive solution y have to be treated.

Because of a simpler structure of nonoscillatory solutions, (1) has been mostly studied in canonical form and much less efforts have been undertaken for noncanonical equations. The approach is twofold: first, it is possible to use factorization techniques [5] in order to transform the noncanonical equation

into a canonical one and apply some known principles. This is possible only for linear equations ($\alpha = 1$), see the works of Ohriska [48], Dzurina [49], and Kusano and Naito [50]. The second approach, also applicable for nonlinear equations, consists in extending known results for canonical ones, by eliminating both classes of possible nonoscillatory solutions; each by an appropriate condition, see [51], [52], [53], [54], [55], [56], [32], [31], [57].

In particular, Sun et al. [31] showed that if (28) holds and there exists a continuously differentiable function $\rho(t)$ such that $\rho(t) > 0$, $\rho'(t) \geq 0$ for $t \geq t_0$,

$$\int_{t_0}^{\infty} \rho(t)q(t)\mathrm{d}t = \infty \qquad (31)$$

and

$$\int_{t_0}^{\infty} \left(\frac{1}{r(t)\rho(t)} \int_{t_0}^{t} \rho(s)q(s)\mathrm{d}s \right)^{1/\alpha} \mathrm{d}t = \infty, \qquad (32)$$

then every solution $y(t)$ of (1) oscillates or $\lim_{t\to\infty} y(t) = 0$. Note that (28) in the above criterion eliminates existence of positive increasing solutions of (1), while conditions (31)–(32) ensure that any positive decreasing solution converges to zero in the neighborhood of infinity.

Recently, Mařík [55] revised Sun's result and proved that with no lack of generality we can put $\rho \equiv 1$ and the pair of conditions (31) and (32) can be safely and with no lack of generality replaced by one condition

$$\int_{t_0}^{\infty} \left(\frac{1}{r(t)} \int_{t_0}^{t} q(s)\mathrm{d}s \right)^{1/\alpha} \mathrm{d}t = \infty. \qquad (33)$$

Erbe et al. [53], [52], Hassan [54], and Saker [56] independently obtained the analogue of Mařík's result in such sense that (28) has been replaced by another condition resulting from the use of different techniques. Ye and Xu [57] proved that if (28) and

$$\int_{t_0}^{\infty} q(s)\pi^{\alpha+1}(s)\mathrm{d}s = \infty \qquad (34)$$

hold, then (1) is oscillatory.

It is well-known that the second-order Euler differential equation

$$(r(t)y'(t))' + \frac{q_0}{r(t)\pi^2(t)}y(t) = 0, \quad q_0 > 0, \qquad (35)$$

is oscillatory if and only if $q_0 > 1/4$. However, nor (33) neither (34) can be applied to (35), since

$$\int_{t_0}^{\infty} \frac{\ln t}{t^2}\mathrm{d}t < \infty.$$

Using the generalized Riccati substitution, Agarwal et al. [51] very recently proved oscillation of (1) via two independent conditions of limsup type, which are sharp when applied to (35). Namely, they showed that if $\alpha \geq 1$ and there exist functions $\rho, \delta \in C^1([t_0, \infty)(0, \infty))$ such that

$$\limsup_{t\to\infty} \int_{t_0}^{t} \left(\rho(s)q(s) - \frac{((\rho'(s))_+)^{\alpha+1} r(\tau(s))}{(\alpha+1)^{\alpha+1}\rho^\alpha(s)(\tau'(s))^\alpha} \right) \mathrm{d}s = \infty$$

and

$$\limsup_{t\to\infty} \int_{t_0}^{t} \left(\psi(s) - \frac{\delta(s)r(s)\left((\varphi(s))_+\right)^{\alpha+1}}{(\alpha+1)^{\alpha+1}} \right) \mathrm{d}s = \infty,$$

where

$$\psi(t) := \delta(t)\left(q(t) + \frac{1-\alpha}{r^{1/\alpha}(t)\pi^{\alpha+1}(t)} \right),$$

$$\varphi(t) := \frac{\delta'(t)}{\delta(t)} + \frac{1+\alpha}{r^{1/\alpha}(t)\pi(t)},$$

then (1) is oscillatory.

Finally, for comparison purposes, we consider the most general Euler-type half-linear delay differential equation

$$\left(r(t)\left(y'(t)\right)^\alpha \right)' + \frac{q_0}{r^{1/\alpha}(t)\pi^{\alpha+1}(t)}y^\alpha(\tau(t)) = 0, \qquad (36)$$

where $q_0 > 0$, the functions r and τ are such that $\pi(\tau(t))/\pi(t) = \lambda$. As previously, we find that (36) has a nonoscillatory solution $y(t) = \pi^m(t)$ if there is a real root of the equation

$$c(m) := \alpha m^\alpha(1-m)\lambda^{-\alpha m} = q_0,$$

what happens if

$$q_0 \leq \max\{c(m) : 0 < m < 1\}. \qquad (37)$$

It is worthwhile to mention that any of works [51], [52], [58], [53], [54], [55], [56], [32], [31], [57] gives at best

$$q_0 > \left(\frac{\alpha}{\alpha+1} \right)^{\alpha+1}$$

for (36) with $r(t) = t^{\alpha+1}$ and $\tau(t) = kt$, $k \in (0, 1]$, which is sharp only in the ordinary case (16) ($\tau(t) = t$). Here, it is easy to see that the influence of the delay is completely removed.

D. Review of known results for half-linear ADEs (2)

So far, most of the literature has been devoted to the investigation of differential equations with delay argument and very less is known up to now about their advances counterparts. For these equations, only partial results based on the application of classical comparison theorems or integral techniques have been published.

For the linear advanced-analogue of equation (18), namely,

$$y''(t) + q(t)y(\sigma(t)) = 0 \qquad (38)$$

(which is a special case of (2) with $\alpha = 1$ and $r(t) = 1$), two main approaches have appeared (see [8, Chapter 2], [59]). Taking Kusano's and Naito's comparison theorem [59, Theorem 1] into account, the oscillatory behavior of (38) can be deduced from that of the ordinary differential equation

$$y''(t) + q(t)y(t) = 0. \qquad (39)$$

On the other hand, another approach (see [8, Theorem 2.1.12]) was based on the comparison with the first-order advanced differential equation

$$y'(t) - \left(\int_{t}^{\infty} q(s)\mathrm{d}s \right) y(\sigma(t)) = 0, \qquad (40)$$

in the sense that oscillation of (38) is inherited from that of (40). Here, the advance may generate oscillations. In particular, by applying classical Kneser (10) and Hille's result (12), and the well-known oscillation criterion due to Ladas to (39) and (40), respectively, one can immediately get the following triple of oscillation criteria for (38):

$$\liminf_{t \to \infty} t^2 q(t) > \frac{1}{4}, \tag{41}$$

$$\liminf_{t \to \infty} t \int_t^\infty q(s)\mathrm{d}s > \frac{1}{4}, \tag{42}$$

and

$$\liminf_{t \to \infty} \int_t^{\sigma(t)} \int_u^\infty q(s)\mathrm{d}s\,\mathrm{d}u > \frac{1}{\mathrm{e}}. \tag{43}$$

Hence, a similar problem as for noncanonical delay equations occurs; that is, there has been no criterion which simultaneously takes into account the presence of the advance and the second order nature of the equation studied as well. The above types of results can be extended to half-linear equation (2) without much effort. The immediate application of Kusano and Wang's comparison theorem [29, Lemma 1] yields that (2) is oscillatory if the half-linear differential equation

$$\left(r(t)\left(y'(t)\right)^\alpha\right)' + q(t)y^\alpha(t) = 0 \tag{44}$$

is oscillatory. Using Theorem A from the previous section, we get a following variant of the classical Kneser oscillation criterion for (2), namely, (2) is oscillatory if

$$\liminf_{t \to \infty} r^{1/\alpha}(t)R^{\alpha+1}(t)q(t) > \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}. \tag{45}$$

The oscillation constant in (45) is a critical one in ordinary case; however, in case of $\sigma(t) \neq t$, the impact of the advanced argument is completely lost and the criterion is not sharp. To determine the gap between the oscillation and nonoscillation for advanced equations, we consider once again the Euler-type half linear advanced equation

$$\left(r(t)\left(y'(t)\right)^\alpha\right)' + \frac{q_0}{r^{1/\alpha}(t)R^{\alpha+1}(t)}y^\alpha(\sigma(t)) = 0 \tag{46}$$

with $R(\sigma(t))/R(t) = \lambda$. It can be directly verified that (46) has a nonoscillatory solution $y(t) = R^m(t)$ if m satisfies the indicial equation

$$c(m) := m^\alpha(1-m)\lambda^{-\alpha m} = q_0,$$

meaning that (46) has a nonoscillatory solution if

$$q_0 \leq \max\{c(m) : 0 < m < 1\} \tag{47}$$

and, by (45), it is oscillatory if

$$q_0 > \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1}. \tag{48}$$

In 2013, Cheng and Xu [60] showed that (2) has a nonoscillatory solution if and only if the sequence $\{v_n(t)\}_0^\infty$ defined by

$$v_0(t) = \int_t^\infty q(s)\mathrm{d}s,$$

$$v_n(t) = \alpha \int_t^\infty \frac{v_{n-1}^{(\alpha+1)/\alpha}(s)}{r^{1/\alpha}(s)}\mathrm{d}s$$
$$+ \int_t^\infty q(s)\exp\left(\alpha \int_s^{\sigma(s)}\left(\frac{v_{n-1}(\tau)}{r(\tau)}\right)^{1/\alpha}\mathrm{d}\tau\right)\mathrm{d}s,$$
$$n \in \mathbb{N}. \tag{49}$$

exists and converges, i.e., there exists a limit $\lim_{n\to\infty} v(t) = v(t) < \infty$. As a consequence, they formulated the following general oscillation result.

Theorem C (See [60]): Let

$$\int_{t_0}^\infty q(s)\mathrm{d}s < \infty.$$

Then (2) is oscillatory if one of the following condition holds:
1) There is an integer m such that v_n is defined for $n = 0, 1, \ldots, m-1$ but $v_m(t)$ does not exists.
2) The sequence $\{v_n(t)\}_0^\infty$ is defined for $n = 0, 1, \ldots$, but for arbitrarily large $t_* \geq t_0$,

$$\lim_{n\to\infty} v_n(t_*) = \infty.$$

However, the result is just theoretical since the integrals in (49) are difficult to compute, even for the Euler-type equation (46) with $\sigma(t) = kt$, $k \geq 1$, so only suitable estimates can be used instead. As a corollary of Theorem C, the authors obtained a variant of the Hille-type oscillation criterion stating that (**??**) is oscillatory if

$$\liminf_{t \to \infty} R^\alpha(t) \int_t^\infty q(s)\mathrm{d}s > \frac{\alpha^\alpha}{(\alpha+1)^{\alpha+1}},$$

being assumed to be safistied if the integral diverges. Again, it just requires (48) for (46) to be oscillatory. By reviewing the literature, one can notice that other works [8, Chapter 2], [61], [62], [63] give at best the same criterion for (46).

IV. PROBLEM STATEMENT

Second-order differential equations play a special role in the oscillation theory. Since the results given for these equations often form a basis for an appropriate generalization on higher-order equations; any modification of existing techniques which would yield sharper results and the observation of new asymptotic properties of solutions, is considered as highly valuable. Mathematical models of real processes involving FDEs cannot often explain at detail various qualitative phenomena caused by the response delay, due to the lack of corresponding mathematical framework. The main reason is that in subsequent analysis, the delay effect starts to be (at least partially) inadvertently neglected.

In spite of intense research on the oscillation theory of half-linear second-order differential equations with deviating

arguments which we outlined in the previous section, various open problems have occurred. Regardless on whether these equations were delayed or advanced, canonical or noncanonical, their common basis just lies in a (partial) neglect of the influence of the argument deviation to the qualitative properties of solutions.

Below, we formulate the main purpose of our research on the subject of qualitative theory of second-order differential equations.

1) Comparison principles are especially powerful tools in the oscillation theory since Sturm's initial contribution to the subject. Their underlying feature is to deduce oscillatory properties of the studied equation from those of a simpler one whose behavior is known in advance. In case of the half-linear equations (1) and (2) investigated in this chapter, these methods have been mostly applied:

   a) comparison with a second-order half-linear ordinary differential equation (13), directly or indirectly via a generalized Riccati generalized inequality;

   b) comparison with a second-order linear differential equation (18), by employing linearization techniques;

   c) comparison with a first-order linear differential equation with deviating argument, where the deviation is essential, but the information about the second-order nature of the equation is lost.

   It turns out to be crucial in all the above methods to obtain as sharp as possible estimates (lower and upper bounds) of nonoscillatory solutions. During the last decades, authors have tried to provide better and better approximations. So far, no method has been proposed which would allow to establish sharp *oscillation constants* for equations (1) and (2), even in a simpler canonical form.

2) The absolute majority of oscillation results in the literature concerns canonical case, since in this case the positive solutions of (1) exhibit simpler behavior. In noncanononical case, it is typical that the oscillation criterion consists of two relatively independent conditions. Also, a typical conclusion in the literature is weaker: all nonoscillatory solutions, if exist any, tend to zero. The secondary open problem arising here is to propose an unified approach allowing to reduce the number of conditions ensuring certain qualitative property of solutions of noncanonical equations.

3) One can observe there is a markable difference between investigation of oscillatory properties of first and second-order differential equations with deviating argument. While in the first-order case; the results for advaced-type equations are often analogous (to such an extent that the proof itself is often omitted); the situation is quite different in the second-order case. The oscillation theory of half-linear ADEs of the form (2) is generally less-developed. When the oscillation property is transferred

from second-order ordinary equations, the influence of the advanced argument is typically removed, even in the linear case.

4) The same as for first-order equations, a requirement on the monotone growth of the deviating argument has been posed in most criteria. It would be of interest to allow more complex behavior of the deviating function by relaxing this assumption.

V. METHODS AND ACHIEVED RESULTS

Our research mainly reflects the above-stated open problems occurring in the qualitative theory of second-order differential equations. Here, we tried to further develop and extend existing techniques, whose improvement lies in establishing and subsequence application of sharper estimates of a nonoscillatory solution. The newly introduced iterative estimates are used in specific directions, which allow us, for example, to obtain oscillation criteria for studied equations (1) and (2) which, in the ideal case, involve critical oscillation constants, and hence are unimprovable in some sense. The verification of the strength of the iterative results for which the construction of the limit case is not possible, is a formal process using the computational methods of MATLAB computing program.

Except for the use of computational tools in contructing desired estimates, we employed various techniques of mathematical and functional analysis that are modifications of standard methods used in the study of qualitative properties of functional differential equations (the use is made of transformation and linearization techniques, comparison principles, the method involving the generalized first-order Riccati inequality, integral averaging techniques, and various integral inequalities).

In summary, our contribution is contained in the following works:

[M1] **Improved oscillation results for second-order half-linear delay differential equations**. G.E Chatzarakis, I. Jadlovská - 2019. In: *Hacettepe Journal of Mathematics and Statistics*. Roč. 48, č. 1 (2019), s. 170-179. - ISSN 1303-5010

[M2] **Kneser-type oscillation criteria for second-order half-linear delay differential equations**, J. Džurina, I. Jadlovská, to appear in *Applied Mathematics and Computation*.

[M3] **A note on oscillation of second-order delay differential equations**. J. Džurina, I. Jadlovská - 2017. In: *Applied Mathematics Letters*. Vol. 69, no. 1 (2017), pp. 126-132. - ISSN 0893-9659

[M4] **Oscillatory results for second-order noncanonical delay differential equations**. J. Džurina, I. Jadlovská, I. P. Stavroulakis - 2019. In: *Opuscula Mathematica*. - Krakov (Poľsko): Wydział Matematyki Stosowanej Roč. 39, č. 4 (2019), s. 483-495. - ISSN 1232-9274

[M5] **A sharp oscillation result for second-order half linear noncanonical delay differential equations**, J.Džurina, I. Jadlovská, submitted to *Electronic Journal of Qualitative Theory of Differential Equations*.

[M6] **Iterative oscillation results for second-order differential equations with advanced argument**. I. Jadlovská - 2017. In: *Electronic Journal of Differential Equations*. Vol. 2017, no. 162 (2017), pp. 1-11. - ISSN 1072-6691

[M7] **A sharp oscillation criterion for second-order half-linear advanced differential equations**. G. E. Chatzarakis, I. Jadlovská, submitted to *Nonlinear Analysis*.

[M8] **New oscillation criteria for second-order half-linear advanced differential equations**. G. E Chatzarakis, J. Džurina, I. Jadlovská - 2019. In: *Applied Mathematics and Computation*- Amsterdam (Holandsko): Elsevier č. 347 (2019), s. 404-416. - ISSN 0096-3003

[M9] **Oscillation criteria of Kneser type for second-order half-linear advanced differential equations with non-canonical operators**. I. Jadlovská, to appear in *Applied Mathematics Letters*.

Depending on survey sections, these papers can be divided into several groups:

A Results for half-linear DDEs (1) in canonical form: [M1]–[M3]

B Results for half-linear DDEs (1) in noncanonical form: [M4]–[M5]

C Results for half-linear ADES (2) in noncanonical form: [M6]–[M9]

A. Part A: Results for half-linear DDEs (1) *in canonical form*

In the sequel, we will try to comprehensibly extract our contribution described in a couple of works [M1] and [M2]. Through the results of this section, we will briefly illustrate the general process of investigation. It starts with the classification of nonoscillatory, let us say positive solutions; which is indeed very simple in case of (1) in canonical form: any solution is strictly increasing.

Lemma 1: Assume that (1) has an eventually positive solution y. Then

$$y(t), \quad y'(t) > 0, \quad \left(r(t)\left(y'(t)\right)^\alpha\right)' < 0, \qquad (50)$$

eventually.

As already mentioned, one of the techniques in the oscillation theory of second-order differential equations which produces reasonably sharp results is the transformation to the first-order Riccati-type equation or inequality. This method is often combined with a suitable transformation, involving the ratio $r(t)\left(y'(t)\right)^\alpha/y^\alpha(\tau(t))$ or $r(t)\left(y'(t)\right)^\alpha/y^\alpha(t)$ for an eventually positive solution of (1). To generalize the theory of ordinary half-linear equations to those with deviating argument, the next step consists in employing a suibtable lower-bound of the corresponsponding quantity

$$\frac{r(\tau(t))\left(y'(\tau(t))\right)^\alpha}{r(t)\left(y'(t)\right)^\alpha} \qquad (51)$$

and

$$\frac{y(\tau(t))}{y(t)}, \qquad (52)$$

respectively. In the paper [M1], we adopted the approach involving the first possibility (51), somehow combined with a comparison principle involving first-order delay differential inequalities. Let the number ρ be defined by

$$\rho := \liminf_{t\to\infty} \int_{\tau(t)}^t q(s)R^\alpha(\tau(s))\mathrm{d}s,$$

and $\lambda(\eta)$ be the smaller positive root of the transcendental equation

$$\lambda = e^{\eta\lambda}, \quad 0 < \eta \le 1/\mathrm{e}.$$

Also, let us define the sequence of constants ρ_k as follows: set

$$\rho_1 := \rho$$

and, for $\rho_n \in (0, 1/\mathrm{e}]$, $n \in \mathbb{N}$, let

$$\rho_{n+1} := \liminf_{t\to\infty} \int_{\tau(t)}^t q(s)R_n^\alpha(\tau(s))\mathrm{d}s,$$

where

$$R_n(t) = R(t) + \frac{\lambda(\rho_n)}{\alpha} \int_{t_0}^t R(s)R^\alpha(\tau(s))q(s)\mathrm{d}s.$$

Lemma 2 (See [M1, Lemma 2.6]): Assume that $\rho > 0$ and (1) has an eventually positive solution. Then for any $n \in \mathbb{N}$,

$$0 < \rho_n \le \frac{1}{e},$$

and

$$\liminf_{t\to\infty} \frac{r(\tau(t))\left(y'(\tau(t))\right)^\alpha}{r(t)\left(y'(t)\right)^\alpha} \ge \lambda(\rho_n). \qquad (53)$$

As a consequence of Lemma 2, we formulated the following results.

Theorem 1 (See [M1, Theorem 2.7]): If, for some $n \in \mathbb{N}$, $\rho_i \le 1/\mathrm{e}$, $i = 0, 1, \ldots, n-1$, and

$$\rho_n > \frac{1}{e},$$

then (1) is oscillatory.

Remark 1: In the ordinary case $\alpha = 1$, Theorem 1 reduces to (20) and (21) for $n = 1$ and $n = 2$, respectively. Hence, it improves and generalizes the results of Koplatadze [23] and Wei [24], and Koplatadze et al. [25], respectively.

As a natural next step, we employed the estimate (53) in the Riccati technique.

Theorem 2 (See [M1, Theorem 2.8]): Assume that for some $n \in \mathbb{N}$, $0 < \rho_i \le 1/\mathrm{e}$, $i = 1, 2, \ldots, n$. If there is a costant $\epsilon > 0$ such that

$$\limsup_{t\to\infty} \int_{t_0}^t \left(R^\alpha(\tau(s))q(s) \right.$$
$$\left. - \left(\frac{\alpha}{\alpha+1}\right)^{\alpha+1} \frac{1}{(\lambda(\rho_n) - \epsilon)} \frac{\tau'(s)}{R(\tau(s))r^{1/\alpha}(\tau(s))} \right)\mathrm{d}s = \infty, \qquad (54)$$

then (1) is oscillatory.

The estimate (53) is crucial as well in the proof of the Hille-type oscillation result.

Theorem 3 (See [M1, Theorem 2.9]): Assume that for some $n \in \mathbb{N}$, $0 < \rho_n \leq 1/\mathrm{e}$, $i = 1, 2, \ldots, n$. If

$$\liminf_{t \to \infty} R^\alpha(\tau(t)) \int_t^\infty q(s)\mathrm{d}s > \frac{1}{\lambda(\rho_n)} \frac{\alpha^\alpha}{(\alpha+1)^{\alpha+1}}, \quad (55)$$

then (1) is oscillatory.

The results of the paper [M1], although disproving the claim that (27) is optimal for oscillation of (1), strongly depend on the properties of first-order delay differential equations and thus fail to apply when $k = 1$. An interesting problem for next research was to establish a different technique for testing oscillations in (**??**) independently on the constant $1/\mathrm{e}$.

Actually, we have made a further improvement of these results just recently in the paper [M2]. Although the idea is very simple and it essentially consists in appropriate selection of "*useful*" constants and sequentially improved monotonicities of a nonoscillatory solution, as we will see, it will yield to unimprovable results for (1). Also, we would like to stress that we have relaxed the assumption that τ is increasing.

Let us define

$$\begin{aligned} \beta_* &:= \frac{1}{\alpha} \liminf_{t \to \infty} r^{1/\alpha}(t) R^\alpha(\tau(t)) R(t) q(t) \\ \lambda_* &:= \liminf_{t \to \infty} \frac{R(t)}{R(\tau(t))}. \end{aligned} \quad (56)$$

Analogously to Lemma 1, we provided a more specific classification of nonoscillatory solutions, under the assumption of positive β_*.

Lemma 3 (See [M2, Lemma 1]): Let $\beta_* > 0$. If (1) has an eventually positive solution y, then

(i) y is eventually increasing;

(ii) $y(t)/R(t)$ is eventually decreasing such that $\lim_{t\to\infty} y(t)/R(t) = 0$;

(iii) the function

$$z(t) := y(t) - r^{1/\alpha}(t) y'(t) R(t) \quad (57)$$

is eventually positive and

$$z'(t) \geq \frac{\beta}{r^{1/\alpha}(t) R^\alpha(\tau(t))} \left(r^{1/\alpha}(t) y'(t) \right)^{1-\alpha} y^\alpha(\tau(t)) \quad (58)$$

for any $\beta \in (0, \beta_*)$ and t large enough.

Now, we dare to skip the process (to come back to it later) and just clearly state the main results: first, for the case of $\lambda_* = \infty$ and second, separately for $\alpha \geq 1$ and $\alpha \leq 1$.

Theorem 4 (See [M2, Theorem 2]): Let $\alpha \geq 1$ and $\lambda_* < \infty$. If

$$\beta_* > \max\{c(m) : 0 < m < 1\}, \quad (59)$$

where

$$c(m) := m(1-m)^\alpha \lambda_*^{-\alpha m},$$

then (1) is oscillatory.

Remark 2: In view of definitions of β_* and λ_*, condition (59) for Euler equation (25) becomes

$$\begin{aligned} q_0 &> \max\{\alpha m(1-m)^\alpha k^{-\alpha(1-m)} : 0 < m < 1\} \\ &= \max\{\alpha m^\alpha (1-m) k^{-\alpha m} : 0 < m < 1\}. \end{aligned}$$

Combining the above condition with (26), we conclude that (59) is sharp in the sense that the strict inequality cannot be replaced by the nonstrict one without affecting the validity of the theorem. Therefore, Theorem 4 can be viewed as a natural extension of Kneser oscillation theorem to a delay half-linear equation (1).

Corollary 1: By elementary calculus methods, we obtain

$$\max\{c(m) : 0 < m < 1\} = c(m_{\max}),$$

where

$$m_{\max} = \begin{cases} \dfrac{\alpha}{\alpha+1} & \text{for } \lambda_* = 1 \\ \dfrac{\alpha r + \alpha + 1 - \sqrt{(\alpha r + \alpha + 1)^2 - 4\alpha r}}{2\alpha r}, \\ \quad \text{for } \lambda_* \neq 1 \text{ and } r = \ln \lambda_*. \end{cases}$$

Theorem 5 (See [M2, Theorem 3]): Let $\alpha \leq 1$ and $\lambda_* < \infty$. If

$$\beta_* > \max\{d(m) : 0 < m < 1\}, \quad (60)$$

where

$$d(m) := m^\alpha (1-m) \lambda_*^{-\alpha m},$$

then (1) is oscillatory.

Remark 3: In case of $\alpha \leq 1$, there is still some gap between nonoscillation and oscillation constant for (25). How to obtain an oscillation criterion for (1) which would fill this gap, remains open at this moment.

Corollary 2: Here,

$$\max\{d(m) : 0 < m < 1\} = d(m_{\max}),$$

where

$$m_{\max} = \begin{cases} \dfrac{\alpha}{\alpha+1} & \text{for } \lambda_* = 1 \\ \dfrac{\alpha r + \alpha + 1 - \sqrt{(\alpha r + \alpha + 1)^2 - 4\alpha^2 r}}{2\alpha r} \\ \quad \text{for } \lambda_* \neq 1 \text{ and } r = \ln \lambda_*. \end{cases}$$

We conclude the main results with the statement for the remaining case $\lambda_* = \infty$.

Theorem 6 (See [M2, Theorem 4]): Let $\beta_* > 0$ and $\lambda_* = \infty$. Then (1) is oscillatory.

Remark 4: Notice that $\lambda_* = \infty$ if we choose e.g. $R(t) = t^\alpha$ and $\tau = t^\beta$, where $\alpha < 1$ and $\beta \in (0, 1)$.

Remark 5: The above results can be generalized without any difficulty to the nonlinear equation

$$\left(r(t) |y'|^{\alpha-1}(t) y'(t) \right)' + q(t) f(y(\tau(t))) = 0, \quad (61)$$

where $\alpha > 0$, f is a continuous function on $\mathbb{R}$ satisfying $f(u)/(|u|^{\alpha-1} u) \geq \gamma > 0$ for $u \neq 0$. The only change would be in defining β_* by

$$\beta_* = \frac{\gamma}{\alpha} \liminf_{t \to \infty} r^{1/\alpha}(t) R^\alpha(\tau(t)) R(t) q(t).$$

Now let us come back to Theorem 6 to see how it was obtained. In order to contruct the lower-bound of the quantity (52) in the case $\alpha \geq 1$, we defined a sequence $\{\beta_n\}$ by

$$\beta_0 := \beta_*$$
$$\beta_n := \frac{\beta_0 \lambda_*^{\alpha \beta_{n-1}}}{(1 - \beta_{n-1})^\alpha}, \quad n \in \mathbb{N}. \tag{62}$$

By induction on n, it is easy to verify that if, for some $n \in \mathbb{N}_0$, $\beta_i < 1$, $i = 0, 1, \ldots, n$, then β_{n+1} exists and

$$\beta_{n+1} = \ell_n \beta_n, \quad n \in \mathbb{N}_0, \tag{63}$$

where $\ell_n > 1$ satisfies

$$\ell_0 := \frac{\lambda_*^{\alpha \beta_0}}{(1 - \beta_0)^\alpha}$$
$$\ell_{n+1} := \lambda_*^{\alpha \beta_n (\ell_n - 1)} \left(\frac{1 - \beta_n}{1 - \ell_n \beta_n} \right)^\alpha, \quad n \in \mathbb{N}_0.$$

Lemma 4 (See [M2, Lemma 3]): Let $\alpha \geq 1$, $\beta_* > 0$ and $\lambda_* < \infty$. If (1) has an eventually positive solution y, then $y(t)/R^{1-\beta_n}(t)$ is eventually decreasing for any $n \in \mathbb{N}_0$.

We will sketch the outline of the proof. By means of induction, we prove that

$$\left(\frac{y(t)}{R^{1-\varepsilon_n \beta_n}(t)} \right)' < 0, \quad n \in \mathbb{N}_0, \tag{64}$$

eventually, where ε_n given by

$$\varepsilon_0 := \frac{\beta}{\beta_0}$$
$$\varepsilon_{n+1} := \varepsilon_0 \left(\frac{\lambda^{\varepsilon_n \beta_n}(1 - \beta_n)}{\lambda_*^{\beta_n}(1 - \varepsilon_n \beta_n)} \right)^\alpha$$

is an arbitrary constant from $(0, 1)$ approaching 1 as $\beta \to \beta_*$ and $\lambda \to \lambda_*$. Ommiting technical details, the idea lies in sequentially improving the monotonicity of a positive solution. Definitions of β_* and λ_* allow us to analytically compute the integrals in each step of the process. As a second step, we show that (64) implies the Lemma conclusion.

Corollary 3: Let $\alpha \geq 1$, $\beta_* > 0$ and $\lambda_* < \infty$. If (1) has an eventually positive solution y, then

$$\liminf_{t \to \infty} \frac{y(\tau(t))}{y(t)} \geq \lambda_*^{1 - \beta_n}. \tag{65}$$

The next step could consist in emplying the obtained estimate in a particular Riccati-type inequality. It is, however, not necessary. The proof of Theorem 6 is only based on a suitable combination of Lemmas 3 and 4, and properties of the sequence of β_n's.

Concerning the slightly worse result for $\alpha \leq 1$ provided by Theorem 5, it is caused by the fact that another sequence appears in the iterative process of improving monotonicities of a nonoscillatory solution.

To sum up, we have obtained an unimprovable Kneser-type oscillation criterion for (1) in canonical form, which improves all related results reported in the literature. Furthermore, in contrast with existing results, we have relaxed the assumption of the monotone growth of the delay argument, which has been traditionally imposed in related criteria.

B. Part B: Results for half-linear DDEs (1) *in noncanonical form*

First, we recall the correspoding structure of nonoscillatory solutions.

Lemma 5: Assume that (1) has an eventually positive solution y. Then either

$$y(t), \quad y'(t) > 0, \quad \left(r(t) (y'(t))^\alpha \right)' < 0 \tag{66}$$

or

$$y(t), \quad y'(t) < 0, \quad \left(r(t) (y'(t))^\alpha \right)' < 0, \tag{67}$$

eventually.

First results for (1) in noncanonical form are presented in the paper [M3]. Contrarily to the existing results stated in the introductory part of the chapter, we ensured the oscillation of (1) via only one condition. The method is twofold: first, we showed by careful observation that a condition (28) (or any of its analogue), which eliminates the solution class (66), is redundant. Also, we provided a simple refinement of condition (34) to be applicable on Euler-type equations. Second, we attempted to deduce oscillation of (1) in noncanonical form from that of the corresponding canonical linear differential equation

$$(\tilde{r}(t)y'(t))' + \tilde{q}(t)y(\tau(t)) = 0 \tag{68}$$

via suitable linearization technique.

Theorem 7 (See [M3, Theorem 1]): Assume that (33) holds. Then every solution $y(t)$ of (1) oscillates or $\lim_{t \to \infty} y(t) = 0$.

Theorem 8 (See [M3, Theorem 2]): Assume that one of the conditions (33) or (34) holds. Then (1) is oscillatory.

Remark 6: Theorems 7 and 8 essentially simplify a number of related criteria in the literature. In fact,

$$\int_{t_0}^\infty q(s) \mathrm{d}s = \infty \tag{69}$$

removes the possibility that any positive solution y of (1) is increasing. This is well known for canonical equations as Fite-Wintner's Theorem, but it has not been surprisingly used when studying noncanonical equations in works [52], [53], [54], [55], [56], [31], [57]. Note that the assumption (69) clearly gives no restriction to, e.g., Euler equation (35).

Theorem 9 (See [M3, Theorem 3].): Assume that

$$\limsup_{t \to \infty} \left(\pi^\alpha(t) \int_{t_0}^t q(s) \mathrm{d}s \right) > 1. \tag{70}$$

Then (1) is oscillatory.

Theorem 10 (See [M3, Theorem 4]): Let $\alpha \geq 1$. Assume that the equation (68) with

$$\tilde{r}(t) = \pi^2(t) r^{1/\alpha}(t), \quad \tilde{q}(t) = \frac{1}{\alpha} \pi^\alpha(t) \pi(\tau(t)) q(t)$$

is oscillatory. Then (1) is oscillatory.

In the work [M4], we made first improvements of the above results, giving sharper constants for Euler-type equations. They

are partly based on the use of Riccati technique, by which we improved Theorem 9 with the help of constants

$$K := \limsup_{t \to \infty} \pi(t) \left(\int_{t_0}^{t} q(s) \mathrm{d}s \right)^{1/\alpha},$$

$$k := \liminf_{t \to \infty} \frac{1}{\pi(t)} \int_{t}^{\infty} \pi^{\alpha+1}(s) q(s) \mathrm{d}s,$$

and

$$\ell \geq \frac{\pi(\tau(t))}{\pi(t)}.$$

In the sequel, we recall the main results of the paper [M4]. Lemma 6 just reformulates Theorem 7 while Lemma 7 is a key-result, as it provides a sharper information on the monotonicity of nonoscillatory solution. Theorems 11 and 12 give oscillation criteria for (1), directly improving Theorem 6. Finally, Theorem 13, based on a specific Riccati substitution and sharper estimates for nonoscillatory solutions; can be seen as a sort of Hille-type criterion, firstly taking the delay impact on the oscillatory propereties of solutions into account.

Lemma 6 (See [M4, Lemma 2.1]): Assume that

$$\int_{t_0}^{\infty} \frac{1}{r^{1/\alpha}(t)} \left(\int_{t_0}^{t} q(s) \mathrm{d}s \right)^{1/\alpha} \mathrm{d}t = \infty. \tag{71}$$

and (1) has an eventually positive solution y. Then y satisfies (67) eventually and

$$\lim_{t \to \infty} y(t) = 0.$$

Lemma 7 (See [M4, Lemma 2.4]): Let (71) hold. Assume that (1) has an eventually positive solution y, then $y/\pi^{K-\varepsilon}$ is decreasing for any fixed $\varepsilon > 0$ and t large enough.

Theorem 11 (See [M4, Theorem 2.3]): Let (71) hold. If

$$\liminf_{t \to \infty} \frac{1}{\pi(t)} \int_{t}^{\infty} \pi^{\alpha+1}(s) q(s) \mathrm{d}s > \alpha \tag{72}$$

or

$$k \leq \alpha \quad \text{and} \quad K > 1 - \frac{k}{\alpha},$$

then (1) is oscillatory.

Theorem 12 (See [M4, Theorem 2.6]): Let (71) hold. If (72) or

$$k \leq \alpha \quad \text{and} \quad \ell^K K > 1 - \frac{k}{\alpha},$$

then (??) is oscillatory.

Theorem 13 (See [M4, Theorem 2.8]): Let (71) hold. If

$$\ell^K k > \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha+1},$$

then (1) is oscillatory.

In the paper [M5], we offered, analogously to the canonical case, a Kneser-type oscillation criterion for (1) in noncanonical form. Different inequalities and iterative estimates are used to arrive at a final criterion, which, among others, proves a critical oscillation constant for the general Euler-type half-linear differential equation (36).

$$\beta_* := \liminf_{t \to \infty} \frac{1}{\alpha} r^{1/\alpha}(t) \pi^{\alpha+1}(t) q(t)$$

$$\lambda_* := \liminf_{t \to \infty} \frac{\pi(\tau(t))}{\pi(t)}. \tag{73}$$

Lemma 8 (See [M5, Lemma 1]): Let $\beta_* > 0$. If (??) has an eventually positive solution y, then

(i) y is eventually decreasing with $\lim_{t \to \infty} y(t) = 0$;
(ii) y/π is eventually nondecreasing.

Remark 7: Compared to Lemma 6, we replaced the integral condition (71) by the natural requirement of positive β_*. Define a sequence $\{\beta_n\}$ by

$$\beta_0 = \sqrt[\alpha]{\beta_*}$$

$$\beta_n = \frac{\beta_0 \lambda_*^{\beta_{n-1}}}{\sqrt[\alpha]{1 - \beta_{n-1}}}, \quad n \in \mathbb{N}. \tag{74}$$

Lemma 9 (See [M5, Lemma 2]): Let $\beta_* > 0$ and $\lambda_* < \infty$. If (1) has an eventually positive solution y, then for any $n \in \mathbb{N}_0$ y/π^{β_n} is eventually decreasing.

Theorem 14 (See [M5, Theorem 1]): Let $\lambda_* < \infty$ and

$$c(m) := m^\alpha (1 - m) \lambda_*^{-\alpha m}.$$

If

$$\beta_* > \max\{c(m) : 0 < m < 1\}, \tag{75}$$

then (1) is oscillatory.

Remark 8: In view of (37), condition (75) is sharp in the sense that the strict inequality cannot be replaced by the nonstrict one without affecting the validity of the theorem. For the remaining case $\lambda_* = \infty$, we have the following result.

Theorem 15 (See [M5, Theorem 2]): Let $\lambda_* = \infty$ and $\beta_* > 0$. Then (1) is oscillatory.

The above results significantly simplified the investigation of oscillatory properties of (1) in noncanonical form. In contrast with existing ones which consist of two indepedent conditions, we provide an easily verifiable criterion, involving critical oscillation constant for Euler-type half-linear noncanonical equation (36).

C. Part C: Results for half-linear ADEs (2)

A first improvement over the known results for the linear ADE (38) which we mentioned in the survey section, was proposed by Baculikova [64] in 2017. She investigated (38) in the case when (41) fails to apply, i.e., there is a constant $a \leq 1/4$ such that

$$t^2 q(t) \geq a, \tag{76}$$

and proved that if, for sufficiently large $t_1 \geq t_0$,

$$\limsup_{t \to \infty} \left(\sigma^{a-1}(t) \int_{t_1}^{t} s \sigma^{1-a}(s) q(s) \mathrm{d}s \right.$$

$$\left. + \frac{1}{\sigma^b(t)} \int_{t}^{\sigma(t)} s \sigma^b(s) q(s) \mathrm{d}s + \sigma^{1-b}(t) \int_{\sigma(t)}^{\infty} \sigma^b(s) q(s) \mathrm{d}s \right) > 1 \tag{77}$$

where
$$b = \frac{1 - \sqrt{1 - 4a}}{2},$$

then (38) is oscillatory. Further improvements with a quite different proof were obtained in the paper [M6]. Here, we used the recent results developed for first-order advanced differential equations which have been based on the iterative application of the Grönwall's inequality. This technique enables one to obtain sufficient conditions for oscillation of (2) involving $\liminf$, which essentially use the value of the advanced argument.

Let us recall the main results, which are as follows: Theorems 16 and 17 are two different initial improved variants of Kusano and Naito's comparison theorem for (38). In Lemma 10, a basic estimate for the ratio $y(\sigma(t))/y(t)$ is given, to be sequentially improved in Lemma 12. Another kind of iterative estimate is employed in Lemma 11 in order to improve Ladde's criterion (43), see Theorem 18. Lemmas 10 and 11 can be used to improve Theorems 16 and 17, see Theorems 20 and 21, respectively.

For the sake of clarify, we list first the functions and constants we used in work [M6]. That is, we put

$$\tilde{q}(t) = q(t)\left(1 + \int_t^{\sigma(t)} \int_u^\infty q(s)\mathrm{d}s\,\mathrm{d}u\right)$$

and we define the corresponding sequence of functions $\tilde{q}_n(t)$ by

$\tilde{q}_0(t) = \tilde{q}(t)$

$$\tilde{q}_n(t) = q(t)\left(1 + \int_t^{\sigma(t)} \int_u^\infty q(s)a_n(\sigma(s),t)\mathrm{d}s\,\mathrm{d}u\right), \quad n \in \mathbb{N},$$
(78)

where

$$a_1(s,t) = \exp\left(\int_t^s \int_u^\infty q(x)\mathrm{d}x\,\mathrm{d}u\right)$$

$$a_{n+1}(s,t) = \exp\left(\int_t^s \int_u^\infty q(x)a_n(\sigma(x),u)\mathrm{d}x\,\mathrm{d}u\right), \quad n \in \mathbb{N}.$$

Also, for the number ρ defined by

$$\rho := \liminf_{t\to\infty} \int_t^{\sigma(t)} \int_u^\infty q(s)\mathrm{d}s\mathrm{d}u,$$

and $\lambda(\eta)$ being the smaller positive root of the transcendental equation

$$\lambda = e^{\eta\lambda}, \quad 0 < \eta \le 1/\mathrm{e},$$

we define the sequence of constants ρ_n as follows: we set

$$\rho_1 := \rho$$

and, for $\rho_n \in (0, 1/\mathrm{e}]$, $n \in \mathbb{N}$, let

$$\rho_{n+1} := \liminf_{t\to\infty} \int_t^{\sigma(t)} \int_u^\infty q(s)a_n(\sigma(s),\sigma(u))\mathrm{d}s\,\mathrm{d}u.$$

Theorem 16 (See [M6, Theorem 2.3]): Assume that the second-order differential equation

$$y''(t) + \tilde{q}(t)y(t) = 0$$
(79)

is oscillatory. Then (38) is oscillatory.
Corollary 4: If

$$\liminf_{t\to\infty} t \int_t^\infty \tilde{q}(s)\mathrm{d}s > \frac{1}{4},$$
(80)

then (38) is oscillatory.
Lemma 10 (See [M6, Lemma 2.6]): Let $y(t)$ be an eventually positive solution of (38). Then

$$\rho \le \frac{1}{\mathrm{e}}$$
(81)

and

$$\liminf_{t\to\infty} \frac{y(\sigma(t))}{y(t)} \ge \lambda(\rho).$$
(82)

Theorem 17 (See [M6, Theorem 2.7]): Let (81) hold. Assume that the second-order differential equation

$$y''(t) + k\lambda(\rho)q(t)y(t) = 0$$
(83)

is oscillatory for some $k \in (0,1)$. Then (38) is oscillatory.
Corollary 5: Let (81) hold. If

$$\liminf_{t\to\infty} t \int_t^\infty q(s)\mathrm{d}s > \frac{1}{4\lambda(\rho)},$$
(84)

then (38) is oscillatory.
Lemma 11 (See [M6, Lemma 2.9]): Let $y(t)$ be an eventually positive solution of (38). Then

$$y(s) \ge y(t)a_n(s,t), \quad s \ge t,$$
(85)

for t large enough.
Theorem 18 (See [M6, Theorem 2.10]): Assume that the first-order advanced differential equation

$$y'(t) - \left(\int_t^\infty q(s)a_n(\sigma(s),\sigma(t))\mathrm{d}s\right)y(\sigma(t)) = 0$$
(86)

is oscillatory for some $n \in \mathbb{N}$. Then (38) is oscillatory.
Corollary 6: If

$$\liminf_{t\to\infty} \int_t^{\sigma(t)} \int_u^\infty q(s)a_n(\sigma(s),\sigma(u))\mathrm{d}s\,\mathrm{d}u > \frac{1}{\mathrm{e}},$$
(87)

for some $n \in \mathbb{N}$, then (38) is oscillatory.
Remark 9: The above theorem permits us to deduce oscillation of (38) from that of the first-order advanced differential equation (86). One can see that, even for $n = 1$, the criterion (87) is sharper than (43) and thus provides a better result.
Theorem 19 (See [M6, Theorem 2.13]): Assume that the second-order differential equation

$$y''(t) + q(t)a_n(\sigma(t),t)y(t) = 0$$
(88)

is oscillatory for some $n \in N$. Then (38) is oscillatory.
Corollary 7: If

$$\liminf_{t\to\infty} t \int_t^\infty q(s)a_n(\sigma(s),s)\mathrm{d}s > \frac{1}{4}$$
(89)

for some $n \in \mathbb{N}$, then (38) is oscillatory.
Theorem 20 (See [M6, Theorem 2.15]): Assume that the second-order differential equation

$$y''(t) + \tilde{q}_n(t)y(t) = 0$$
(90)

is oscillatory for some $n \in \mathbb{N}_0$. Then (38) is oscillatory.

Corollary 8: If

$$\liminf_{t \to \infty} t \int_t^\infty \tilde{q}_n(s)\mathrm{d}s > \frac{1}{4} \qquad (91)$$

for some $n \in \mathbb{N}$, then (38) is oscillatory.

Lemma 12 (See [M6, Lemma 2.17]): Let $y(t)$ be an eventually positive solution of (38). Then for any $n \in \mathbb{N}$

$$\rho_n \leq \frac{1}{\mathrm{e}} \qquad (92)$$

and

$$\liminf_{t \to \infty} \frac{y(\sigma(t))}{y(t)} \geq \lambda(\rho_n).$$

Theorem 21 (See [M6, Theorem 2.18]): Let $\rho_i \leq 1/e$ for $i = 0, 1, 2, \ldots, n$. Assume that the second-order differential equation

$$y''(t) + k\lambda(\rho_n)q(t)y(t) = 0 \qquad (93)$$

is oscillatory for some $k \in (0, 1)$. Then (38) is oscillatory.

Corollary 9: Let $\rho_i \leq 1/e$ for $i = 0, 1, 2, \ldots, n$. If

$$\liminf_{t \to \infty} t \int_t^\infty q(s)\mathrm{d}s > \frac{1}{4\lambda(\rho_n)}, \qquad (94)$$

then (38) is oscillatory.

Example 1: Consider the second-order advanced Euler differential equation

$$y''(t) + \frac{q_0}{t^2}y(kt) = 0, \quad k \geq 1, \quad q_0 > 0, \quad t \geq 1. \qquad (95)$$

Known oscillation criteria (42) and (43) give

$$q_0 > \frac{1}{4} \qquad (96)$$

and

$$q_0 \ln k > \frac{1}{\mathrm{e}}, \qquad (97)$$

respectively. Besides, the recent result (77) gives

$$q_0 \left(\frac{k^\beta - 1}{\beta} + \frac{1}{1 - q_0} + \frac{k^\beta}{1 - \beta} \right) > 1, \qquad (98)$$

where

$$\beta = \frac{1 - \sqrt{1 - 4q_0}}{2}$$

and $q_0 \leq 1/4$.

From Corollary 4, we have that (95) is oscillatory if

$$q_0(1 + q_0 \ln k) > \frac{1}{4}. \qquad (99)$$

To apply Corollary 5, we set

$$\rho := q_0 \ln k \leq 1/e.$$

Then the smaller root of the equation $\lambda = \mathrm{e}^{\rho\lambda}$ is given by

$$\lambda(\rho) = -\frac{W(-\ln \mathrm{e}^\rho)}{\ln \mathrm{e}^\rho} = -\frac{W(-\rho)}{\rho},$$

TABLE I
COMPARISON OF THE STRENGTH OF CRITERIA (96)–(104) FOR A GIVEN
$q_0 = 0.23$

criterion	required k
(96)	inapplicable
(97)	4.950436
(98)	2.274700
(99)	1.459467
(100)	1.395881
(101)	3.426695
(102)	1.436966
(103)	1.304194
(104)	1.292806

where $W(\cdot)$ standardly denotes the principal branch of the Lambert function. Consequently, the oscillatory criterion (84) becomes

$$-q_0 \frac{W(-\rho)}{\rho} > \frac{1}{4},$$

that is,

$$-\frac{W(-q_0 \ln k)}{\ln k} > \frac{1}{4}. \qquad (100)$$

Now, set $n = 1$. Performing numerical computations, the following conditions for oscillation of (95), i.e.,

$$\frac{q_0}{1 - q_0} \ln k > \frac{1}{e}, \qquad (101)$$

$$q_0 k^{q_0} > \frac{1}{4}, \qquad (102)$$

$$q_0 \left(1 + \frac{k^{q_0}(k^{q_0} - 1)}{1 - q_0} \right) > \frac{1}{4}, \qquad (103)$$

and

$$\frac{(q_0 - 1)W\left(\frac{q_0}{q_0 - 1} \ln k\right)}{\ln k} > \frac{1}{4}, \qquad (104)$$

where

$$\frac{q_0}{1 - q_0} \ln k \leq 1/e,$$

result from Corollaries 6, 7, 8 and 9, respectively. A comparison of the effectiveness of the above-mentioned criteria in terms of the required value k for a given coefficient $q_0 = 0.23$ is shown in the Table I.

On the other hand, if we set $q_0 = 0.19$ and $k = 2$ in Eq. (95), then it is easy to verify that all criteria (97)−(104) fail. In such a case, it is interesting to compare the length of the iteration process in particular cases corresponding to Corollaries 6-9. As can be seen from Table II, 13 iteration steps are necessary when applying Corollary 6, Corollary 7 requires 7 steps, while Corollaries 8 and 9 ensure the oscillation of (95) after the same number of iterations (6 steps).

In accordance with the results presented for half-linear DDEs (1), we obtained sharp oscillation criteria for (2) in canonical

TABLE II
COMPARISON OF ITERATIVE PROCESSES FOR (95) RESULTING FROM
COROLLARIES 6-9, RESPECTIVELY.

n	critical value $1/e$ (Cor. 6)
1	0.162590
2	0.179813
3	0.191500
4	0.200467
5	0.208003
6	0.214830
7	0.221447
8	0.235846
9	0.244837
10	0.256514
11	0.273525
12	0.302947
13	0.372771

n	critical value $1/4$ (Cor. 7)
1	0.216745
2	0.228721
3	0.235918
4	0.241002
5	0.245011
6	0.248452
7	0.251627

n	critical value $1/4$ (Cor. 8)	n	critical value $1/4$ (Cor. 9)
1	0.231658	1	0.227666
2	0.237998	2	0.235188
3	0.24264	3	0.240441
4	0.246414	4	0.244541
5	0.249743	5	0.248028
6	0.252893	6	0.251219

and noncanonical form. As usual, we formulated final criteria in terms of constants β_* and λ_* defined by

$$\beta_* := \frac{1}{\alpha} \liminf_{t \to \infty} r^{1/\alpha}(t) R^{\alpha+1}(t) q(t)$$
$$\lambda_* := \liminf_{t \to \infty} \frac{R(\sigma(t))}{R(t)} \tag{105}$$

or

$$\beta_* := \frac{1}{\alpha} \liminf_{t \to \infty} r^{1/\alpha}(t) \pi^\alpha(\sigma(t)) \pi(t) q(t)$$
$$\lambda_* := \liminf_{t \to \infty} \frac{\pi(t)}{\pi(\sigma(t))}, \tag{106}$$

in canonical and noncanonical case, respectively. Corresponding criteria are stated for both canonical and noncanonical case of (2) in Theorem 22. In both cases, they provide a necessary and sufficient condition for oscillation of corresponding half-linear advanced Euler differential equation.

Theorem 22 (See [M7, Theorem 1]): Assume that $\lambda_* < \infty$ and

$$\beta_* > \max\{c(m) : 0 < m < 1\}, \tag{107}$$

where

$$c(m) := \begin{cases} m^\alpha (1-m) \lambda_*^{-\alpha m} & \text{in canononical case} \\ m(1-m)^\alpha \lambda_*^{-\alpha m} & \text{in noncanonical case} \end{cases}$$

Then (2) is oscillatory.

Theorem 23: Let $\beta_* > 0$ and $\lambda_* = \infty$. Then (2) is oscillatory.

As a final remark, we have relaxed the assumption on σ to be increasing.

VI. CONCLUSION

In the paper, we have dealt with the oscillation problem for second-order half-linear differential equations with deviating argument of the form (1) and (2).

Here, we have proposed a unified approach for investigation of these nonlinear equations, no matter whether they are in canonical or noncanonical form, or with delay or advanced argument. For each possible case, we have provided unbeatable oscillation constants. Our findings are of importance in the following specific directions, related to arising open problems stated in Section IV.

- We have provided a detailed asymptotic characterization of nonoscillatory solutions of (1) and (2) and their quasiderivative(s).
- The newly proposed procedure brought new results even for the simplest (and mostly studied) linear delay differential equations of the second order.
- We have built the concise oscillation theory for second-order advanced differential equations which essentially uses the value of the advanced argument.
- We have offered effective techniques for investigating nonlinear differential equations in noncanonical form, which essentially simplify and improve existing results.
- We have relaxed any restriction posed on the deviating function (boundedness, monotone growth).
- We have extended these results to a class of neutral differential equations.

To sum up, the main benefits of the work consist of developing a complex methodology for investigating oscillatory and asymptotic properties of solutions to studied equations. The basis of newly employed techniques mostly lies in establishing and subsequent application of iterative estimates relating a nonoscillatory solution with its derivatives, which, in the limit case, lead to the necessary and sufficient conditions for the oscillation of these equations. The resulting criteria have been verified on suitable comparison equations, using the computational methods. The novelty of the approach allows to obtain sharp results for several classes of nonlinear equations.

REFERENCES

[1] Y. Kuang, *Delay differential equations: with applications in population dynamics.* Academic Press, 1993, vol. 191.

[2] V. Kolmanovskii and A. Myshkis, *Introduction to the theory and applications of functional differential equations.* Springer, 1999, vol. 463.

[3] J. Chiasson and J. J. Loiseau, *Applications of time delay systems.* Springer, 2007, vol. 352.

[4] C. Sturm, "Mémoire sur les équations différentielles linéaires du second ordre," *J. Math. Pures Appl.*, no. 1, p. 106–186, 1936.

[5] W. F. Trench, "Canonical forms and principal systems for general disconjugate equations," *Trans. Amer. Math. Soc.*, vol. 189, pp. 319–327, 1973.

[6] W. B. Fite, "Concerning the zeros of the solutions of certain differential equations," *Trans. Amer. Math. Soc.*, vol. 19, no. 4, pp. 341–352, 1918.

[7] R. P. Agarwal, S. R. Grace, and D. O'Regan, *Oscillation theory for second order linear, half-linear, superlinear and sublinear dynamic equations*. Kluwer Academic Publishers, Dordrecht, 2002.

[8] ——, *Oscillation theory for second order dynamic equations*, ser. Series in Mathematical Analysis and Applications. Taylor & Francis, Ltd., London, 2003, vol. 5.

[9] R. P. Agarwal, M. Bohner, and W.-T. Li, *Nonoscillation and Oscillation: Theory for Functional Differential Equations*, ser. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 2004, vol. 267.

[10] R. P. Agarwal, S. R. Grace, and D. O'Regan, *Oscillation theory for difference and functional differential equations*. Springer Science & Business Media, 2013.

[11] O. Došlý and P. Rehák, *Half-linear differential equations*, ser. North-Holland Mathematics Studies. Elsevier Science B.V., Amsterdam, 2005, vol. 202.

[12] I. Győri and G. Ladas, *Oscillation theory of delay differential equations*, ser. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1991, with applications, Oxford Science Publications.

[13] J. K. Hale, *Functional differential equations*. Springer, 1971.

[14] A. Kneser, "Untersuchungen über die reellen nullstellen der integrale linearer differentialgleichungen," *Mathematische Annalen*, vol. 42, no. 3, pp. 409–435, 1893.

[15] R. E. Bellman and K. L. Cooke, "Differential-difference equations," 1963.

[16] P. Hartman, *Ordinary differential equations*, 2nd ed. Birkhäuser, Boston, Mass., 1982.

[17] I. T. Kiguradze and T. A. Chanturia, *Asymptotic properties of solutions of nonautonomous ordinary differential equations*, ser. Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1993, vol. 89, translated from the 1985 Russian original.

[18] E. Hille, "Non-oscillation theorems," *Transactions of the American Mathematical Society*, vol. 64, no. 2, pp. 234–252, 1948.

[19] J. S. Wong, "On the generalized emden–fowler equation," *Siam Review*, vol. 17, no. 2, pp. 339–360, 1975.

[20] Á. Elbert, "Asymptotic behaviour of autonomous half-linear differential systems on the plane," *Studia Sci. Math. Hungar*, vol. 19, no. 2-4, pp. 447–464, 1984.

[21] J. Mirzov, "On some analogs of sturm's and kneser's theorems for nonlinear systems," *Journal of Mathematical Analysis and Applications*, vol. 53, no. 2, pp. 418–425, 1976.

[22] L. H. Erbe, Q. Kong, and B. G. Zhang, *Oscillation theory for functional-differential equations*, ser. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 1995, vol. 190.

[23] R. G. Koplatadze, "Criteria for the oscillation of solutions of differential inequalities and second-order equations with retarded argument," *Tbiliss. Gos. Univ. Inst. Prikl. Mat. Trudy*, vol. 17, pp. 104–121, 1986.

[24] J. J. Wei, "Oscillation of second order delay differential equation," *Ann. Differential Equations*, vol. 4, no. 4, pp. 473–478, 1988.

[25] R. Koplatadze, G. Kvinikadze, and I. P. Stavroulakis, "Oscillation of second order linear delay differential equations," *Funct. Differ. Equ.*, vol. 7, no. 1-2, pp. 121–145, 2000.

[26] J. S. Wong, "Second order oscillation with retarded arguments," in *Ordinary differential equations*. Elsevier, 1972, pp. 581–596.

[27] L. Erbe, "Oscillation criteria for second order nonlinear delay equations," *Canadian Mathematical Bulletin*, vol. 16, no. 1, pp. 49–56, 1973.

[28] M. Kulenovic, "Oscillation of the euler differential equation with delay," *Czechoslovak Mathematical Journal*, vol. 45, no. 1, pp. 1–6, 1995.

[29] T. Kusano and J. Wang, "Oscillation properties of half-linear functional-differential equations of the second order," *Hiroshima Math. J.*, vol. 25, no. 2, pp. 371–385, 1995.

[30] W. E. Mahfoud, "Comparison theorems for delay differential equations," *Pacific Journal of Mathematics*, vol. 83, no. 1, pp. 187–197, 1979.

[31] Y. G. Sun and F. W. Meng, "Note on the paper of J. Džurina and I. P. Stavroulakis: "Oscillation criteria for second-order delay differential equations"," *Appl. Math. Comput.*, vol. 174, no. 2, pp. 1634–1641, 2006.

[32] J. Džurina and I. P. Stavroulakis, "Oscillation criteria for second-order delay differential equations," *Appl. Math. Comput.*, vol. 140, no. 2-3, pp. 445–453, 2003.

[33] L. Erbe, T. S. Hassan, A. Peterson, and S. Saker, "Oscillation criteria for half-linear delay dynamic equations on time scales," *Nonlinear Dyn. Syst. Theory*, vol. 9, no. 1, pp. 51–68, 2009.

[34] R. P. Agarwal, S.-L. Shieh, and C. C. Yeh, "Oscillation criteria for second-order retarded differential equations," *Math. Comput. Modelling*, vol. 26, no. 4, pp. 1–11, 1997.

[35] B. Baculíková and J. Džurina, "Oscillation theorems for second-order nonlinear neutral differential equations," *Comput. Math. Appl.*, vol. 62, no. 12, pp. 4472–4478, 2011.

[36] T. Candan, "Oscillatory behavior of second order nonlinear neutral differential equations with distributed deviating arguments," *Appl. Math. Comput.*, vol. 262, pp. 199–203, 2015.

[37] S. Fišnarová and R. Mařík, "Oscillation of half-linear differential equations with delay," *Abstr. Appl. Anal.*, pp. Art. ID 583 147, 6, 2013.

[38] L. Erbe, A. Peterson, and S. H. Saker, "Kamenev-type oscillation criteria for second-order linear delay dynamic equations," *Dynam. Systems Appl.*, vol. 15, no. 1, pp. 65–78, 2006.

[39] A. Tiryaki, "Oscillation criteria for a certain second-order nonlinear differential equations with deviating arguments," *Electron. J. Qual. Theory Differ. Equ.*, no. 61, pp. 1–11, 2009.

[40] Z. Han, T. Li, S. Sun, and Y. Sun, "Remarks on the paper [Appl. Math. Comput. 207 (2009) 388–396] [mr2489110]," *Appl. Math. Comput.*, vol. 215, no. 11, pp. 3998–4007, 2010.

[41] L. Liu and Y. Bai, "Erratum to: "New oscillation criteria for second-order nonlinear neutral delay differential equations" [J. Comput. Appl. Math. 231 (2009) 657–663] [mr2549731]," *J. Comput. Appl. Math.*, vol. 233, no. 10, p. 2755, 2010.

[42] H. Wu, L. Erbe, and A. Peterson, "Oscillation of solution to second-order half-linear delay dynamic equations on time scales," *Electron. J. Differential Equations*, pp. Paper No. 71, 15, 2016.

[43] J.-G. Dong, "Oscillation behavior of second order nonlinear neutral differential equations with deviating arguments," *Comput. Math. Appl.*, vol. 59, no. 12, pp. 3710–3717, 2010.

[44] R. Xu and F. Meng, "Some new oscillation criteria for second order quasi-linear neutral delay differential equations," *Appl. Math. Comput.*, vol. 182, no. 1, pp. 797–803, 2006.

[45] S. Sun, Z. Han, and C. Zhang, "Oscillation of second-order delay dynamic equations on time scales," *Journal of Applied Mathematics & Computing*, vol. 30, no. 1-2, p. 459, 2009.

[46] Z. Han, S. Sun, and B. Shi, "Oscillation criteria for a class of second-order emden–fowler delay dynamic equations on time scales," *Journal of Mathematical Analysis and Applications*, vol. 334, no. 2, pp. 847–858, 2007.

[47] S. Fisnarova and R. Mařík, "Modified riccati technique for half-linear differential equations with delay," *Electronic Journal of Qualitative Theory of Differential Equations*, vol. 2014, no. 64, pp. 1–14, 2014.

[48] J. Ohriska, "Oscillation theorems for non-canonical self-adjoint differential equations of second order," *Zeitschrift für Analysis und ihre Anwendungen*, vol. 21, no. 2, pp. 515–520, 2002.

[49] J. Džurina, "Oscillation of a second order delay differential equations," *Archivum Mathematicum*, vol. 33, no. 3, pp. 309–314, 1997.

[50] T. Kusano and Y. Naito, "Oscillation and nonoscillation criteria for second order quasilinear differential equations," *Acta Math. Hungar.*, vol. 76, no. 1-2, pp. 81–99, 1997.

[51] R. P. Agarwal, C. Zhang, and T. Li, "Some remarks on oscillation of second order neutral differential equations," *Appl. Math. Comput.*, vol. 274, pp. 178–181, 2016.

[52] L. Erbe, T. S. Hassan, and A. Peterson, "Oscillation criteria for nonlinear damped dynamic equations on time scales," *Appl. Math. Comput.*, vol. 203, no. 1, pp. 343–357, 2008.

[53] L. Erbe, A. Peterson, and S. H. Saker, "Oscillation criteria for second-order nonlinear delay dynamic equations," *J. Math. Anal. Appl.*, vol. 333, no. 1, pp. 505–522, 2007.

[54] T. S. Hassan, "Oscillation criteria for second-order nonlinear dynamic equations," *Adv. Difference Equ.*, pp. 2012:171, 13, 2012.

[55] R. Mařík, "Remarks on the paper by Sun and Meng, Appl. Math. Comput. 174 (2006)," *Appl. Math. Comput.*, vol. 248, pp. 309–313, 2014.

[56] S. H. Saker, "Oscillation criteria of second-order half-linear dynamic equations on time scales," *J. Comput. Appl. Math.*, vol. 177, no. 2, pp. 375–387, 2005.

[57] L. Ye and Z. Xu, "Oscillation criteria for second order quasilinear neutral delay differential equations," *Appl. Math. Comput.*, vol. 207, no. 2, pp. 388–396, 2009.

[58] P. Drábek, A. Kufner, and K. Kuliev, "Oscillation and nonoscillation results for solutions of half-linear equations with deviated argument," *Journal of Mathematical Analysis and Applications*, vol. 447, no. 1, pp. 371–382, 2017.

[59] T. Kusano and M. Naito, "Comparison theorems for functional-differential equations with deviating arguments," *J. Math. Soc. Japan*, vol. 33, no. 3, pp. 509–532, 1981.

[60] A. Cheng and Z. Xu, "Existence of non-oscillatory solutions for second-order advanced half-linear differential equations," *Electronic Journal of Differential Equations*, vol. 2013, no. 54, pp. 1–10, 2013.

[61] J. Džurina, "Oscillation of second order differential equations with advanced argument," *Mathematica Slovaca*, vol. 45, no. 3, pp. 263–268, 1995.

[62] R. Koplatadze, G. Kvinikadze, and I. P. Stavroulakis, "Properties a and b of nth order linear differential equations with deviating argument," *Georgian Mathematical Journal*, vol. 6, no. 6, pp. 553–566, 1999.

[63] T. Kusano and M. Naito, "Comparison theorems for functional differential equations with deviating arguments," *Journal of the Mathematical Society of Japan*, vol. 33, no. 3, pp. 509–532, 1981.

[64] B. Baculíková, "Oscillatory behavior of the second order functional differential equations," *Applied Mathematics Letters*, vol. 72, pp. 35–41, 2017.

Intelligent Human-System Interoperability with Edge Computing Support

Erik Kajati
Technical University of Kosice
Faculty of Electrical Engineering and Informatics
Department of Cybernetics and Artificial Intelligence
Kosice, Slovakia
email: erik.kajati@tuke.sk

Iveta Zolotova
Technical University of Kosice
Faculty of Electrical Engineering and Informatics
Department of Cybernetics and Artificial Intelligence
Kosice, Slovakia
email: iveta.zolotova@tuke.sk

Abstract—Recent industrial evolution changes the habits of humans jobs. With developing interactions, there is also a need to discuss the issue of blending people into this new ecosystem. Throughout past years the technologies and techniques that can support intelligent interoperability between humans and systems have emerged. The new concept of work assisted by machines demands "smart interaction" with devices principally characterized as physical and cognitive interactions. Assigning a vast number of sensors in the environment can contribute to a better understanding of the entire space. We think that the leading role will have edge computing along with varying methods of indoor localization. In this paper, we have described Human-System Interoperability and several types of interactions. After that, we have concentrated on edge computing along with the network evaluation study. One of the case studies that arise from the ideas were introduced. These studies can help to achieve the cognitive support of operators.

Keywords—Industry 4.0, Operator 4.0, Edge computing, Human-System Interoperability, Indoor localization

Abstrakt—Aktuálna priemyselná evolúcia mení spôsob ako ľudia pracujú. S meniacimi sa interakciami prichádza taktiež potreba riešiť otázku začlenenie ľudí do tohto nového ekosystému. Počas posledných rokov sa objavili nové technológie a metódy, ktoré dokážu podporiť inteligentnú interoperabilitu medzi človekom a systémom. Nový koncept práce asistovanej strojmi si vyžaduje inteligentnú interakciu so zariadeniami, charakterizovanú ako fyzickú a kognitívnu interakciu. Začlenenie veľkého množstva senzorov do prostredia môže dopomôcť k pochopeniu takéhoto prostredia. Myslíme si, že výpočty na hrane siete spolu s rôznymi metódami lokalizácie v interiéri zohrá dôležitú úlohu. V tejto práci popisujeme interoperabilitu medzi človekom a systémom a taktiež niekoľko druhov interakcií. Ďalej sa sústredíme na výpočty na hrane siete spolu a taktiež na hodnotenie obmedzení sieťového prenosu. Predstavená je taktiež jedna z viacera prípadových štúdií, ktoré vznikli počas štúdia spomenutých tém. Všetky popísané štúdie môžu poskytovať operátorom kognitívnu podporu.

Kľúčové slová—Industry 4.0, Operator 4.0, Edge computing, Interoperabilita človeka a systému, Lokalizácia v interiéri

I. Introduction

Emerging technologies enable new types of interactions between humans and systems, interactions that are already transforming the humans' work environment. With developing interactions, there is also a need to address the issue of blending people into this new intelligent environment so that they can accommodate to these developments as efficiently as possible. The focus should be on treating computerization as a further enhancement of the human's physical, sensorial, and cognitive skills using Human Cyber-Physical System (H-CPS) integration.

One instance can be the modernizing of the Human-Machine Interface (HMI). Nevertheless, operators that work with these HMIs in the production sector are mainly older – mostly, they are not used to advanced technologies and devices. Also, a growing amount of devices are connected, and a lot of data needs to be represented in these HMIs. HMIs are becoming more and more difficult. Moreover, mobile accessories are substituting static HMIs panels that were directly attached to the machines.

One way how this problem can be resolved is by achieving cognitive HMIs whose intelligent design and presented information can be modified according to the actual requirements of the operators. To be able to achieve this, there is a need to precisely locate the operator in the newly created intelligent space. We should focus on the fusion of various technologies that could provide this location information accurately in the indoor environment. As a next step, the HMIs' intelligent and cognitive design based on the location of the operator could be achieved. We should also focus on emerging technologies like Augmented Reality (AR) because nowadays, it looks like the AR will be the new standard for HMIs in the new everchanging dynamic environments.

The paper is structured as follows: Section 2 summarizes the literature review and deals with topics like Industry 4.0, Human-System Interoperability, Intelligent space, Operator 4.0, Edge-enabled computing and also deals with theory behind indoor localization; Section 3 describes measurements made to asses the network connection needed to perform cloud control operations; Section 4 offer a closer look on Edge-enabled computing and also presents Edge-enabled computing solutions criteria selection and evaluation; Section 5 describes realized case studies; and Section 6 concludes this work.

II. Background

The literature review in this chapter summarises the gained findings. Chapter is divided into three sections relevant to our

work and deals with – Industry 4.0 [1] [2], Human-System Interoperability [1], Edge-enabled computing [3].

A. *Industry 4.0*

Industry 4.0 is a highly debated concept proposed to computerize further and automate production. The purpose is to recognize innovative industrial strategies and bring new system solutions to create the factory of the future. Due to the recent advancements in Information Technology (IT) paradigms, including the Internet of Things (IoT) [4], cloud computing [5], and cognitive computing [6], the Industry 4.0 [7] enables advantages such as high customization in mass production and completely automatized industry. The feasibilities for new advances affect all sectors of manufacturing, but mainly the following [1]: Smart products, Smart machines, Smart planners, Smart operator.

B. *Human-System Interoperability*

Human-machine Interoperability (interaction) emerges from the cooperation work to work supported by machines [8], which forms the concept of Operator 4.0. Accordingly, we need to embrace the role of human cooperation in real-time system communication.

Operator 4.0 is a moderately new term [8], which emerged from a resemblance to the industrial revolution concept – Industry 4.0. The new concept of work assisted by machines, which is characteristic for the Operator 4.0, demands "smart interaction" with machines principally characterized as physical and cognitive interactions [1][9].

1) Physical Interaction: The central purpose of physical interaction with machines is to eliminate the operator's restrained physical energy and provide him with mechanical strength. Two concepts introduced in [9] are explaining ideas for empowering operators with [1]:

- exoskeletons,
- collaborative robots,
- task shifting.

2) Cognitive Interaction: Smart cooperation depends on the cognitive abilities of the operator rather than his physical strength. Cognitive abilities extended by machines enhance a human's capability to perceive the surroundings. Wearable technology, additional sensors, or virtual reality equipment enables an operator to collect appropriate data, which cannot be obtained and processed in real-time otherwise [1]:

- AR and VR technology – tools and accessories processing digital knowledge from machines, decreasing human failures and the dependency on operator memory in real-time and intermediating a digitally replicated appearance with decreased risks,
- Intelligent personal assistants, and big data – incorporating AI within HMI with speech, language, image recognition software for task achievement in a human-like behavior and combining the human operator to participating smart devices information and machine learning techniques for decision-making assistance.

Both the physical and cognitive interaction capabilities embedded in machines and production operations are assisting human operators with health-related monitoring and tracking. Wearable trackers can measure human performance under stressful or demanding circumstances analyzing them and sending notifications if required [1].

Human assistance in production is, and will always prevail essential and irreplaceable. Although significant technological advance has already been made. Therefore, the concept of Balanced Automation Systems [10], having people at subservient positions to machine and process control, now evolves into a research question and technology development. This matter will set the human-machine Interoperability to find approaches to improve people's job satisfaction, product quality, and cost minimization at the same time [1].

3) Intelligent space: Cognitive operator support can be achieved by placing a large number of additional sensors in the environment to provide a better understanding of the entire space. This could create an Intelligent Space, Intelligent Environment, or a Digital Twin, of physical space that will describe the individual processes in such an environment. It can be created, for example, by thoroughly identifying and locating objects, people, and machines in space.

According to [11], Intelligent Environment are spaces with embedded systems and information and communication technologies that create interactive spaces that bring computing systems to the physical world and enhance the experience of the people there. An intelligent Environment is an area where technologies are used to improve and simplify everyday activities. One of the driving forces behind the emerging interest in a highly interactive environment is that computers are not only user-friendly but virtually invisible to the user [12].

Intelligent space describes the physical environment in which information and communication technologies and sensor systems are integrated into physical objects, infrastructures, and the environment. The goal is to enable natural communication between humans and the system. One reason for creating such an environment is that the systems and computers can be truly user-friendly and intuitive to use, without a need for computer skills background.

Intelligent environments are characterized by systems and technologies that are [13] [14]:

- built-in – many devices are integrated into the environment,
- context-aware – they can recognize objects, people and their context,
- personalized – can be adapted to different individual needs,
- adaptable – can be adapted to different conditions,
- foresight – they can anticipate and adapt to future situations.

Technologies enabling the creation of intelligent space may include hardware resources (cameras, RFID, BLE, WiFi, ...) or software resources (PDR, Fog / Edge / Cloud computing, UI methods). Since the creation of intelligent space requires,

in particular, precise location and identification of objects on the scene, the following technologies will be described later, mainly in terms of possible location and identification [15].

C. Edge-enabled Computing

The concept of sending all data over the Internet to cloud centers and the pay-as-you-go cloud computing principle has had a few problems. After all, all intelligence was centered in the cloud, and it was a centralized architecture. In recent years, ideas have arisen that have attempted to decentralize this architecture by reducing the amount of cloud-processed data. It has been achieved, in particular, by the increasing computational power of the end devices. Among the most well-known concepts that have tried to decentralize the cloud-based architecture are micro data centers, cloudlet, fog computing [16] [17] and edge computing [18] [19] [20] [21]. These new concepts have tried to solve problems that we can summarize as follows [22]:

- centralized architecture,
- security [23],
- various IoT devices were connected to the cloud,
- economic reasons.

The computational capabilities of the end devices are growing; thus, there is an opportunity to implement smart algorithms. Devices are constantly growing more advanced; so, they should help to:

- decrease network latency and the amount of data transmitted over the network,
- decrease the amount of data sent to the cloud,
- decrease cloud computing costs,
- enhance system management through better response times,
- enhance data security as part of data remains "home".

III. Network Evaluation

In this chapter, we aim to analyze the latency to some of the biggest cloud providers' services. Three cloud providers, i.e., Amazon Web Services, Google Cloud Platform, and Microsoft Azure, were chosen. Data centers located in west Europe (WestEU), east United States (EastUS), and eastern Australia (EastAU) were selected. The findings described in this chapter have already been published in a publication [2] where further measurements and verification can be found, so we will focus on a short description of this part of the research.

In order to evaluate the properties of the network, the theoretical and physical limitations need to be identified first. If two outermost sites in the world can be connected without intermediary network devices, the theoretical latency is 134ms, as calculated in (1). Two outermost places are $\frac{C}{2}$ apart (C is Earth's circumference), in fact, this distance is much bigger, due to pathways of the cables. Latency is the interval between sending and returning a message, so the packet must pass that distance two times ($2\frac{C}{2}$). c is the speed of light, and L is latency. The distance between east Slovakia and cloud providers' datacenters along with the theoretical minimum latency according to (1) can be seen in following TABLE I.

$$L = \frac{2\frac{C}{2}}{c} = \frac{40075km}{299792km/s} = 134ms \tag{1}$$

TABLE I
DISTANCE AND THEORETICAL MINIMAL LATENCY TO DATA CENTERS.

Data center	Distance	Theoretical min. latency
WestEU	1300km	9ms
EastUS	9000km	60ms
EastAU	15000km	100ms

The data center in Europe symbolizes the closest data center. The data center in the United States symbolizes intercontinental communication. The data center in Australia symbolizes communication with the outermost data center. In the first phase, many measurements were made at varying times of the day and for every day of the week throughout February 2019 from Technical University of Kosice, Slovakia, to compare different providers datacenters' connection speed. Results of the mean latency and the 10th percentile (90% of the values are better) are summarized in TABLE II [2].

TABLE II
LATENCY BETWEEN EAST SLOVAKIA AND CLOUD PROVIDERS' DATA CENTERS [2].

Provider	Location	Mean latency	10th perc.
Amazon Web Services	WestEU	64.33ms	65ms
	EastUS	145.67ms	146ms
	EastAU	545.52ms	740ms
Google Cloud Platform	WestEU	37.58ms	38ms
	EastUS	111.92ms	113ms
	EastAU	**301.25ms**	302ms
Microsoft Azure	WestEU	**30.51ms**	31ms
	EastUS	**106.07ms**	107ms
	EastAU	316.54ms	339ms

Since Microsoft Azure provided the lowest latency to multiple locations, it was chosen for thorough testing. In the second phase, we have selected four communication protocols suitable for IoT solutions; namely HTTP, WCF, OPC-UA, AMQP. Again, multiple measurements were made. We found out that the time of the day, and also the day of the week, have a negligible influence on the latency. Therefore, the table is not divided by the time of the measurement. The experimental results were summarized and analyzed. The minimum and maximum latency, the median, mean, mode, and standard deviation for each protocol have been summarized and are thoroughly described in [2].

Cloud requirements as well as latency and data amount requirements for the cloud control have been analyzed. Experimental results showed that control is possible in many solutions. Technological processes that do not require latency or control period lower than 1s can be managed through cloud services without dependence on the type of transmission service or the data centers' position. Processes that need latency or control period lower than 40ms cannot be controlled directly from cloud services. Technological processes which require times within 1s and 40ms are reliant on the transfer

protocol and position of the cloud data center, or the speed of the transmission channel between physical processes and data centers' location. The latency is subject to transmission channel usage, not just the distance between corresponding joints; sometimes, the faster channel could be a more distant one. That suggests that not everything can be controlled through cloud services, and edge computing can make a bridge between local and cloud services. The intelligence on the edge of the network will play a significant role. The computing capacities of end devices are increasing; fog/edge computing will shift intelligence closer to systems which will reduce the amount of data that need to be transferred as well as shorten the latency [2].

IV. Edge AI

This chapter discusses a selection of fundamental criteria for edge-enabled IoT gateway, which should be implemented to any edge-enabled solution. From the scientific work carried out and the analysis of scientific articles, a set of 14 criteria for edge-enabled architecture was chosen. These criteria fell into four main categories (device connectivity, data pre-processing, data analytics, special hardware requirements). For a better understanding, we ranked these criteria as follows [3]:

- Device connectivity
  - Network standardization and protocols [24] [25] [26] [27] [28] – services that offer tools to standardize the various protocols coming from connected devices into protocols that are supported and sent to the cloud. This can contain data formatting and message brokers.
  - Configuration of the connected devices [24] [25] [26] [27] – tools for organizing connected devices to recognized and user-defined groups. This should have a better visibility over the monitored system.
- Data pre-processing
  - Data filtering [26] [29] [30] [27] – tools providing different algorithms that allow different types of filtering to be applied to incoming data without losing the information value of the data.
  - Data consolidation [24] [25] [26] [29] [31] [27] – tools providing inbound data aggregation based by topic. This can contain data fusion and grouping.
- Data analysis
  - Data pre-analytics [25] [26] [30] [27] [32] – tools that provide machine learning algorithms for flowing data analysis and pattern search in these data. It can be a similar principle to predictive maintenance.
  - Local database [24] [25] [30] [27] – local storage of processed data. This repository can serve as a source database for a local machine learning or a backup in case of a lost connection, after the resume, the stored data is sent to the cloud.
  - Notifications [25] [26] [30] [27] – tools for publishing quick notifications to end devices immediately after processing at IoT gateway.

- Data compression [33] [26] [34] [30] – tools that allow data to be partially saved and compressed by a suitable method after collecting enough data.
- Data encryption [33] [25] [31] [30] – tools that offer features to further encryption of data.
- Data watermarking [33] [25] – sent data will carry a watermark, which guarantees a clear identification of the source.

- Special hardware requirements
  - High computation power [28] – powerful processor for tasks that requires much computational power.
  - Long-term battery power supply [28] – battery power supply for long-term operation without the need of an external power supply.
  - Backup internet connection [27] [35] – in the event of the primary connection fails, it will provide an opportunity to be further connected to the Internet. In case of long-term failure of the primary connection, send an emergency report.
  - Backup power supply [27] [35] – an extra battery that will also keep the device functioning when it has been disconnected from the main power supply.

This set of criteria has been evaluated by the experts regarding the usefulness and complexity of the implementation. The individual criteria were evaluated on a scale of 1 to 5. The usefulness should take into account the overall impact of the solution on functionality. Difficulty should take into account the implementation difficulty, the estimated number of implementation hours, and the financial difficulty of running the service, as well as the impact on the final cost of the IoT gateway.

Standardized weights were calculated. We have evaluated each criterion from a usefulness point of view throughout the IoT solution and the implementation-financial difficulty. Difficulty takes into account the implementation difficulty, the estimated number of implementation hours, and the financial difficulty of running the service, as well as the impact on the final cost of the IoT gateway. From these evaluations, we have calculated normalized weights according to the defined formulas, and from these weights, the importance of the given criterion was determined.

In TABLE III, we have sorted the criteria by importance. However, we do not include the criteria with less than zero normalized weight because these criteria have not fulfilled the condition.

In this way, a system of recommendations was created for creating edge-enabled IoT gateway by experts. This system captures expert thinking when designing edge-enabled IoT gateway. It should help systematically design the right gateway. These 12 criteria could be implemented to any edge-enabled IoT gateway. There is a possibility that these criteria will not be best for the edge-enabled gateway that needs to be used in environments with special conditions. For example, the long-lasting battery may be needed. In that case, the system can be reevaluated with the help of the additional experts from

TABLE III
IMPORTANCE OF CRITERIA

Criterion	Importance
Data filtering	1
Local notifications	2
Data pre-analytics	3
Data compression	4
Local storage	5
Network standardization and protocols	6
Data consolidation	7
Backup internet connection	8
Configuration of the connected devices	9
Data encryption	10
Data watermarking	11
Backup power supply	12

that field and some criteria may change, or criteria can be added or removed.

Using the mentioned criteria, we determined the weights of each phase during IoT gateway data processing. The individual weights represent the advancement of the given phase concerning the concept of edge-enabled devices. Based on the rating defined by each data processing stage, we have produced a ranking that shows the order in which the phases can be performed. These criteria are hierarchical. After defining all IoT gateway data flow functions, we calculate the cumulative scores we get by counting the weights as shown in the formula (2). Where s_e is the final score of edge-enabled IoT gateway which is the sum of the ε_i criteria and m is a count of fulfilled criteria.

$$s_e = \sum_{i=1}^{m} \varepsilon_i \qquad (2)$$

In this case, the solution that achieved a higher score is technologically more advanced. Thus, the number of functions/criteria is the maximum theoretical score of 33. Based on this cumulative score for IoT gateway, we have defined four categories of IoT gateway as follows: *normal IoT gateway*, *smart IoT gateway*, *intelligent IoT gateway*, and *edge-enabled IoT gateway*.

V. CASE STUDY

In this case study, we focus on augmented reality human-robot interoperability (AR-HRI). The case study is thoroughly described in the conference paper [36].

We aim to create an environment where AGV have planned their path, and people can see their paths, and mark restricted areas by virtual means if needed, for these mobile robots and humans to move and stand safely around a smart warehouse. We have proposed an AR environment for people, where they can see an AGV planned path, and they can add virtual obstacles and walls to the mobile robots' view. These virtual obstacles and walls can be used to determine restricted areas for mobile robots, which can be seen, for example, as safe areas for humans' and/or robots' stationary work [36].

The AR-HRI system architecture design consists of three separate, but interrelated applications: (i) a Core app installed in a PC, and responsible for creating the AR view for the humans and for the mobile robots; (ii) the AR app installed in a mobile phone (or smart-glasses), allowing the humans to view any virtual object added in the smart cyber-physical space in their view of the real-world; and (iii) the Robot app installed in an AGV or drone, which uses computer vision to also provide the mobile robots to have in this case a cyber-physical view of the environment as they move around it (see Fig. 1). AR-HRI system have a goal to provide humans and mobile robots with safe dynamic routes for movement and delivery products at the smart environment, avoiding any collision by means of AR visual controls (i.e. virtual paths and virtual walls) [36].
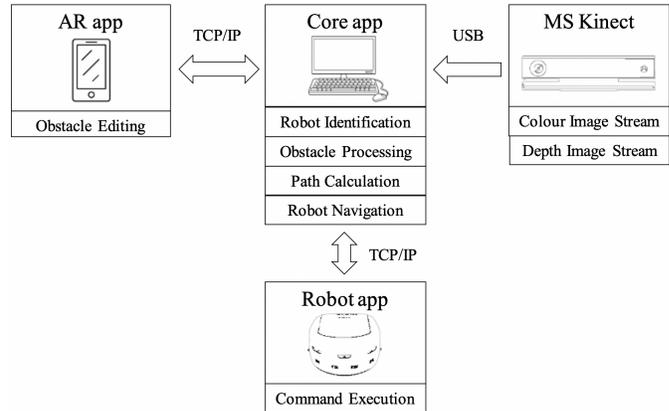

Fig. 1. Architecture.

For developing and testing the proposed AR-HRI system architecture, the following software and hardware elements were considered for the laboratory platform development: (i) for the Core app, C# programming language was used for the desktop application development, and a desktop computer with the following characteristics: 4-cores Intel i7-4710HQ CPU with base frequency at 2.50GHz, 16GB of RAM memory, GeForce GTX970M GPU, Windows 10 Pro OS; (ii) for the AR application, Unity development platform was chosen due to its multiplatform support, which allow the possibly to extend the solution to other AR platforms such as Microsoft HoloLens, and a smart mobile phone with the following characteristics: 6-cores Snapdragon 650 w/ frequency at 1.80GHz, 3GB of RAM memory, Adreno 510 GPU, Android 6.0 OS; and (iii) for the Robot app C programming language was used, and a Khepera III AGV, which runs an embedded Linux OS, and offers all standard C libraries for its programming Fig. 1. The test application can be seen in Fig. 2. In this figure, we can see the AGV identification, physical and virtual obstacles, and also planned path.

Thanks to the AR app developed, the user can add and see virtual walls in the cyber-physical environment by simply drawing on the mobile phone screen. Furthermore, we have tested different placement, size, and shapes of the virtual obstacles. Fig. 3 shows an example scene. The three blue barriers are virtual obstacles.

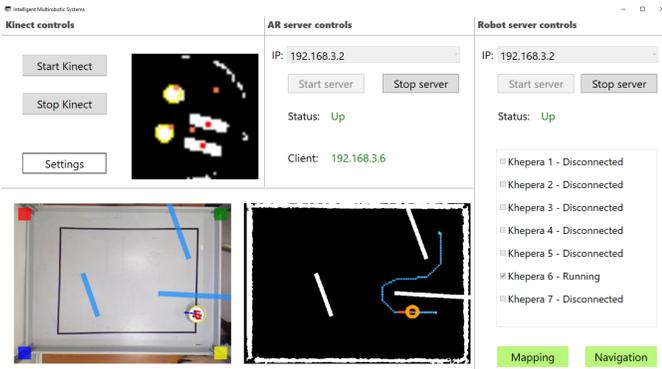Although not yet widely adopted, AGVs and drones will
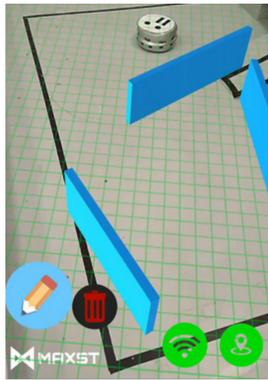
Fig. 2. The Core App interface.

Fig. 3. The AR App interface.

transform the way goods are transported. Nevertheless, not all smart environments will become "unmanned", some may still require humans. Thus, AGVs and people collaboration for their optimal performance is needed. So it is in these cases where our AR solution for humans-robots interoperability aims to contribute.

VI. Contribution

During Ph.D. study we worked on an Edge-enabled approach to Human-System Interoperability. The motivation was to asses technologies and techniques that can be used to provide cognitive support for the operators in different fields. Some of the contributions are presented in this paper, others are thoroughly described in dissertation and cited publications.

The main scientific contribution is as follows:

- Intelligent Human-System Interoperability was analyzed by the means of physical and cognitive interaction and is presented.
- Cloud and edge computing were analyzed, and architectures were defined and presented. Four different BLE indoor positioning architectures that use principles of edge computing and cloud computing were defined.
- To asses edge solutions and to help with designing these solutions, the evaluation criteria were defined and are presented..

The main technological contribution is as follows:

- Technologies and techniques that can provide cognitive support for operators were described. Beforehand, we needed to asses the network parameters, so we understand the limits of the cloud solutions. In further research, this will help with the optimization of the distribution of computing power and capabilities. Network evaluation study was conducted and is described.
- Several case studies were conducted that uses Edge-enabled computing, we have presented five of them and are as follows: indoor localization, AR human-robot interoperability, cognitive healthcare platform, distribution of NN models, and security in intelligent space.
- I have been principal investigator in two projects, which supported the acquisition of a large number of intelligent sensors and actuators and their subsequent use in applications using AI methods, especially ML. Students, teachers, and researchers have full access to this – ThinkLab – at Technical University of Kosice.

VII. Conclusion

Throughout past years the technologies and methods that can support intelligent Human-System Interoperability forced by the present industrial revolution were introduced. With developing interactions, there is also a need to address the issue of blending people into this new intelligent environment so that they can accommodate to these developments as efficiently as possible. Architectures that depend solely on cloud technologies have a few problems, like the increasing amount of the transferred data and higher latency. Numerous ideas have arisen that have attempted to overcome these problems. Among the most well-known concepts is edge computing.

Human-machine Interoperability emerges from the cooperation work to work supported by machines, which forms the concept of Operator 4.0. Accordingly, we need to embrace the role of human cooperation in real-time system communication. The new concept of work assisted by machines, which is characteristic for the Operator 4.0, demands "smart interaction" with machines principally characterized as physical and cognitive interactions. Cognitive operator support can be achieved by placing a large number of additional sensors in the environment to provide a better understanding of the entire space. This should create an Intelligent Space, Intelligent Environment, or a Digital Twin, of physical space that will describe the individual processes in such an environment. It can be created, for example, by thoroughly identifying and locating objects, people, and machines in environment.

The motivation was to asses technologies and techniques that can be used to provide cognitive support for the operators in different fields. We think that the leading role will have edge computing along with varying methods of indoor localization. We have described Human-System Interoperability and various types of interactions. After that, we have focused on edge computing along with the network evaluation study. One of the several case studies that arise from the ideas was presented. These studies can help to achieve the cognitive

support of operators with the goal of intelligent human-system interoperability.

Further research should focus on the comparison of the fusion of different indoor localization technologies and techniques to provide faster, more accurate, and efficient ways of indoor localization. The UWB and Bluetooth 5.1 should also be implemented into studies since UWB starts to appears in mobile devices, and Bluetooth 5.1 offers builtin technology to provide better localization. Further research can also focus on implementing all of the mentioned in one application that can support operators' work on a large scale.

REFERENCES

[1] Iveta Zolotova et al. "Smart and cognitive solutions for Operator 4.0: Laboratory H-CPPS case studies". In: *Computers & Industrial Engineering* (2018).

[2] Erik Kajati et al. "Cloud based cyber-physical systems: Network evaluation study". In: *Advanced Engineering Informatics* 42 (2019), p. 100988.

[3] Peter Papcun et al. "Edge-enabled IoT gateway criteria selection and evaluation". In: *Concurrency and Computation: Practice and Experience* (2019), e5219.

[4] Jayavardhana Gubbi et al. "Internet of Things (IoT): A vision, architectural elements, and future directions". In: *Future generation computer systems* 29.7 (2013), pp. 1645–1660.

[5] Qusay Hassan. "Demystifying cloud computing". In: *The Journal of Defense Software Engineering* 1 (2011), pp. 16–21.

[6] John E Kelly III and Steve Hamm. *Smart machines: IBM's Watson and the era of cognitive computing*. Columbia University Press, 2013.

[7] Henning Kagermann et al. *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group*. Forschungsunion, 2013.

[8] David Romero et al. "Towards a human-centred reference architecture for next generation balanced automation systems: human-automation symbiosis". In: *IFIP International Conference on Advances in Production Management Systems*. Springer. 2015, pp. 556–566.

[9] David Romero et al. "Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies". In: *International conference on computers and industrial engineering (CIE46) proceedings*. 2016.

[10] Stamatis Karnouskos et al. "The IMC-AESOP architecture for cloud-based industrial cyber-physical systems". In: *Industrial Cloud-Based Cyber-Physical Systems*. Springer, 2014, pp. 49–88.

[11] Juan C Augusto et al. "Intelligent environments: a manifesto". In: *Human-centric Computing and Information Sciences* 3.1 (2013), pp. 1–18.

[12] Alan Steventon and Steve Wright. *Intelligent spaces: The application of pervasive ICT*. Springer Science & Business Media, 2010.

[13] Quan Z Sheng, Wei Emma Zhang, and Elhadi Shakshuki. *Practices and applications in ambient and intelligent information systems*. 2017.

[14] Zoe Falomir et al. "Thematic issue on human-centred ambient intelligence: Cognitive approaches, reasoning and learning". In: *Journal of Ambient Intelligence and Smart Environments* 9.5 (2017), pp. 539–542.

[15] Erik Kajati. "Interoperabilita cloveka a systemu - inteligentny priestor, bezkontaktne technologie, servisne architektury". Dissertation proposal. TUKE, 2018.

[16] Flavio Bonomi et al. "Fog computing and its role in the internet of things". In: *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM. 2012, pp. 13–16.

[17] Yannis Nikoloudakis et al. "A fog-based emergency system for smart enhanced living environments". In: *IEEE Cloud Computing* 6 (2016), pp. 54–62.

[18] Weisong Shi et al. "Edge computing: Vision and challenges". In: *IEEE Internet of Things Journal* 3.5 (2016), pp. 637–646.

[19] Weisong Shi and Schahram Dustdar. "The promise of edge computing". In: *Computer* 49.5 (2016), pp. 78–81.

[20] Mahadev Satyanarayanan. "The emergence of edge computing". In: *Computer* 50.1 (2017), pp. 30–39.

[21] Rodrigo Roman, Javier Lopez, and Masahiro Mambo. "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges". In: *Future Generation Computer Systems* 78 (2018), pp. 680–698.

[22] Maryam M Najafabadi et al. "Deep learning applications and challenges in big data analytics". In: *Journal of Big Data* 2.1 (2015), p. 1.

[23] Eva Chovancova et al. "Securing distributed computer systems using an advanced sophisticated hybrid honeypot technology". In: *Computing and Informatics* 36.1 (2017), pp. 113–139.

[24] Eun-Young Jung et al. "Home health gateway based healthcare services through Uhealth platform". In: *Wireless Personal Communication* 73.2 (2013), pp. 207–218.

[25] Tuan Nguyen Gia et al. "Low-cost fog-assisted healthcare IoT system with energy-efficient sensor nodes". In: *13th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE. 2017, pp. 1765–1770.

[26] Behailu Negash et al. "Leveraging fog computing for healthcare iot". In: *Fog Computing in the Internet of Things*. Springer, 2018, pp. 145–169.

[27] Tomas Lojka, Martin Miskuf, and Iveta Zolotova. "Industrial IoT gateway with machine learning for smart manufacturing". In: *IFIP International Conference on Advances in Production Management Systems*. Springer. 2016, pp. 759–766.

[28] Jozef Mocnej et al. "Impact of Edge Computing Paradigm on Energy Consumption in IoT". In: *IFAC-PapersOnLine* 51.6 (2018), pp. 162–167.

[29] Suryadip Chakraborty et al. "Fog networks in healthcare application". In: *13th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE. 2016, pp. 386–387.

[30] Amir M Rahmani et al. "Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach". In: *Future Generation Computer Systems* 78 (2018), pp. 641–658.

[31] Mahmood Ahmad et al. "Health Fog: a novel framework for health and wellness applications". In: *The Journal of Supercomputing* 72.10 (2016), pp. 1–18.

[32] Adrian Pekar et al. "Adaptive aggregation of flow records". In: *Computing & Informatics* 37.1 (2018).

[33] M Shamim Hossain and Ghulam Muhammad. "Cloud-assisted industrial internet of things (iiot)–enabled framework for health monitoring". In: *Computer Networks* 101 (2016), pp. 192–202.

[34] Harishchandra Dubey et al. "Fog data: Enhancing telehealth big data through fog computing". In: *Proceedings of the ASE BigData & SocialInformatics 2015*. ACM. 2015, p. 14.

[35] Martin Miskuf. "Edge-enabled platform for monitoring quality of healthcare in Home Environment". In: *Proceedings from 18th Scientific Conference of Young Researchers*. SCYR 2018 TUKE. 2017, pp. 92–93.

[36] Peter Papcun et al. "Augmented Reality for Humans-Robots Interaction in Dynamic Slotting "Chaotic Storage" Smart Warehouses". In: *IFIP International Conference on Advances in Production Management Systems*. Springer. 2019, pp. 633–641.

Processing of Digital Images Obtained by Observing the Solar Corona

Matúš Kozák

Department of electronics and multimedia telecommunications
Faculty of Electrical Engineering and Informatics
042 00 Košice, Slovakia
matus.kozak@student.tuke.sk

Dušan Kocur

Department of electronics and multimedia telecommunications
Faculty of Electrical Engineering and Informatics
042 00 Košice, Slovakia
dusan.kocur@tuke.sk

Abstract - **Photographing the solar corona is one of the most complicated tasks of astrophotography. Preparing for image capture is an important part of producing a quality image of the solar corona. This consists of the preparation of the observation system, image calibration and additional image enhancements. The thesis contains an overview of the observed object and the way of capturing the image of the solar corona from the Earth. It discusses the possibilities of improving results of observation image and, if necessary, the calibration of images. The result of this thesis is a method that can visualize the fine structures of the solar corona in the processing of a single image. The advantage of the method is its application to a single image, which does not have to be calibrated or combined with other images from the observation process. The method uses adaptive histogram balancing and the goal is a simple calculation for the fastest possible processing and creation of a preview. The results of this method are presented in several images at different exposure times.**

Abstrakt - **Fotografovanie slnečnej koróny patrí medzi najzložitejšie úlohy astrofotografie. Dôležitou časťou vyhotovenia kvalitného snímku slnečnej koróny je príprava snímku. Tá pozostáva z prípravy pozorovacej sústavy, kalibrácie snímku a dodatočných úprav snímku. Práca obsahuje prehľad o prehľad o pozorovanom objekte a o spôsoboch akými sa dá vyhotoviť obraz slnečnej koróny zo zemského povrchu. Rozoberá možnosti vylepšenia výsledného obrazu a kalibráciu obrazov pre ich ďalšiu úpravu. Výsledkom tejto práce je metóda, ktorá umožňuje zviditeľniť jemné štruktúry slnečnej koróny v spracovaní jedného obrazu. Výhoda tejto metódy spočíva v jej aplikovateľnosti s dobrými výsledkami aj na jeden snímok bez nutnosti jeho kalibrácie alebo kombinácie s inými obrazmi s pozorovania. Metóda využíva adaptívne vyváženie histogramu a zohľadňuje výpočtovú nenáročnosť pre čo najrýchlejšie spracovanie a poskytnutie náhľadu aktuálneho snímku. Výsledky tejto metódy sú prezentované na viacerých snímkach pri rôznych expozičných dobách.**

Keywords— solar corona, observation, histogram, adaptive filters, image processing, histogram equalization

I. INTRODUCTION

The image contains a lot of information that can be interpreted differently. The interpretation depends on many factors. The image can be used for teaching, simple transmission of information or it can be a work of art. It all depends on the purpose for which the image was created.

Image processing is already part of modern electronics. This requires the development of new and faster image processing methods. It is used in mobile devices that offer many filters in addition to basic image capture. Software is required to improve hardware or human imperfections. Image quality can be affected in two ways, by improving the hardware or the software. Hardware modification is limited by electronics, mechanics, optics and, finally, lighting conditions when creating an image. The second way to enhance an image is by software editing. Unlike solid-state hardware, software modifications can be modified at any time.

Solar corona research has a long tradition and is still relevant. Observation of the solar corona is possible during natural solar eclipses or with the use of specifically designed telescopes. In both cases, the quality of the observation and thus the production of the images depends on the most accurate setting of the observation technique.

The processing of solar corona images is specific due to the properties of these images. There are overexposed areas near the sun center and darker areas at greater distances. These contain essential information for the study of the corona.

Several methods are currently used to process solar corona images. The basic method is a combination of images with different exposure times, i.e. simply a combination of brighter and darker images. The emphasis on fine structures in the dark areas of the image resulted in the creation of filters focused primarily on these areas, such as the tangential filter. This filter produces a deficiency, which is the neglect of the area parallel to the edge of the sun. On the contrary, it highlights radially oriented structures in the solar corona. This raised investigation of methods that work with the selected area of the corona separately in order to preserve and highlight the original information of the image. These methods can be used on data sets created during observation. They often require calibration images, such as flat field and dark current. All this increases the complexity of the calculation and requires additional scanning of calibration images.

The analysis of current methods for processing solar corona images raises two issues that are discussed in this work. The first is to create a relevant view of a single image of the solar

corona, so it is not necessary to combine it with other images from the observation or use calibration. The second issue is the creation of the most accurate view with minimal or no modification of the fine structures of the solar corona.

The aim of the work is to create an immediate preview of the image from the observation, which is useful in the analysis of the selected observation program directly during the observation. Thus, it can immediately evaluate the relevance of the use of the capacity of the telescope or the region of interest of observation. Because it is a fast and efficient method, it can be applied directly to observation software.

The inspiration for the creation of the Fast corona image preview method with histogram equalization was the analysis of the current state of the problem, especially the method of obtaining images of the solar corona and the specifics of these images. The biggest contribution was the knowledge from the High-dynamic-range processing method and the Noise adaptive fuzzy equalization method. The idea is that the overexposed pixels in the image are taken separately during processing, thus eliminating their undesired effect on the processed part of the image during histogram equalization. The resulting image is thus a combination of several images with a separate histogram equalization. Finally, a smoothing filter is applied.

II. DATA SOURCE AND OBJECT OF THE OBSERVATION

A. The Sun

The Sun has strong magnetic field, which differs across the surface. In the place which has bigger concentration of magnetic fields than his surroundings are observable darker spots, the sunspots. Although the solar magnetic fields give rise to many different solar phenomena, it is the sunspots which provide the best detailed information about the behavior of the solar magnetic fields [1].

The Sun's magnetic field invokes many effects called solar activity. The solar activity has major role in the formation of Solar system. The solar magnetic fields play fundamental role in determining the characteristics of the plasmas that make up the corona and the wind [2]. Solar activity also causes geomagnetic storms, which can lead to communications disruptions, electric power network problems, satellite orbit shifts and, sometimes, satellite failure. Predicting storm conditions requires understanding of solar magnetism and its fluctuations [3].

B. Solar corona

As mentioned above, the Sun has layers and whatever comes from the Sun to the Earth, goes through the atmosphere of the Sun. Atmosphere surrounds the Sun and is composed of photosphere, chromosphere, the transition region and the corona. The chromosphere is the thin layer above the surface of the Sun, where dominates spectrum of emission and absorption lines. The temperature here raises from 4 100 K to almost 20 000 K. Higher is corona, the plasma that surrounds the Sun and is extended to millions of kilometers into outer space. It has extremely low density and the temperature here increases over 1 000 000 K. This temperatures increase is found to occur very

short distance, with raising from hundred thousand to millions K within 500km [4].

The Sun has another influence on the Earth, which is called space weather. This is responsible for perturbations and affects communications, space flight and power transmission. To understand how the solar corona and the solar wind (i.e., the steady supersonic outflow of ionized gas from the corona) are produced and maintained, one must have detailed empirical knowledge about the properties of the plasma [5].

Observing of the solar corona could be done from the Earth as well as from the space. There are limitations for the solar observations, which have their origin mostly in the Earth´s atmosphere, but sometimes in the instrumentation itself [6]. Non - uniform brightness compared to the disk of the Sun and Earth atmosphere conditions causes specifics ways for observations. Then the observed data contain many influences which are necessary to modify with hardware of software processing.

C. Observation devices

The solar corona can be visible only during the solar eclipse. The eclipse occurs when the disk of the Sun is covered by the moon and observer on the earth is in this shadow. Of course, the coverage of solar disk by the moon can be total or partial, therefore we distinguish following types of solar eclipse: a total eclipse, an annular eclipse (size of the moon is smaller than solar disk), a hybrid eclipse (combination both previous) and partial solar eclipse.

The total solar eclipse is not common phenomena, and it does not occur on every place on the earth. It occurs every 18 months in average, however partial solar eclipses are more frequent [7]. Duration of the total solar eclipse does not take more than 7 minutes. These limitations have leaded to discover new telescope or technique of the observations.

The telescope provides both a source of new information about the universe and means of confirming theories that explain previous discoveries [8]. The telescope that was first used for this observation is called coronagraph, it was discovered by B. Lyot early in 1930's, with a special grouping of the lenses and using artificial moon in the telescope tube. The optical scheme of the coronagraph is shown in Figure 1. The inclination of the artificial moon is designed to reflect light incident from the solar disk.
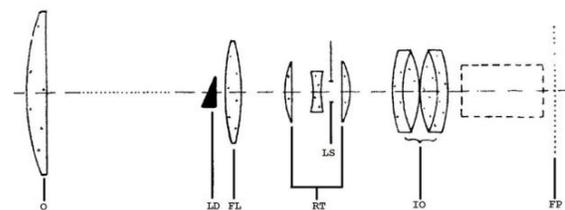

Fig. 1.   Figure 1 The optical scheme of the Zeiss-200/3000 coronagraph. 4 O: Objective lens. LD: Lyot diaphragm/artificial moon. FL: Field lens. RT: Reimaging triplet with Lyot's Stop (LS). IO: Imaging objective lens. FP: Focal plane. [Downloaded from https://www.researchgate.net/]

In present observations, modern telescopes, thanks to appropriate optics (beamsplitters), can acquire image of the Sun by more cameras at once. For example, at the Themis at the Observatorio del Teide, a three - like quiescent prominence was selected as a target for a double-beam spectropolarimetry in the He i D3 line [9]. By using optical filters and multiple cameras as a detector, creation of the modern telescope is possible.

It is well known that one of few coronagraphs is located in Slovakia, at the Lomnicky peak. The resolution of the Earth based telescopes is insufficient by the atmosphere [10]. Hence, the limits of the location are hidden in relatively low attitude, despite being the highest peak in Slovakia. The coronagraph is equipped with post focus instrument called CoMP-S (Coronal Multipolarimeter for Slovakia), which is able to measure the linear polarization, circular polarization and Doppler shift of spectral lines. A schematic of the single detector CoMP model is shown in Figure 2. Such measurements will be used for computation of the plane of sight magnetic field direction, the line of sight magnetic field strength and the line of sight velocity, respectively [11].

Fig. 2. Solid model rendering of the CoMP instrument package. Light enters through the blue tube on the left [12].

III. DATA RPEPROCESSING

The base preprocessing of the image data consists of removing noise produced by the sensor, bad pixels caused by manufacture and by the removing of any artefacts created by camera's optic. Any of these images is created for the process called calibration of observed images and consists of the dark current, flat fields and bias. Image calibration is optional but increases quality of final image.

A. Image calibration

In the processing of astronomical images, the calibration or preprocessing is usually done. It consists of elimination of all known influences of optics, telescope and camera described above as flat field, bias and dark frames. The basic image correction is given by following formula:

$$R = (P - G) / ((F - G)/A) \qquad (1)$$

where R is the resulting reduced spectrum (image), P is the observed particular raw spectrum, F is the flat-field source frame, A is the average pixel value of the corrected flat-field source frame used for normalization, i.e. to scale the flat-field source frame to a mean intensity of 1.0 and

$$G = D + B + L \qquad (2)$$

is the background composed of dark current D, bias B and scattered light L[13]. Background is not constant during the observation due to the scattered light and the dark current, that depend on the optic system, respectively on the temperature or on the exposure time.

B. Grayscale conversion

In this work, the input image is a three-dimensional matrix with a resolution of 4928 x 7380 x 3. A grayscale image is tied for processing. The transformation uses constants that express the ratio of each color component in the gray level component. There are several types of conversions with different constants for RGB components, for this paper there are as follows [14]:

- red = 0.2989

- green = 0.5870

- blue = 0.1140

For faster processing, it is possible to exclude a useful part of the image, based on the size of the image, i.e. to crop the image based on the position of the center of the solar disk. This can only be done with sufficiently large images where it is possible to lose a part of the image. If multiple images need to be processed, it is convenient to set the cropping position automatically.

IV. METHOD

The aim of this paper is proposing, implementing and testing a method suitable for fast processing of the images obtained by observing the solar corona. The result is the method called Fast Corona Image Preview with Histogram Equalization (shortly FCIPHE). This method is based on the enhancement of the fine structures of the images from the observation of solar eclipse. It uses a histogram equalization to adjust image contrast and in addition, the method is extended by adaptive histogram equalization that enhances the fine structures in the images. These structures on the image background are suppressed by sunlight what is a dominant source of light in the image. Usually they are located at a greater distance from the center of the sun on the image. An imbalance in brightness is typical of solar corona images. The aim of this work is to make these fine elements in the image visible to the naked eye.

The method is performed on a single input image. The steps of the computation of the method are following.

1) A set of images.

A set of images is created, based on the omission of the selected value of pixels in the image. This will reduce the effect of overexposed or hot pixels on the resulting image. Each image from the set is processed separately. An image in a set of images is created, based on omitting a selected group of

pixels according to the pixel level. The level of pixels that are omitted corresponds to the exponential function. Histogram analysis shows that most pixels belong to the dark part of the image and that there are areas where the pixels are overexposed. The created set of images allows separate processing, where unwanted pixels are eliminated.

The prerequisite is an input image with a defined shutter speed saved in NEF format. The image consists of a three-dimensional matrix, each matrix is for one color component. The image is transformed into a grayscale.

A set of N image $u(x, y)$ is formed from the input image based on the pixel level value.

$$u(x,y) = \{u_1(x,y),\ u_2(x,y),\ ...,\ u_i(x,y),\ ..\ u_N(x,y)\} \quad (3)$$

Where N is the number of separate images created and depends on the bit depth. The maximum number of created images can be equal to the gray levels L in the image, $N = \langle 0, L_{max} \rangle$.

The occurrence of overexposed pixels is visible from the histogram of the image. Based on this, a set of independently processed images is created according to the exponential function. The values inside the separately processed image are created, based on the condition

$$u_i(x, y) = \begin{cases} u(x,y) & if\ u(x,y) < L_{max} - 2^{2^m} \\ 0 \end{cases} \quad (4)$$

where $m = \{0,1,2, ... k\}$, k refers bits per pixel.

2) Histogram equalization.
Histogram equalization (HE) is performed on all images in the set of images. A set of images is created from the original input image. Each image in the set is processed separately by adjusting the histogram. The equation (4) condition ensures that the equalization histogram includes each pixel of the original image.

HE is calculated based on cumulative distribution function (CDF) corresponding to p_u as

$$cdf_u(i) = \sum_{j=0}^{i} p_u(u = j) \quad (5)$$

Transformation of the form $v = T(u)$ to produce a new image $\{v\}$, with a flat histogram have a linearized CDF across the value range, i.e.

$$cdf_v(i) = iK \quad (6)$$

for some constant K. The properties of the CDF allow us to perform such a transform:

$$cdf_v(y') - cdf_v(T(k)) = cdf_u(k) \quad (7)$$

here k is in the range $[0, L]$. Transformation T maps the levels into the range $[0,1]$, since a normalized histogram of $\{u\}$ is used.

3) Rank equalization.
An adaptive histogram (AHE) is then calculated from the set of images.

After this processing, an AHE is performed. The AHE algorithm is as follows:

- Calculate a grid size based on the maximum dimension of the image. The minimum grid size is 32 pixels square.

- If a window size is not specified chose the grid size as the default window size.

- Identify grid points on the image, starting from top-left corner. Each grid point is separated by grid size pixels

- For each grid point calculate the CDF of the region around it, having an area equal to a window size and centered at the grid point.

- After calculating the mappings for each grid point, repeat steps 6 to 8 for each pixel in the input image.

- For each pixel find the four closest neighboring grid points that surround the pixel.

- Using the intensity value of the pixel as an index, find its mapping at the four grid points based on their cumulative distribution functions.

- Interpolate among these values to get the mapping at the current pixel location. Map this intensity to the range *[min : max]* and put it in the output image.

4) Image composition
The composition of the final image is formed in the next step. This is created by multiplying each image in the set by the original image and the result is the sum of the image. It can be written as follows:

$$g(x, y) = \frac{1}{N} \sum_{i=1}^{N} h_i(x, y) . f(x, y) \quad (8)$$

Where $f(x, y)$ is the input image, $h(x, y)$ is the image after adjusting the histogram from the set according to equation (3). The original image includes all the information, both useful and useless. The described processing means that the pixels of the image undergo certain changes and contain additional information.

Adjusting the contrast also means that useful pixels should be more visible, and the pixels with less informative value should be darker. By multiplying each image of the set created, based on the mask with the processing of a separate image, and then adding them together, an almost final image is created. This is then further processed.

5) Smoothing filter.
The image in this section has its useful pixels and the contrast adjusted and it is ready for interpretation. As a result of the image visualization, there is still a useful way to modify it by applying image smoothing. The smoothing filter was selected from several available filters that interpret a multidimensional Gaussian filter. The multidimensional filter is implemented as a sequence of one-dimensional convolution filters.

V. Results

A series of these steps creates the final image. Depending on the size of the processed data, this method can be applied to the image section or to the entire image. The described processing is performed on a single sample image, which includes the specifics of the image - too bright areas and overexposed pixels, on the other hand also dark, slightly visible to the naked eye, parts of the image. The results of the method applied to images with different exposure times are shown in Figures 5, 6 and 7.

A. Creating set of images

The properties of the observation images of solar corona show that the high values of the pixel level should not be at a great distance from the center of the sun and, conversely, darker places should be located throughout the image. The distribution of pixels in the image according to the pixel value is shown in the Figure 3. The high value pixels that are overexposed are close to the center and, conversely, the lower value pixels are mostly in the corner of the image.

Fig. 3. The pixel distribution in the raw image, a) the pixel value level is in the range of 600 - 6000, b) pixels with a value of 6000 - 60,000 and c) pixels with a value higher than 60,000.

B. Adaptive histogram equalization

AHE includes a constant corresponding to the number of processed neighborhood pixels. The contrast increases with smaller scales, and the blur increases with larger numbers of this constant. This can be seen in Figure 4, where there is a comparison of three different values of this constant. It can be seen in the image that a low value enhances fine structures but increases the noise, on the contrary, a high value smoothes these fine structures.
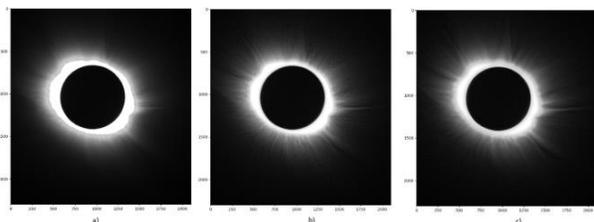

Fig. 4. FCIPHE image processing with three different constants (The size of the neighborhood region) for values a) 5, b) 30 and c) 50.

C. Exposure times

FCIPHE processing is intended for various types of images taken during solar corona observation. Different image input formats should not affect the resulting image quality, as well as the exposure time of the captured image. The principle of the method is shown in the image with an exposure time of 1/160 second, and the result is shown in Figure 5. The FCIPHE method is also shown at other exposure times, 1/60 second in Figure 6 and 1/25 second in Figure 7.
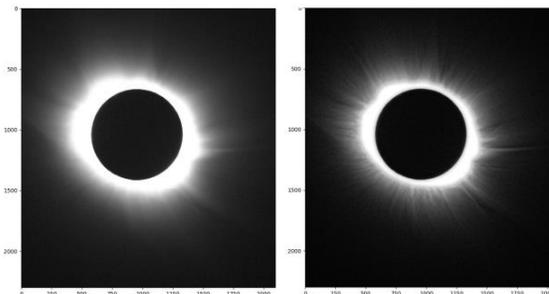

Fig. 5. Comparison of the original raw image a) and the image with the applied FCIPHE method b) on an image with an exposure time of 1/160 second.
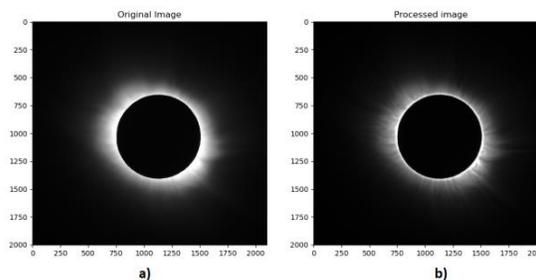

Fig. 6. Comparison of the original raw image a) and the image with the applied FCIPHE method b) on an image with an exposure time of 1/60 second.
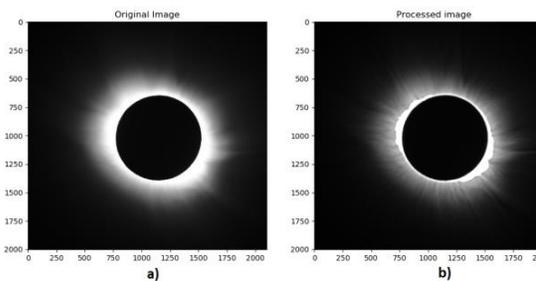

Fig. 7. Comparison of the original image a) and the image processed by the FCIPHE method b) in an image with an exposure time of 1/25 second.

VI. Conclusion

The FCIPHE method is applied to the images from the solar eclipse, which are characterized mainly by large differences in contrast and brightness. It is also difficult to identify the thin structures of the solar corona in the raw image. The result is a single image with visible fine structures that are independent of their position from the solar disk in the image. The results showed that the developed method provides a better view of the image of the solar corona in a single image.

The summary of the goals of the method is following.

- The method provides good results even after its use on a single image, so it is not necessary to calibrate the image or to process several input images. The searched objects can be visually recognized after applying the method to one image.

- The method does not require only one type of image format. This means that the method does not depend on the dimensions of the image or the data from the header of the saved image. The only requirement for processing is sufficient image bit depth.

- The method is timesaving, it can be calculated and presented in a short time. Compared to the NAFE method, it is almost 6 times faster. This, of course, depends on the technical parameters of the computer on which the method is calculated.

- The method can make more visible fine structures of the solar corona that were not previously recognizable due to blurring or their poor visibility.

- The method includes a sequence of steps that is applicable to the observation software. It can provide an instant preview of images without the need to install additional specific software.

The FCIPHE method is designed to work with a single uncalibrated image. Of course, the availability of more observational data or the use of calibration images will also improve the result of this method. In addition to this option, there are three other areas that affect the resulting image quality.

- Number of processed images. Although FCIPHE works with a single input image, it creates a set of individually processed images. The creation of this set of images is universal and therefore does not depend on the quality of the image and its properties. By applying a more sophisticated algorithm for creating a set of images, it is possible to achieve more efficient processing with a better visible result.

- Data modification and noise. In this case, further use of the image using FCIPHE is considered. In addition to the visual presentation of the observed object, the method also provides information on the shape of structures in the solar corona. However, the processing modifies the data, so quantitative analysis of the resulting image is not reliable. Noise is often present in the image in areas that contain useful information about the subject. To preserve useful information, this noise is not eliminated during image processing.

- Fine structures in dark areas. This is the area at a greater distance from the edge of the solar disk. Depending on the exposure time of the input image, these areas are more or less visible due to overexposed pixels around the center of the sun. The method considers the overall uniform image, but in this case an adaptive approach to the individual regions in the image is required, similar to, for example, in Normalizing - a radial filter-graded filter.

Finally, the method provides an efficient way of image processing, which can be the basis for automation of solar corona observations, in evaluating the quality of the selected observation program or in evaluating already acquired images.

References

[1] A. R. CHOUDHURI. An Elementary Introduction to Solar Dynamo Theory. 2007.

[2] M. MAHAJAN, S., MIKLASZEWSKI R., I. Nikol K., Shatashvili N. Primary Flows, the Solar Wind and the Corona. 1999.

[3] David M. RUST, Solar Physics at APL. JOHNS HOPKINS APL TECHNICAL DIGEST, VOLUME 20, NUMBER 4 (1999).

[4] L. GOLUB, J. M. PASACHOFF. The Solar Corona. 2010. ISBN 978-0-521-88201-9.

[5] S.R. CRANMER. Coronal holes and the high-speed solar wind. 2002. Space Science Reviews 101: 229–294.

[6] M. STIX. The Sun, An Introduction. 2002. ISBN 3-540-20741-4.

[7] F. ESPENAK. Periodicity of Solar Eclipses. 2012. NASA Goddard Space Flight Center.

[8] J. W. HARDY. Adaptive optics for astronomical telescopes. Oxford University Press 1998. ISBN 0-19-509019-5.

[9] J. KOZA, J. RYBÁK, P. GÖMÖRY, M. KOZÁK, A. LÓPEZ ARISTE. Spectral Characteristics of the He I D3 Line in a Quiescent Prominence Observed by THEMIS. 2017. Solar Physics, Vol. 292, p. 98.

[10] C.R. KITCHIN. Telescopes and Techniques. 1995. ISBN 978-3-540-19898-7.

[11] A. KUČERA, J. AMBRÓZ, P. GÖMÖRY, M. KOZÁK, J. RYBÁK. CoMP-S - the coronal multi-channel polarimeter for Slovakia. 2010. Contributions of the Astronomical Observatory Skalnaté Pleso, vol. 40, no. 3, p. 135-138.

[12] TOMCZYK S., CARD G., DARNELL T., ELMORE D., LULL R., NELSON P., STREANDER K., BURKEPILE J., CASINI R., JUDGE P. An Instrument to Measure Coronal Emission Line Polarization. 2008. Solar Physics. 247. 411-428. 10.1007/s11207-007-9103-6.

[13] H. WOHL, A. KUCERA, J. RYBAK, AND A. HANSLMEIER. Precise reduction of solar spectra obtained with large CCD arrays. 2002. In: A&A 394, 1077 – 1091.

[14] J. MIANO. Compressed image file formats. 2000. Addison Wesley Longman Inc. ISBN 0201604434.

Magnetic properties of selected rapidly quenched composite materials with amorphous and nanocrystalline structure

Branislav Kunca

Department of Physics
Faculty of Electrical Engineering and Informatics TUKE
Košice, Slovak Republic
branislav.kunca@tuke.sk

Ivan Škorvánek

Department of Applied Magnetism and Nanomaterials
Institute of Experimental Physics, SAS
Košice, Slovak Republic
skorvi@saske.sk

Abstract — **In this paper, we investigate impact of rapid annealing on the magnetic properties of selected rapidly quenched Fe-(Co)-B-(Cu) alloys. Two processing devices, both utilizing pair of preheated copper blocks were used in order to demonstrate importance of correctly tuned annealing conditions, namely time and temperature. Advantages and disadvantages of the furnace system used so far are discussed. New upgraded experimental setup for rapid annealing under vacuum/protective gas atmosphere is proposed.**

Keywords — *nanocrystalline alloys, rapid annealing, microstructure, soft magnetic materials, coercivity*

Abstrakt — **Táto práca sa venuje štúdiu dopadu rýchleho žíhania na magnetické vlastnosti vybraných rýchlo chladených zliatin Fe-(Co)-B-(Cu). K tomuto účelu boli použité dve žíhacie metódy využívajúce pár prednahriatych medených blokov. Na základe výsledkov bola ukázaná dôležitosť správne nastavených parametrov tepelného spracovania, najmä žíhacieho času a teploty. Boli vyhodnotené výhody a nevýhody zariadenia ONAM na rýchle tepelné spracovanie. Je načrtnutý návrh novej experimentálnej aparatúry pre rýchle tepelné spracovanie vo vákuu/ochrannej atmosfére.**

Kľúčové slová — *nanokryštalické zliatiny, rýchle žíhanie, mikroštruktúra, magneticky mäkké materiály, koercivita*

I. INTRODUCTION

Nanocrystalline soft magnetic alloys have become one of the most perspective groups of materials for research and development due to possibilities to obtain excellent soft magnetic properties. They may be achieved by combination of appropriate alloy composition and suitable thermal processing procedure. Three decades of research proved application potential of several classes of nanocrystalline soft magnetic alloys for various technical, power or sensory applications [1-4]. Specifically designed composition of FINEMET alloy (Fe-Si-Nb-B-Cu) [5] helped to attain nanocrystalline structure of ferromagnetic α – Fe(Si) grains with mean diameter of 12 – 15nm which lead to extremely low values of coercivity and high permeability. Higher content of Fe in NANOPERM alloys (Fe-M-B-(Cu), where M is metalloid) [6] helped to increase saturation magnetization. Further enrichment by Co led to production of HITPERM (Fe-Co-M-B-(Cu), where M is metalloid) [7] with improved thermal stability of magnetization and significantly increased Curie temperature T_C.

Conventional annealing of nanocrystalline alloys consists of processing of the samples for 0.5 – 1h at temperatures slightly above the temperature of primary crystallization [8]. This leads to formation of nanocrystalline ferromagnetic grains embedded in the residual amorphous matrix. Primary crystallization of the as – cast ribbons are governed by nucleation and growth mechanisms, both driven by atomic diffusion [9]. Subsequent coercivity reduction and increase of permeability may be explained by the Herzer's random anisotropy model RAM [10]. Grain refinement effectively decreases influence of magnetocrystalline anisotropy and thus enhances long-range effect of exchange interaction between nanocrystals.

Recently, rapid annealing method has been utilized for processing of soft magnetic (nanocrystalline) alloys. Several approaches were designed, including millisecond current pulses [11], flash annealing under tensile stress [12] or fast Joule heating [13]. However, set of preheated copper blocks shows up to be the most promising due to potential for achieving extremely high heating rates ($10^2 – 10^4$ K/s) and at the same time prevention against undesirable self - heating of the sample caused by latent heat released as a consequence of undergoing crystallization processes [14]. It has been shown [9, 15] that due to rapid increase of heating rate and application of temperatures near glass transition temperature, crystallization process is governed predominantly by viscous flow rather than atomic diffusion. At this point, nucleation rate of ferromagnetic grains increases rapidly. Low processing times and damping of growth rate help to develop nanocrystalline structure even in alloys which do not respond well to conventional annealing.

In this work we summarize results for $Fe_{87}B_{13}$ and $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloys, processed using a rapid annealing method. We discuss positives and negatives of the experimental approach and processing conditions. Design of new experimental apparatus is proposed for improvement of rapid annealing technique at Department of Applied Magnetism and Nanomaterials, IEP SAS.

II. EXPERIMENTAL

Rapidly quenched amorphous precursor ribbons were produced by planar flow casting method. $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ samples were provided by Institute of Physics, SAS, while $Fe_{87}B_{13}$ ribbons were prepared at Department of Materials Science and Engineering (DMSE), Monash University. Two furnace systems were utilized to conduct rapid annealing of the as – quenched samples.

First rapid annealing furnace system was used at Department of Applied Magnetism and Nanomaterials (DAMN), Institute of Experimental Physics SAS in Košice, see Fig. 1. The furnace system uses pre – heated element consisting of two massive, hinge connected Cu blocks placed inside the muffle furnace. Setpoint annealing temperature was stabilized for at least 15 minutes before sample processing while being constantly monitored by thermocouple mounted on the side of bottom block. Annealing time has been defined as a time period between point of insertion and point of removal of the sample. The heating rate has been estimated to more than
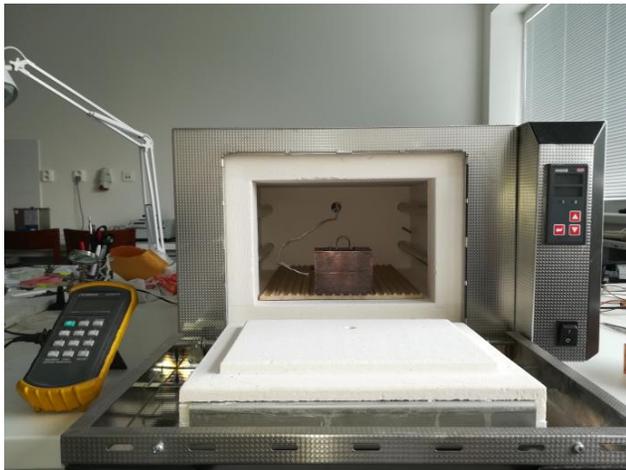

Fig. 1. Furnace system with heating element for rapid annealing situated at DAMN, Institute of Experimental Physics, SAS.

100 K/s. Minimum processing time to obtain reproducible results was found at 5s

Second processing procedure was realized at Department of Materials Science and Engineering (DMSE), Monash University, with courtesy of prof. Kiyonori Suzuki. This furnace system uses similar processing technique as the one used at DAMN, with several important improvements including protective Ar atmosphere and pneumatic cylinder operated Cu block system. The heating rate has been estimated

to 104 K/s [16]. Annealing time was controlled by automated timing mechanism. More detailed description of the construction and operating mechanism can be found elsewhere [17].

Individual ribbons were foremost wrapped in the Cu foil in order to provide its uniform heating. Afterwards, specimens were clutched between the preheated two copper blocks for required annealing time. After annealing, removed samples were rapidly cooled down using either separate Cu block held at room temperature (DNAM furnace) or fan (DMSE system).This ensured termination of undefined heating of the ribbons caused by heat from the Cu envelope.

Magnetic properties of as – quenched and rapidly annealed ribbons were studied using the measurement equipment of both DAMN and DMSE. Saturation magnetization and coercivity values of the studied ribbons were obtained using the MicroSense EV9 vibrating sample magnetometer (VSM) and BHS-40 DC hysteresis loop tracer by Riken Denshi. To ensure enough statistical data for evaluation of magnetic parameters, 3 samples were used for VSM measurements. Microstructure of selected processed alloys was studied using the X-Ray diffraction method (XRD) using Bruker D8 Advance diffractometers (Cu - K_α radiation) with Goebel mirror in the incident beam and LiF monochromator in the diffracted path.

III. RESULTS AND DISCUSSION

A. Annealing temperature

Despite the many positive impacts of glass forming, non-ferromagnetic elements on the development of microstructure and magnetic properties of the nanocrystalline alloys, its content needs to be kept as low as possible in order to maintain high saturation magnetization. Certain amount of alloying elements is necessary for achievement of glassy state in the ribbons produced by planar flow casting method. Considering the simple binary Fe-B system, in order to prepare samples with fully amorphous structure, minimum required B content was found to be approx. 10 at% for thin films [18] or 12-13 at% for ribbons [19]. Subsequent conventional annealing of these binary alloys for 0.5 – 1h near temperature of primary crystallization leads to undesirable increase of coercivity mainly due to large grain size [16, 17, 20] and/or formation of Fe_2B or Fe_3B phases [21, 22] with large magnetocrystalline anisotropy [23]. Significant improvement of the soft magnetic characteristics (and of other ferromagnetic alloys) has been reported after rapid annealing using a set of preheated copper blocks [14, 17, 24 - 29]. Fe-(Co)-B systems are currently studied in more detail for the unprecedented reduction of coercivity values H_c and high saturation magnetization M_s, comparable to that of Si steels [27, 28]. The main reasons for this effect were found to be i.a. high heating rate [29] and suppression of undesirable self - heating of the sample [14] during rapid annealing.

In order to observe positive effects of rapid annealing on magnetic properties of studied alloys, processing conditions need to be optimized first. Importance of correctly chosen temperature will be demonstrated on $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloy. Addition of 1 at% Cu works as a nucleation agent during

nanocrystallization process and effectively helps to increase volume density of the nanocrystalline grains [30]. Primary crystallization temperature of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloy was observed at 367 degC [31]; yet magnetic softening and formation of nanocrystalline structure of ferromagnetic grain was achieved even after annealing for 1 h at 320 degC in the presence of longitudinal magnetic field. Rapid annealing conducted at DAMN was performed at 380 – 420 degC, close to primary crystallization temperature. This was meant to be compensated by application of rather extended annealing times 10 – 30s. Coercivities of the samples annealed at 380 and 400 degC oscillate between 12 and 17 A/m without any observable dependence on the processing time and/or temperature, see Fig. 2. The steep increase of H_c is more pronounced after processing for 20 and 30s at 420 degC. Development of saturation magnetization depicted in the Fig. 3 clearly shows dependence on the processing time yet the values remain more or less constant for each annealing temperature. M_s increases from 189 Am^2/kg (10 s) to 194 Am^2/kg (30 s).

Fig. 4 shows XRD data obtained for samples rapidly annealed at 380 degC which seems to be the most suitable for processing of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloy. Broad pattern present at around $2\theta = 45°$ clearly indicate an amorphous structure in all studied samples. Some minute traces of nanocrystalline grains are present after rapid annealing; nevertheless the overall fraction of nanocrystalline phase is well below the resolution capacity of the diffractometer (which is typically around 4%). Increase of coercivity after
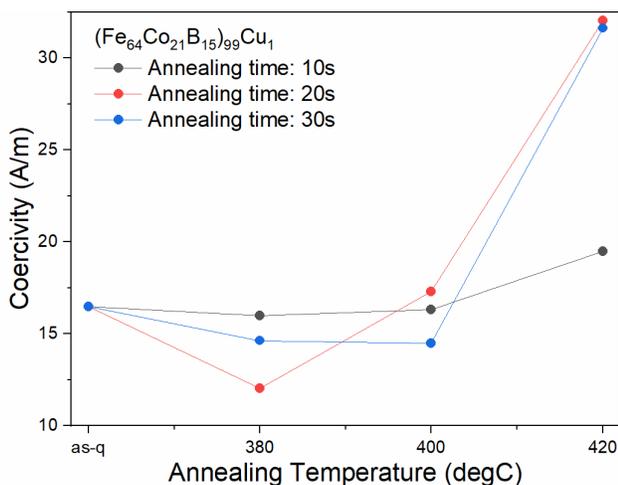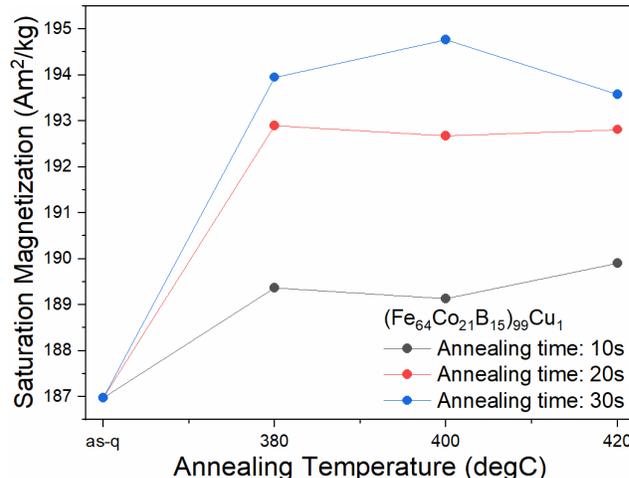

Fig. 3. Saturation magnetization values of $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$, rapidly annealed system for 10 – 30 s at 380 – 420 degC using the DAMN furnace.

rapid annealing above 400 degC occurs as a result of development of coarse – grained nanocrystalline structure [32]. Rapid thermal processing of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ was executed using the furnace apparatus at DMSE as well. Considering previous mediocre results, higher annealing temperatures combined with ultra-low processing time (0.5s) were chosen in order to achieve improvement of soft magnetic properties. Such approach was shown to be favorable for processing of Fe-B - based alloys [27].

Coercivity values of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ ribbons, rapidly annealed for 0.5s between 460 and 530 degC are depicted in the Fig. 5. A significant coercivity decrease is achieved instantly after treatment at 460 degC. H_c values range between 2 and 4 A/m for samples rapidly annealed up to 500 degC. Temperature scans of magnetization (not shown here) conducted on the processed samples unveiled that rapid annealing at 480 degC and below is insufficient to ensure full nanocrystallization of the studied alloy. The ideal temperature
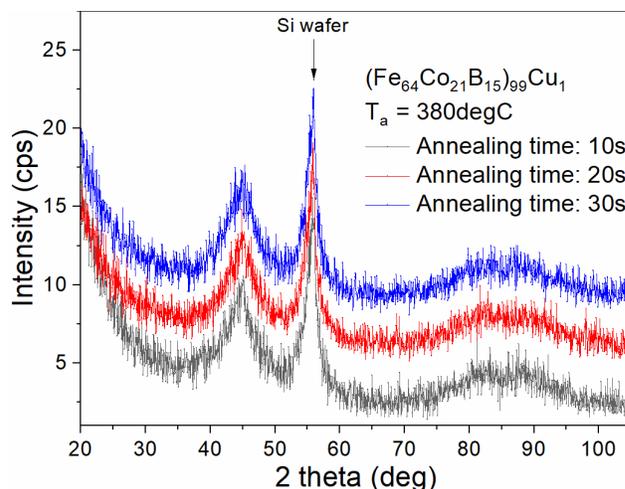

Fig. 2. Coercivity values obtained for $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$, rapidly annealed using the DAMN furnace system for 10 – 30 s at 380 – 420 degC.

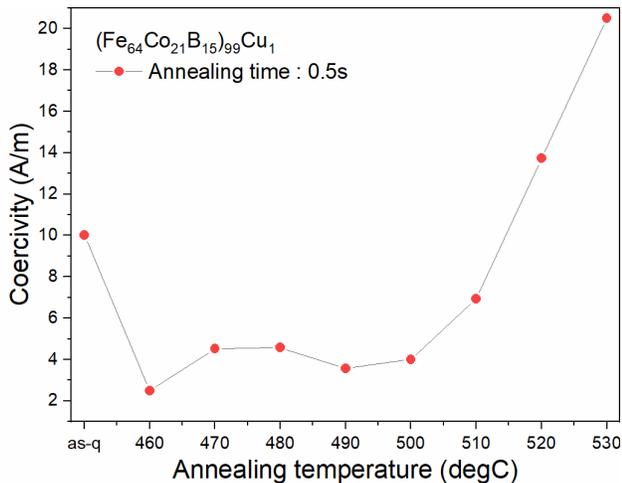Fig. 4. XRD patterns obtained for $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$, rapidly annealed for 10 – 30 s at 380 degC using the DAMN furnace system.

Fig. 5. Coercivity values obtained for $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$, rapidly annealed for 0.5 s at 460 - 540 degC using the DMSE furnace system.

range for rapid thermal processing of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloy shows up to be between 490 – 500degC. The saturation magnetization values in the Fig. 6 increased steadily up to approx. 200 Am^2/kg, which are higher than values obtained for those rapidly annealed at 380 – 420 degC.

Improvement of the soft magnetic properties achieved for $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ after rapid annealing for 0.5s in the temperature range 460 – 500 degC can be described by work of Köster et al. [15]. Nucleation and growth rate during crystallization may be tuned by selection of processing temperature. The main process governing formation of nanocrystalline grains below glass transition temperature T_g is an atomic diffusion. Increasing the temperature to the region near T_g and above, the viscous flow mechanism becomes more pronounced. Formation of nuclei is enhanced due to increase of the nucleation rate while growth rate of the grains is at the same time markedly suppressed [33]. Rapid annealing of the metallic alloys should be performed at temperatures100 degC
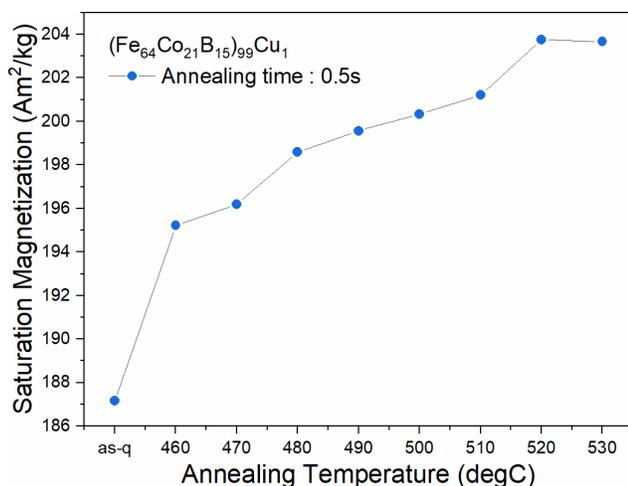

Fig. 6. Saturation magnetization values obtained for $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$, rapidly annealed for 0.5 s at 460 - 540 degC using the DMSE furnace system.

higher than primary crystallization temperature obtained under conventional thermal analysis (such as DSC), ideally above its glass transition temperature T_g [24]. The main issue with annealing at these temperatures withstands from the fact that most metallic alloys could experience not only primary but secondary (and higher) crystallization processes.

Therefore the onset of crystallization needs to be shifted above glass transition temperature, which may be achieved by elevated heating rate of the rapid annealing technique. Under such conditions it is possible to accomplish suitable nanocrystallization process even in binary metallic alloys which usually do not respond well to the conventional annealing techniques [9].

Because primary crystallization temperature of the $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ alloy obtained by the thermomagmnetic measurements at a constant heating rate of 10 K/min was observed at 367 degC [31], nucleation kinetics during crystallization process occuring during rapid annealing between 380 and 420 degC could not be fully governed by viscous flow and therefore no significant magnetic softening has been observed. Compensation of lower annealing temperature by relatively longer processing times cannot be effective either as it does not have impact on the prevailing crystallization mechanism and it only manifest in increase of the grain size.

B. Annealing time

Improvement of soft magnetic properties in $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ was apparently achieved by increase of the processing temperature. Higher nucleation rate led to formation of nanocrystalline grains embedded in the residual amorphous matrix and thus magnetic softening of the alloy. We have observed that longer annealing time cannot compensate for low processing temperature and have significant impact on the grain size.

Importance of appropriately chosen annealing times will be demonstrated on rapidly annealed binary Fe-B ribbons. Alloy with nominal composition $Fe_{87}B_{13}$ was chosen in order to maximize content of the ferromagnetic element (Fe) [19] and at the same time ensure presence of amorphous structure of the ribbon after its production using the planar flow casting method [21]. Produced amorphous precursor samples were rapidly annealed for 0.5s using the annealing furnace system at DMSE, and for 5s and 8s at DAMN. Both furnace systems are based on the same principle, i.e. clamping of the sample wrapped in the Cu foil between two pre-heated massive copper blocks. Processing temperature range of 480 – 500 degC was chosen based on the previous research on this alloy reported in [28].

Fig. 7 shows coercivity dependence on annealing temperature at selected processing times. Very short annealing for 0.5 s led to decrease of H_c to 10 – 13 A/m as a result of formation of nanocrystalline α – Fe grains. These values are similar to those reported in [27]. The importance of low annealing time is shown on the samples rapidly annealed at the same temperatures for 5s and 8s. Significant coercivity
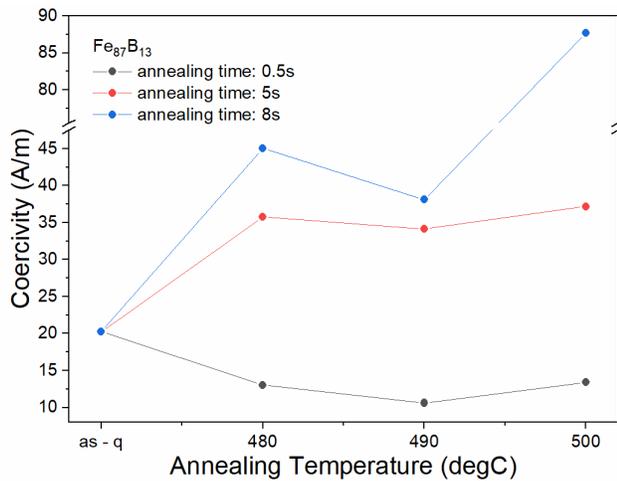
Fig. 7.   Coercivity values obtained for $Fe_{87}B_{13}$, rapidly annealed for 0.5, 5 and 8s at 480 – 500 degC.

increase with increasing annealing time caused by grain coarsening makes the difference between unsuccessful processing and material with potential for use in technical applications. H_c increases threefold after 5s rapid annealing and in the most extreme case, the coercivity of $Fe_{87}B_{13}$ approaches 87 A/m which makes it practically unusable for technical applications. Saturation magnetization depicted in the Fig. 8 shows very similar development for all processed samples and ranges between 197 and 199 Am^2/kg which corresponds to 1.92T [27].

It is therefore evident, that processing time used for rapid annealing of the metallic ribbons should be kept as low as possible, as even 0.5s is sufficient for creation of desired nanocrystalline structure. These results show that $Fe_{87}B_{13}$, rapidly annealed for 0.5s at 480 – 500 degC has a great potential for use in industrial use. Because of very low core losses compared to Fe-Si steels [26] and use of inexpensive and naturally abundant elements, rapidly annealed Fe-B alloys
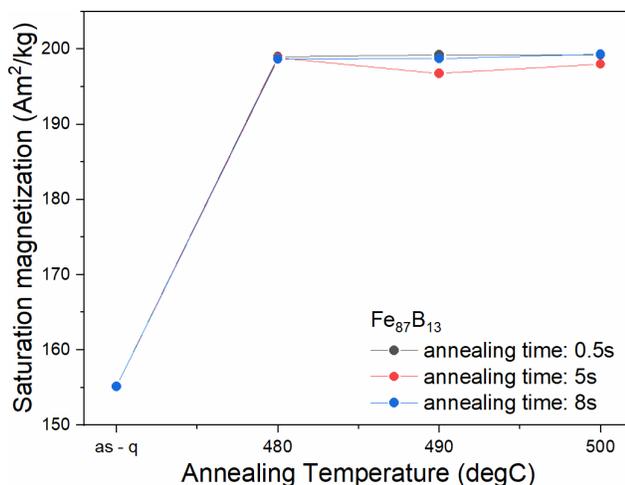

Fig. 8.   Saturation magnetization values obtained for $Fe_{87}B_{13}$, rapidly annealed for 0.5, 5 and 8s at 480 – 500 degC.

are promising candidates for applications with low energy loss requirements. For these reasons, study of thermal performance and stability of its magnetic properties at elevated temperatures is currently undergoing at DAMN.

C. New furnace system

Upgraded experimental apparatus and thus the annealing procedure at DAMN has been proposed in order to improve several disadvantages of the currently used device:

- Annealing atmosphere: originally designed apparatus only allowed processing in the air. As has been shown in [34] oxidation of the samples during rapid annealing up to 60s did not show up. Nevertheless, processing for extended period of time led to formation of Fe_3O_4 and CuO oxides due to exposure to air, see Fig.9. Absence of protective atmosphere may be even more striking in case of alloys sensitive to oxidation, e.g. those containing Zr.

- Annealing time: processing of the samples was conducted manually by the operator. As a result, lowest annealing time yielding reproducible results (5s) was much higher
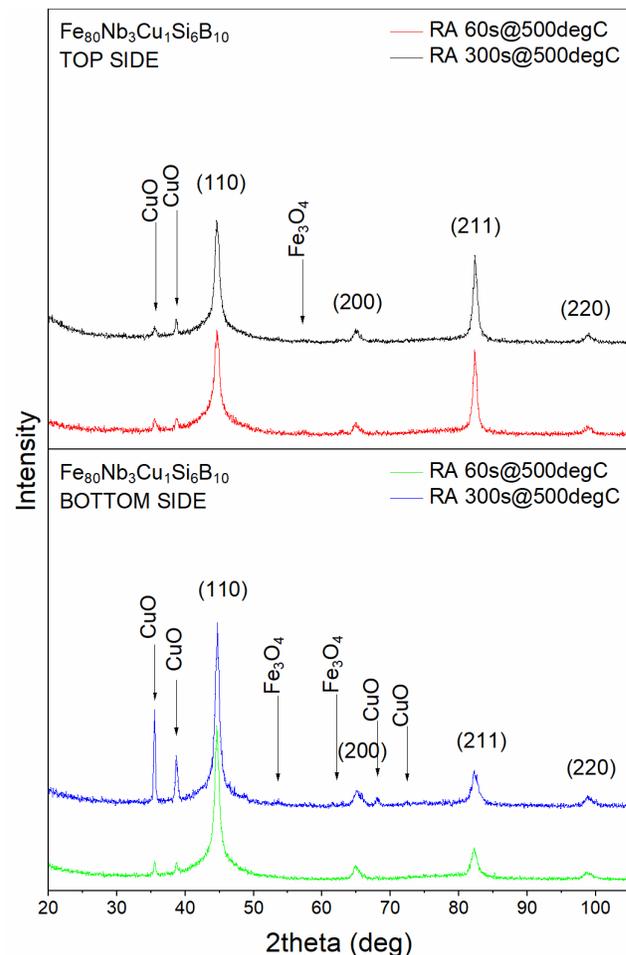

Fig. 9.   XRD patterns from the TOP and BOTTOM side of the $Fe_{80}Nb_3Cu_1Si_6B_{10}$ ribbon, rapidly annealed for 60s and 300s at 500°C.

than required (0.5s). Due to handling inconsistency, error of approx. 1– 2s may arise as well. Well defined processing procedure, independent of operator needed to be established.

- Thermal stability: processing of the samples required opening the front door of the muffle furnace, thus exposing the heating element to external ambient atmosphere at room temperature Cooling of the block surface led to temperature error estimated up to 5 degC.

For these reasons a new experimental setup has been proposed, see fig. 10. Furnace allows rapid annealing both under high vacuum and in the protective gas atmosphere. Annealing technique remained the same. Two copper blocks (each approx. 1.5 kg) are placed into a frame which allows vertical movement of the upper one. Heating of the Cu blocks is arranged by resistive wires supplied from the external power sources and controlled by custom PID program inside the LabVIEW software environment. Temperature is detected by two separate thermocouples, both placed on the side of individual blocks. Block frame is placed inside the furnace tube (green) and is operated via the rotating feedthrough using the stepper motor. This upgrade makes improvement in definition of annealing time, as the processing period during which is the sample clutched between the copper blocks is now determined by software program. Linear feedthrough transports the sampleholder from the sample chamber (yellow cross) to the annealing area placed inside the furnace tube. After annealing, processed sample is dragged back to the sample chamber where it is cooled down by blow of the noble gas. System of pneumatic gates and set of valves has been carefully proposed in order to prevent aeriation or other vacuum deterioration (or eventual contamination of the gas atmosphere, if used). They also separate furnace apparatus to two independent parts (sample chamber and furnace tube) which allows continuous pumping of the furnace tube and at the same time pressurizing the sample chamber in order to exchange the samples for further processing. The furnace apparatus is currently under construction at DAMN.
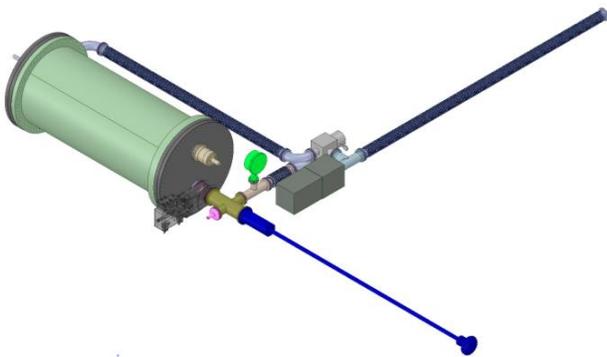

Fig. 10.   Design of new furnace system for rapid annealing.

IV. Summary

Impact of processing time and temperature on development of soft magnetic characteristics was demonstrated in selected Fe-(Co)-B - based alloys. Importance of suitable processing temperature was shown on microstructure formation and magnetic properties of rapidly quenched $(Fe_{64}Co_{21}B_{15})_{99}Cu_1$ ribbons. Despite compensation of low annealing temperatures close to primary crystallization temperature by extended processing time, formation of nanocrystalline structure leading to significant magnetic softening has not been observed. Rapid annealing for 0.5s close to 500degC resulted in substantial decrease of coercivity close to 3 A/m and increase of saturation magnetization to approx. 200 Am^2/kg. Impact of annealing times was demonstrated on rapidly quenched $Fe_{87}B_{13}$ ribbons. Increase from 0.5s to 5s caused threefold increase of coercivity. After 8s rapid annealing, maximum H_c of 87 A/m was achieved, making the material unsuitable for technical applications. Experimental results have been utilized for design and construction of new furnace system for rapid annealing under vacuum or protective gas atmosphere.

References

[1]   T. Eggers et al., "Impact of the transverse magnetocrystalline anisotropy of a Co coating layer on the magnetoimpedance response of FeNi-rich nanocrystalline ribbon", J. Alloys Compd., vol. 741, pp. 1105-1111, Apr. 2018.

[2]   A. Nishiyama, K. Tanimoto and A. Makino, "Outstanding efficiency in energy conversion for electric motors constructed by nanocrystalline soft magnetic alloy "NANOMET®" cores", AIP Advances, vol. 6, is. 5, art. no. 055925, Jan. 2016.

[3]   Y. Naitoh, T. Bitoh, T. Hatanai, A. Makino and A. Inoue, "Application of nanocrystalline soft magnetic Fe–M–B (M=Zr, Nb) alloys to choke coils", J. Appl. Phys., vol. 83, is. 11, pp. 6332 – 6334, Oct. 1998.

[4]   G.H. Kim, T.H. Noh, G.B. Choi and K.Y. Kim, "Magnetic properties of FeCuNbSiB nanocrystalline alloy powder cores using ball-milled powder", Journal of Applied Physics, vol. 93, is. 10, pp. 7211 – 7213, May 2003.

[5]   Y. Yoshizawa, S. Oguma and K. Yamamuchi, "New Fe- based soft magnetic alloys composed of ultrafine grain structure", J. Appl. Phys., vol. 64, is. 10, pp. 6044-6046, Sep. 1988.

[6]   K. Suzuki, N. Kataoka, A. Inoue, A. Makino and T. Masumoto, "High Saturation Magnetization and Soft Magnetic Properties of BCC Fe-Zr-B Alloys With Ultrafine Grain Structure", Mater. T. JIM, vol. 31, is. 8, pp. 743 – 746, Aug. 1990.

[7]   M.A. Willard et al., "Magnetic properties of HITPERM $(Fe,Co)_{88}Zr_7B_4Cu_1$ magnets", J. Appl. Phys., vol. 85, is. 8, pp. 4421 – 4423, Apr. 1999.

[8] M.E. McHenry, M.A. Willard and D.E. Laughlin, "Amorphous and nanocrystalline materials for applications as soft magnets", Prog. Mat. Sci., vol. 44, is. 4, pp. 291-434, Oct. 1999.

[9] U. Köster and J. Meinhardt, "Crystallization of highly undercooled metallic melts and metallic glasses around the glass transition temperature", Mat. Sci. Eng. A, vol. 178, is. 1 – 2, pp. 271 – 278, Apr. 1999.

[10] G. Herzer, "Grain size dependence of coercivity and permeability in nanocrystalline ferromagnets", IEEE Trans. Magn., vol. 26, is. 5, pp. 1397 – 1402, Sept. 1990.

[11] W.L. Johnson et al., "Beating crystallization in glass-forming metals by millisecond heating and processing", Science, vol. 332, is. 6031, pp. 828 – 833, May 2011.

[12] G. Herzer, V. Budinsky and Ch. Polak, "Magnetic properties of FeCuNbSiB nanocrystallized by flash annealing under high tensile stress", Phys. Status Solidi B, vol. 248, is. 10, pp. 2382-2388, Oct. 2011.

[13] J. Bednarcik, M. Cesnek and P. Sovak, "Soft magnetic amorphous alloys in X-ray light: Insights from ultra-fast Joule heating experiments", J. Magn. Magn. Mat., vol. 499, art. no. 166282, Apr. 2020.

[14] R. Parsons et al., "Effect of latent heat during primary crystallization on the nanostructural formation process in nanocrystalline soft magnetic materials", AIP Advances, vol. 10, is. 1, art. no. 015105, Jan. 2020.

[15] U. Köster and U. Herold, "Crystallization of metallic glasses" in Glassy Metals I, H.J. Güntherodt and H. Beck, Springer – Verlag, 1981, pp. 225 - 259.

[16] R. Parsons et al.: Nano-crystallisation and magnetic softening in Fe–B binary alloys induced by ultra-rapid heating, J. Phys. D: Appl. Phys., vol. 51, art. no. 415001, Sep. 2018.

[17] B. Zang, R. Parsons, K. Onodera, H. Kishimoto, A.C.Y. Liu and K. Suzuki, "Effect of heating rate during primary crystallization on soft magnetic properties of melt-spun Fe-B alloys", Scripta Mater., vol. 132, pp. 68 – 72, Apr. 2017.

[18] C.L. Chien and K.M. Unruh, "Magnetic properties of amorphous Fe_xB_{100-x} (0 < x < ~90)", Phys. Rev. B, vol. 24, is. 3, pp. 1556 – 1558, Aug. 1981.

[19] R. Hasegawa and R. Ray, "Iron- boron metallic glasses", J. Appl. Phys., vol. 49, is. 7, pp. 4174 – 4179, Aug. 2008.

[20] A.L. Greer, "Crystallisation kinetics of $Fe_{80}B_{20}$ glass", Acta Mater., vol. 30, is.1, pp. 171 – 192, Jan. 1982.

[21] T. Nakajima, E. Kita and H. Ino, "Crystallization process and magnetic properties of $Fe_{100-x}B_x$ (10 ≦ × ≦ 35) amorphous alloys and supersaturated state of boron in α-Fe", J. Mat. Sci., vol. 23, pp. 1279 – 1288, Apr. 1988.

[22] M. Takahasi, M. Koshimura and T. Abuzuka, "Phase Diagram aof Amorphous and Crystallized Fe-B Alloy System", Jap. J.Appl. Phys., vol. 20, no. 10, pp. 1821 – 1832, Oct. 1981.

[23] I. Skorvanek, J. Kovac, J. Marcin, P. Duhaj and R. Gerling, "Annealing effects on the magnetic properties of nanocrystalline FeNbB alloys", J. Magn. Magn. Mat., vol. 203, is. 1 – 3, pp. 226 – 228, Aug. 1999.

[24] K. Suzuki et al., "Nanocrystalline soft magnetic materials from binary alloy precursors with high saturation magnetization", AIP Advances, vol. 9, is. 3, art. no. 035311, Mar. 2018.

[25] L. Morsdorf et al., "Phase selection and nanocrystallization in Cu-free soft magnetic FeSiNbB amorphous alloy upon rapid annealing", J. Appl. Phys. 119, is. 12, art. no. 124903, Mar. 2016.

[26] R. Parsons et al., "Core loss of ultra-rapidly annealed Fe-rich nanocrystalline soft magnetic alloys", J. Magn. Magn. Mat., vol. 476, pp. 142 – 148, Apr. 2019.

[27] R. Parsons, Z. Li and K. Suzuki, "Nanocrystalline soft magnetic materials with a saturation magnetization greater than 2 T", J. Magn. Magn. Mat., vol. 485, pp. 180 – 186, Sep. 2019.

[28] K. Suzuki, R. Parsons, B. Zang, K. Onodera, H. Kishimoto and A. Kato, "Copper-free nanocrystalline soft magnetic materials with high saturation magnetization comparable to that of Si steel", Appl. Phys. Lett., vol. 110, is. 1, art. no. 012407, Jan. 2017.

[29] R. Parsons, B. Zang, K. Onodera, G. Kishimoto, A. Kato and K. Suzuki, "Soft magnetic properties of rapidly-annealed nanocrystalline Fe-Nb-B-(Cu) alloys", J. Alloys Compd., vol. 723, pp. 408 – 417, Nov. 2017.

[30] T.H. Noh , M.B. Lee, H.J. Kim and I.K. Kang, "Relationship between crystallization process and magnetic properties of Fe- (Cu- Nb)- Si- B amorphous alloys", J. Appl. Phys., vol. 67, is. 9, pp. 5568 – 5570, May 1990

[31] J. Marcin et al., "Melt-Spun Fe–Co–B–Cu Alloys with High Magnetic Flux Density for Relax-Type Magnetometers," IEEE Trans. Magn., vol. 46, is. 2, pp. 416 – 419, Feb. 2010

[32] I. Skorvanek, B. Kunca, J. Marcin and P. Svec, "Impact of rapid annealing on the soft magnetic properties of Fe-Co-B-Cu melt-spun alloys". Presented on 24th Soft Magnetic Materials Conference. Poznan, September 4-7, 2019.

[33] T. Kulik, "Nanocrystallization of metallic glasses", J. Non-Cryst. Solids, vol. 287, is. 1 – 3, pp. 145 – 161, Jul 2001

[34] B. Kunca, J. Marcin, P. Svec and I. Skorvanek, "Effect of rapid annealing on magnetic properties of the nanocrystalline $Fe_{80}Nb_3Cu_1Si_6B_{10}$ alloy", Acta Electrotech. Inform., vol. 19, is. 2, pp. 33 – 37, Jun 2019

Explainable Classifier Supporting Decision-making for Colorectal Cancer Diagnosis from Histopathological Images

Patrik Sabol
Department of Cybernetics and Artificial Intelligence
Technical University of Košice, Košice, Slovakia

Peter Sinčák
Department of Cybernetics and Artificial Intelligence
Technical University of Košice Košice, Slovakia

Pitoyo Hartono
School of Engineering
Chukyo University
Nagoya, Japan

Abstract—**Pathologists are responsible for cancer type diagnoses from histopathological cancer tissues. However, it is known that microscopic examination is tedious and time-consuming. In recent years, a long list of machine learning approaches to image classification and whole-slide segmentation has been developed to support pathologists. Although many showed exceptional performances, the majority of them are not able to rationalize their decisions. In this study, we developed an explainable classifier to support decision making for medical diagnoses. The proposed model does not provide an explanation about the causality between the input and the decisions, but offers a human-friendly explanation about the plausibility of the decision. Cumulative Fuzzy Class Membership Criterion (CFCMC) explains its decisions in three ways: through a semantical explanation about the possibilities of misclassification, showing the training sample responsible for a certain prediction and showing training samples from conflicting classes. In this paper, we explain about the mathematical structure of the classifier, which is not designed to be used as a fully automated diagnosis tool but as a support system for medical experts. We also report on the accuracy of the classifier against real world histopathological data for colorectal cancer. We also tested the acceptability of the system through clinical trials by 14 pathologists. We show that the proposed classifier is comparable to state of the art neural networks in accuracy, but more importantly it is more acceptable to be used by human experts as a diagnosis tool in the medical domain.**

I. Introduction

Histopathological images of cancer tissue are routinely analysed by pathologists who are responsible for the cancer type diagnosis and prognosis [1]. Different types of tissues can be distinguished from histopathological evaluations of Hematoxylin and Eosin (H&E) stained tissue sections. In colorectal cancer (CRC), the tumour architecture changes during tumour progression and is related to patient prognosis [2]. Microscopic examination of tissue sections is known to be tedious and time-consuming [3]. Furthermore, the outcome of the analysis may be affected by the levels of experience of the pathologists involved. Therefore, with the advancement of digital pathology, computer-aided analysis of the histopathological images and machine learning-based diagnostic systems, the fidelity and efficacy of medical diagnoses can be significantly improved.

A long list of machine learning approaches to image classification and whole-slide segmentation has been developed to support pathologist in interpreting histopathological images [4], [5]. Especially, in the recent years, conventional classification approaches, which mainly rely on manually-engineered features, were outperformed by Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) [6]. The high performances of deep models is due to their ability to automatically extract representations that are strongly relevant to the their predictions from the learning data. However, their performance is not supported by their ability to explain their decisions and hence, may prevent their applicability in real world clinical settings [7]. Our objective in this study is to build a reliable classifier that is also able to provide explanations about its decision in human-friendly forms. We believe that in real world clinical settings, which require accountability, the accuracy of a classifier should be paired with its explainability [8].

A. Relevant studies

In [9], *Holzinger et al.* distinguished two types of explainable AI: *Ante-hoc systems* which incorporate explainability directly into the structure of an AI-model; these are systems that are interpretable by design. Typical examples include linear regression, decision trees and fuzzy inference systems. They are commonly referred to as *white-boxes* or, currently, *glass-boxes* [10]. *Posthoc systems*, on the other hand, aim to explain and interpret *black-box* classifiers which provide local explanations for their specific decision. The majority of the explanation approaches seek to link a particular output of the classifier to input variables to see the impact of features on the final decision outcome. For instance, in [11], *G. R. Vásquez-Morales et al.* used neural network-based classifier to predict whether a person is at risk of developing chronic kidney disease. Here, a black-box machine-learning method was complemented by Case-Based Reasoning, a white-box method that is able to find explanatory cases for an explanation-by-example justification of a neural network's prediction. In [12], *Mullenbach et al.* presented an attentional convolutional

network that predicted medical codes from clinical texts. Using an attention mechanism, the most relevant segments of the clinical text for each of the medical codes were selected and used as an explanation mechanism. Through an interpretability evaluation by a physician, they showed that the attention mechanism identified meaningful explanations. In [13], *Lundberg et al.* presented an ensemble-model-based machine learning method using deep learning that predicts the near-term risk of hypoxaemia during anaesthesia care and explains the patient and surgery-specific factors that led to that risk. The system improved the performance of anesthesiologists by providing interpretable hypoxaemia risk and the contributing factors. In [7], *Hagele et al.* utilized Layer-wise Relevance Propagation (LRP) to provide pixel-level explanation heatmaps for the classification decision of the CNN in digital histopathology analyses of tumour tissue. These explanations were used to improve the generalization of the classifier by detecting and removing the effects of hidden biases in used datasets. A similar approach to visualize parts of the input image responsible for the prediction was used in [14], where LIME (Local Interpretable Model-agnostic Explanations) was utilized to provide a global understanding for the CNN model by providing explanations for individual instances in the context of in-vivo gastral image analysis.

It is natural that in the delicate medical domain, prediction models should not only be accurate, but also accountable; they should state uncertainty in their predictions, indicating difficult cases for which further human expert inspections are necessary. Therefore, another approach to the probe and interpret machine learning algorithm is to measure the uncertainty of the prediction for one particular example, the *predictive uncertainty* [9]. In [15], a transparent neural network, S-rRBF, was proposed and applied to DNA microarray data sets. It provides an intuitive explanation through a visualization of its decision process and on the given problem. It allows the users to understand why a certain problem is easy or difficult. Moreover, it makes it possible to see whether a new input is hard to classify or unlikely to be misclassified. However, the visual information still needs to be interpreted and thus is prone to subjective inconsistencies. For the field of digital pathology, in [16], *Raczkowski et al.* proposed an accurate, reliable and active (ARA) image classification framework using a Bayesian Convolutional Neural Network (ARA-CNN) for classifying histopathological images of colorectal cancer. The model is able to achieve reliability by measuring the uncertainty of each prediction. This capability was used to identify mislabelled training samples. In [17], the recently proposed semantically explainable fuzzy classifier called Cumulative Fuzzy Class Membership Criterion (CFCMC) [18], [19] was used to classify histopathology images for breast cancer and to generate additional information about classification reliability in human-friendly terms, in the form of a semantic explanation. It provides a confidence measure for the classification result of a test image followed by a visualization of train image and the most similar images that belong to clusters of the conflicting class with a different confidence degree. In this

paper, we extend the explainability of the CFCMC classifier by defining the *factor of miclassification (FoM)* and the *certainty threshold*. While the *FoM* is a value that describes the possibility of the input sample being misclassified to the one particular conflicting class, *the certainty threshold* is a value of the *FoM*, under which it is a certain that the input sample will not be misclassified. Compared to the concept of the uncertainty measure proposed in [16], in the case of uncertain prediction, our approach is additionally able to suggest the classes in which the input sample could be misclassified. Thus, it offers relevant classes to be further examined. Therefore, we believe that our approach is able to speed up the decision process of the experts, and hence it is suitable for supporting decision-making in time-crucial medical domains. The objective of this article is to apply the explainable CFCMC classifier for the classification of histopathological images of colorectal cancer. We used a publicly available dataset that was released in [2] by *Kather et al.*. It consists of a training set comprised of 5000 small tiles, each of them annotated with one of eight tissue classes and 10 non-annotated whole slide images (WSI) of the tissue.

In [16], it was shown that CNN outperformed the approach in [2], where features derived from images using texture descriptors served as a basis for a support vector machine model to classify colorectal cancer. Moreover, in [20], CNN achieved an exceptional level of performance, 98,7% accuracy, in nine tissue types classifications of colorectal cancer, using the VGG19 model [21], which was pretrained on the ImageNet database [22]. Therefore, to enhance the accuracy of the CFCMC, we employ a Convolutional Neural Network as a feature extractor. We are aware of the problem of losing explainablity of the CNN model by compressing of the data from the feature space to the latent space, which causes that it is hard to track the decision back to the features in the feature space. That is not our case, because we do not provide an explanation about the causality between the features and the decisions but we provide explanation about the classifiability of the data, which is significantly improved by the CNN model by mapping the data from the feature space to the latent space.

Finally, we developed an explanation interface, which provided a semantical and visual explanation that was extracted from the CFCMC classifier that was used to classify the WSIs of the colorectal cancer tissue. We evaluated our XAI (eXplainable Artificial Inteligence) system using common within-subject experimental design [23]; the outcomes from our explanation interface (XAI system) were compared with outcomes from a stand alone CNN (AI system with no explanation) by 14 pathologists at clinical trials through questionnaires.

II. PROPOSED EXPLAINABLE MODEL

In this section, the mathematical description of the Cumulative Fuzzy Class Membership Criterion classifier is explained and followed by the definition of the factor of misclassification and the certainty threshold.

A. Cumulative Fuzzy Class Membership Criterion decision based classifier

The proposed method is based on the assumption that d-dimensional data in the feature space are split into n_c classes, where C_i $(i = 1, \cdots, n_c)$, the i-th class, is divided into n_{cl}^i clusters, where Cl_{ij} $(j = 1, \cdots, n_{cl}^i)$ is the j-th cluster of the i-th class. Each cluster Cl_{ij} comprises training data $\widetilde{p}_{ijk} \in \mathbb{R}^d, (k = 1, \cdots, m_{ij})$, and m_{ij} is the number of training patterns of cluster C_{ij}. Each training pattern $\widetilde{p}$ defines a fuzzy class membership criterion $\kappa_{\widetilde{p}}(\overline{x})$, which is considered as a triangular function as follows:

$$\kappa_{\widetilde{p}_{ijk}}(\overline{x}) = \begin{cases} 1 - \dfrac{\|\widetilde{p}_{ijk} - \overline{x}\|}{a_{ij}} & \|\widetilde{p}_{ijk} - \overline{x}\| < a_{ij} \\ 0 & otherwise \end{cases}, \quad (1)$$

where, $\overline{x} \in \mathbb{R}^d$ is an input vector, a_{ij} is the width of the triangular function for the j-th cluster of i-th class.

Let $(\kappa_{\widetilde{p}_{ij1}}, \ldots, \kappa_{\widetilde{p}_{ijK_{ij}}}, \ldots, \kappa_{\widetilde{p}_{ijm_{ij}}})$ be a desceding order of the fuzzy class membership criterion $\kappa_{\widetilde{p}_{ij}}$ for the j-th cluster of i-th class such that $\kappa_{\widetilde{p}_{ij1}} \geq \cdots \geq \kappa_{\widetilde{p}_{ijm_{ij}}}$, where K_{ij} is the number of first values in the reordering.

Then, the Cumulative Fuzzy Class Membership Criterion (CFCMC) for class C_i is defined as follows:

$$\chi_{C_i}(\overline{x}) = \max_j \left(\frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} \kappa_{\widetilde{p}_{ijk}}(\overline{x}) \right), \quad (2)$$

where, $\chi_{C_i}(\overline{x})$ is the value of CFCMC for an unknown pattern $\overline{x}$ to the class C_i.

Then, the decision rule for winner class CL for the input pattern $\overline{x}$ is as follows:

$$CL(\overline{x}) = C_{\underset{i}{\operatorname{argmax}}(\chi_{C_i}(\overline{x}))}. \quad (3)$$

B. Algorithm description

The algorithm consists of two phases: the initialization and the learning phase. The initialization phase consists of three processes: data splitting, clustering, and parameters initialization. First, input data are divided into three sets: training sets, validation sets, and testing sets. Training patterns are used in Eq. 1 to create a CFCMC decision surface. During the learning phase, the decision surface is optimized (parameters a_{ij} Eq. 1 and $K_{i,j}$ in Eq. 2 are adaptively optimized) in order to cover all validation patterns. The testing set is used for the final evaluation of the created decision surface.

Afterwards, the training data of each class C_i are independently clustered in order to find n_{cl}^i clusters for each class in feature space using the well-known K-means algorithm. The number of clusters, k, is estimated via a gap statistic [24]. This technique uses the output of any clustering algorithm, comparing the change of the within-cluster dispersion with that which is expected under an appropriate reference null distribution. Any other techniques for the estimation of the number of clusters can be used, such as Silhouette analysis [25] or Davies-Bouldin clustering criterion [26].

Next, the parameters a_{ij} and K_{ij} in Eq. 1 and in Eq. 2, respectively, are initialized. These parameters affect the shape of the boundary created by fuzzy class membership criterion $\kappa_{\widetilde{p}}$. Every fuzzy class membership criterion $\kappa_{\widetilde{p}_{ijk}}$ of the j-th cluster of the i-th class shares the same value of parameters a and K. a_{init} is initialized as follows:

$$a_{init} = \frac{1}{n_{\widetilde{p}}} \sum_{ij=1}^{n_{\widetilde{p}}} \min_{j \neq i} \|\widetilde{p}_i - \widetilde{p}_j\|, \quad (4)$$

where, $n_{\widetilde{p}}$ is the number of training patterns. K_{init} value is initialized from the interval $(1; m_{ij})$. The value of threshold θ is set from the interval $(0; 1)$. If the value of CFCMC $\chi(\overline{x})$ of the input pattern $\overline{x}$ is below the threshold θ, thus $\chi(\overline{x}) < \theta$, the pattern is "not classified". Finally, the CFCMC surface is computed using Eq. 1 and Eq. 2.

During the learning phase, adjustment of the CFCMC surface's shape occurs in order to obtain the highest classification accuracy. An assumption of dividing training set into n_c classes and each class C_i into n_{cl}^i clusters $Cl_{i,j}$ generates a set of vectors

$$\overline{p}_i = [a_{i1}, K_{i1}; \cdots; a_{ij}, k_{ij}; \cdots; a_{in_{cl}^i}, k_{in_{cl}^i}], \quad (5)$$

where, $i = 1, \cdots, n_c$. Optimizing of $\overline{p}_i$, the CFCMC surface's shape is adjusted. Any optimizing algorithm can be used, for instance, simulated annealing or hill-climbing methods. We decided to employ a well-known genetic algorithm [27]. The fitness function is defined as follows:

$$Minimize : err\left(S_{valid}\right), \quad (6)$$

where, S_{valid} is the validation set and $err\left(S\right)$ is the error rate evaluated from the data set S as follows:

$$err(S) = \frac{\#\text{of incorrectly classified samples}}{\#\text{of all samples}}. \quad (7)$$

C. Factor of misclassification

The term factor of misclassification *(FoM)* is described as *"the likehood of the input sample, which is assigned to the cluster Cl belonging to the class C, to be misclassified to one of the rest of the classes"* i.e. the possibility that in reality the observation belongs to another class. The factor of misclassification of the input sample, assigned to the cluster Cl_{Ai}, to the reference cluster Cl_{Bj} is defined as follows:

$$FoM(\overline{x}, Cl_{Bj}) = \frac{\chi_{Cl_{Bj}}(\overline{x})}{\chi_{Cl_{Ai}}(\overline{x})} + sim_{Cl}(Cl_{Ai}, Cl_{Bj}), \quad (8)$$

where the first term on thee right hand side describes the *local similarity* as the ratio between memberships of the input sample $\chi(\overline{x})$ to the reference cluster Cl_{Bj} and the winner cluster Cl_{Ai}. The second term on the left hand side describes *global similarity*, which is based on the relationship between the data's clusters.

The similarity between the two clusters Cl_{Ai} and Cl_{Bj} is defined as follows:

$$sim_{Cl}(Cl_{Ai}, Cl_{Bj}) = \frac{A_{intersection(Cl_{Ai}, Cl_{Bj})}}{A_{Cl_{Ai}}}, \quad (9)$$

where A_{Cl} is the area of a hypersphere describing cluster Cl and $A_{intersection}$ is the area of intersection of the two clusters. Here, for simplification, the clusters are described with n-dimensional hypersphere, where n is the data dimensionality. For the center and the radius of a hypersphere, the coordinates of a cluster's centroid $c_{Cl_{ij}}$ and the estimated variance $\hat{\sigma}_{Cl_{ij}}$ of a cluster's data, respectively, are used. For computational purposes, n-hypersphere is transferred into a two dimensional circle. The area of intersection between the two clusters is computed using simple two dimensional trigonometry by using the distance between the centers and radiuses of the circles, and thus, the Euclidean distance $\left\| c_{Cl_{Ai}} - c_{Cl_{Bj}} \right\|$ between the centroids of the clusters Cl_{Ai} and Cl_{Bj} and their estimated variances $\hat{\sigma}_{Cl_{Ai}}$ and $\hat{\sigma}_{Cl_{Bj}}$.

The equation for the estimation of the cluster's variance value $\hat{\sigma}_{Cl_{A,i}}$ was derived in [19] and it is calculated as follows:

$$\hat{\sigma}_{Cl_{ij}} = a_{Cl_{ij}} \left(\frac{k}{\chi_{Cl_{ij}^{max}}} \right)^m \begin{cases} m = 0.7 & \chi_{Cl_{ij}^{max}} \le k \\ m = 2.5 & \chi_{Cl_{ij}^{max}} > k \end{cases},$$
$$k = p_1 * K + p_2$$
$$p_l = a_l * dim^{b_l} + c_l \quad (l = 1, 2)$$
$$a_1 = -0.7621, b_1 = -0.2799, c_1 = 0.0746$$
$$a_2 = 0.8372, b_2 = -0.3729, c_2 = 0.1758$$
(10)

It should be noted that during the variance estimation, the Euclidean distance was replaced with the following distance measure d: Let's have two vectors x_A and x_B with n-dimensionality. Then the distance d is defined as follows:

$$d = \frac{1}{n} \sum_{i=1}^{n} |x_{Ai} - x_{Bi}|,$$
(11)

The value of the factor of misclassification of the input $\overline{x}$ to the i-th class C_i is computed as follows:

$$FoM(\overline{x}, C_i) = \max_j FoM(\overline{x}; Cl_{ij})$$
(12)

The factor of misclassification can also be expressed semantically. It exhibits the values as follows:

$$DFoM \begin{cases} no \quad possibility & FoM(\overline{x}, C_A) \in (0, c_\theta) \\ low \quad possibility & FoM(\overline{x}, C_A) \in (c_\theta, \theta_{mid}] \\ high \quad possibility & FoM(\overline{x}, C_A) \in (\theta_{mid}, FoM_{max}] \end{cases}$$
(13)

where c_θ is the certainty threshold and $\theta_{mid} = (FoM_{max} - c_\theta)/2 + c_\theta$. The FoM_{max} is the maximum value of the FoM computed using the validation samples as follows:

$$FoM_{max} = \max_j \max_i FoM(\overline{x}_j, C_i),$$
(14)

where $\overline{x} \in S_{valid}$.

D. Certainty threshold and certain prediction

The certainty threshold is the value of the FoM, below which it is certain (i.e. there is no possibility) that the input sample $\overline{x}$ is assigned to the class C_A and will not be misclassified to any other classes in the feature space.

Let $\overline{x}$ be samples from the validation set S_{valid} that were misclassified and C_{GT}^k be the ground truth label of the k-th sample $\overline{x}_k$. Then the certainty threshold c_θ is calculated as follows:

$$c_\theta = \min_k FoM(\overline{x}_k; C_{GT}^k).$$
(15)

It follows that if $FoM(\overline{x}, C_i) < c_{FoM}$ holds, it is unlikely that input sample $\overline{x}$ belongs to the class C_i.

Therefore, if it holds that $\forall i_{\in 1, \cdots, n_c}; FoM(\overline{x}, C_i) < c_\theta$, the prediction of the input sample $\overline{x}$ is *certain*, otherwise it is *uncertain*.

E. Representative training sample

Representative training sample $\widetilde{p}^r$ is the training sample that is the most responsible for assigning input sample $\overline{x}$ to the j-th cluster Cl_{ij} of the i-th class C_i. It is computed as follows:

$$\widetilde{p}^r(\overline{x}) = \underset{k}{\operatorname{argmin}} \|\widetilde{p}_{ijk} - \overline{x}\|$$
(16)

where $\|\widetilde{p}_{ijk} - \overline{x}\|$ is the Euclidean distance between $\overline{x}$ and $\widetilde{p}_{ijk}$.

III. COLORECTAL CANCER DETECTION EXPLANATION INTERFACE

This section describes the explanations generated to human experts by the proposed system in colorectal cancer detection tasks. To evaluate the usefulness of the proposed explanations, we designed two systems; first a plain CNN model that only generates decisions without any explanation, second, an X-CFCMC (eXplainable CFCMC) model that complements its decisions with explanations.

For both systems, we developed similar user interfaces; both provide classification results of the whole-slide images (WSI) of colorectal cancer tissue, showing the original image of the WSI and a corresponding label map with a colour code of eight different tissue types. Pathologist can examine an arbitrary area of the WSI by clicking on the desired area. Subsequently, the interfaces show their predictions. Finally, a pathologist can provide the final decision by selecting one of the eight buttons representing the eight different tissue types.

The *non-explainable AI system* interface that uses the plain CNN model is shown in Fig. 1. It provides predicted type of tissue and a probability distribution of the prediction of all tissue types, which was computed in the output layer of the CNN model with the softmax activation function.

The *explainable AI system* interface that uses the X-CFCMC model is shown in Fig. 2. It complements its decision with three types of information:
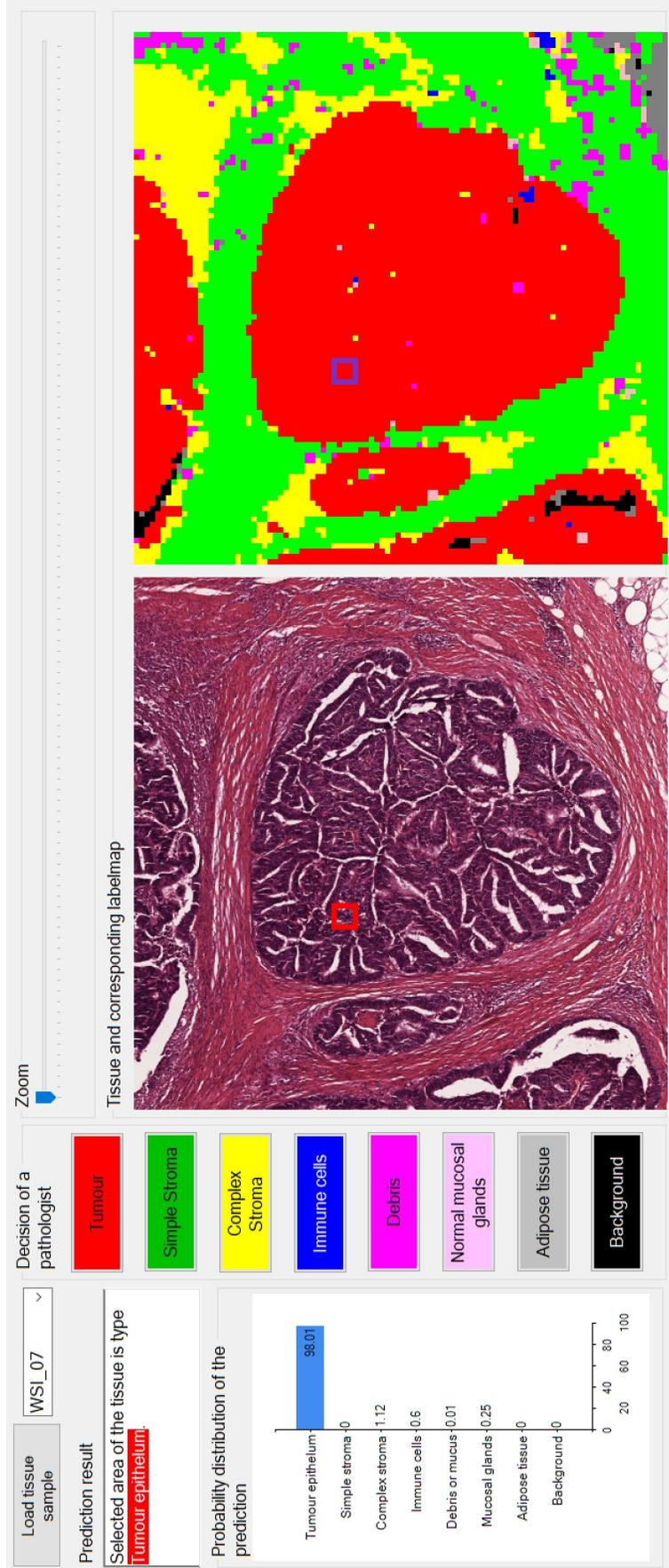
Fig. 1: The interface for the plain CNN, without any explanation, for presenting prediction on histopathological WSI. From the left hand side, it shows a prediction result and the probability distribution of the prediction. Next, eight buttons for making a final decision by a pathologist are located there. Finally, the original image of the WSI and the corresponding label map are visualized.

1) Semantical explanation: To provide user-friendly explanations, prediction results and information regarding the possibility of the misclassification of the examined area of the WSI are semantically explained, for example, with the phrase "there is a low possibility that this classification is wrong". This semantical explanation is generated based on the value of the FoM.

2) Visualization of the training image most responsible for a given prediction: To justify the prediction result, the most responsible training image for a given prediction is displayed to the user. A similar approach to understand the predictions was introduced in [28], where the influence function was used to trace the CNN's prediction through the learning algorithm and back to its training data. In our approach, the most responsible training image for a prediction is a representative training sample $\widetilde{p}^r(\overline{x}_w)$ for the input image $\overline{x}_w$ assigned to the winner tissue type w. It follows that if the training image has very similar context with the input image, it should gain the trust of the pathologist in prediction. Otherwise, the pathologist could consider the certain prediction as not being reliable.

3) Visualization of training images of other types of tissue: The third means of explanation shows the representative training images for the input image to the other tissue types, into which the input image could be misclassified with a high or low possibility. It should visually explain to the pathologist why the input sample could be misclassified to a particular tissue type. In the case of similar context, a pathologist could consider the particular tissue type as the true type of tissue.

IV. EXPERIMENTS

To choose the best performing explainable model to classify colorectal cancer image data, the model has to be accurate and reliable with its explanations. Therefore, in this section, first, we present the results from the task of boosting the performance of the CFCMC classifier. Then, we describe the results of validating the *certainty threshold*. Finally, we show examples of generated explanations.

A. Histopathological data description

We used a publicly available dataset released in [2] by *Kather et al.*. It consists of Hematoxylin and Eosin (H&E) tissue slides, which were cut into 5000 small tiles of the size 150x150 pixels (equivalent to 74 μm 74 μm), each of them annotated with one of eight tissue classes, namely tumour epithelium, simple stroma (homogeneous composition, includes tumour stroma, extra-tumoural stroma and smooth muscle), complex stroma (containing single tumour cells and/or few immune cells), immune cells (including immune-cell conglomerates and sub-mucosal lymphoid follicles), debris (including necrosis, haemorrhage and mucus), normal mucosal glands, adipose tissue, background (no tissue). The data are class-balanced, each of the classes consists of 625 tiles.

B. Boosting the CFCMC's performance

The focus of the first part of experiments was to find the architecture of the CNN with the best performance as a feature extractor to train the explainable CFCMC classifier.

1) Experimental setup: We utilized eight well-known CNN models, pre-trained on ImageNet [22] dataset, specifically, *AlexNet* [29], *VGG-16* [21], *Inception-v3* [30], *ResNet-50* [31], *Xception* [32], *DenseNet121* [33], *Inception-ResNet-V2* [34] and *EfficientNet0* [35]. For all of them, the fully connected layers were cut and replaced with a dense layer containing 1024 neurons with ReLU activation functions and with an output layer containing 8 neurons with the softmax activation function.

Moreover, we created three lighter architectures that were trained from scratch, specifically a *VGG-like* model with 12 convolutional layers and 2 fully-connected layers, an *Inception-like* model with 3 inception layers and a *ResNet20* model with a depth 20. All models were trained using the Adam [36] optimizer to minimize cross-entropy for 100 epochs with the learning rate set to the value 0.0001.

Finally, to train the explainable CFCMC classifier, we extracted the features from the last dense layer of all the CNN models. The experimental setup for the CFCMC algorithm is as follows: number of clusters for each of the classes was set to 1. The value of the threshold θ was set to value $\theta = 0.01$. For the optimization of the CFCMC, MATLAB implementation of the genetic algorithm was used with a population size of 50 individuals. The mutation rate was set to 0.2. Arithmetic crossover and adaptive feasible mutation operators were used for reproduction. Stochastic uniform selection was used to choose parents for the next generation. The algorithm stops at the 30th generation.

2) Experimental results: Table I provides the performance results of different CNN models and the corresponding CFCMC models. The classification results are evaluated with a 10-folds cross validation test. Because of the class-balanced dataset, the accuracy metric was chosen to evaluate the performance.

From Table I, it can be observed that pre-trained and fine-tuned models outperform the ones trained from scratch. Moreover, it can be seen that features extracted from the CNN models significantly boost the performance of the explainable CFCMC models.

C. Validation of the certainty threshold

To validate the certainty threshold, three metrics were defined: certainty rate, certainty error and ground truth label certainty error.

The *certainty rate* c_r is defined as the ratio between the number of certain predictions y_c^* and the number of all predictions y^*.

$$c_r = \frac{\#y_c^*}{\#y^*} \tag{17}$$

The *certainty error* is defined as follows:

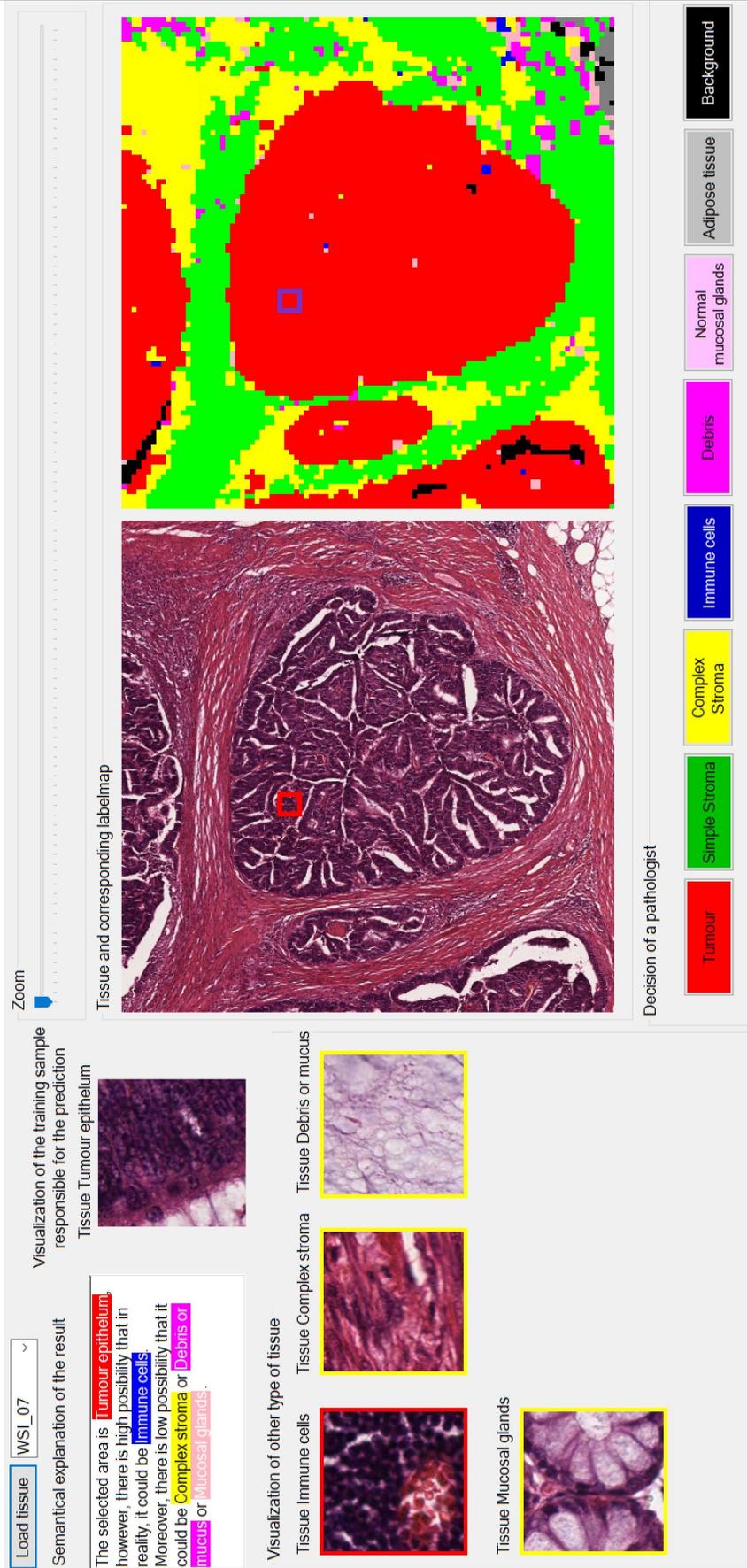$$c_e = \frac{\#\widetilde{y}_c^*}{\#y_c^*} \tag{18}$$

Fig. 2: The interface for the explainable system (X-CFCMC system) for presenting predictions with additional explanations of the CFCMC. From the left hand side, it shows a semantical explanation of the results alongside the visualization of the training sample responsible for the prediction, below which there is a visualization of the other tissue types, which could potentially be the true tissue type. Next, the original image of the WSI and the corresponding label map are visualized, below which eight buttons for making a final decision by a pathologist are located.

	CNN model	CFCMC model
Raw image	—	59.35(3.43)
AlexNet	91.43(2.42)	85.32(3.26)
VGG-16	93.61(1.75)	92.42(2.02)
Inception-v3	92.76(1.44)	90.79(2.04)
ResNet-50	**93.80(1.08)**	91.28(1.64)
Xception	93.58(1.25)	**92.78(1.74)**
DenseNet121	92.76(1.29)	92.06(1.75)
Inception-ResNetV2	92.76(1.02)	91.44(1.95)
EfficientNet0	90.97(1.39)	85.84(5.83)
VGG-like	80.88(1.14)	80.21(2.14)
Inception-like	85.25(2.82)	83.97(3.04)
ResNet20	90.14(1.24)	84.01(5.02)

TABLE I: The performance results of different CNN models and the corresponding CFCMC models

where $\widetilde{y}_c^*$ are certain predictions that were misclassified.

The *ground truth label certainty error* c_e^{GT} is defined as follows:

$$c_e^{GT} = \frac{\#\widetilde{y}_{GT_c}^*}{\#\widetilde{y}^*} \qquad (19)$$

where $\widetilde{y}^*$ are misclassified predictions and $\widetilde{y}_{GT_c}^*$ are misclassified predictions, which for their ground truth label C_{GT} holds that $\phi(\overline{x}; C_{GT}) < c_\theta$, i.e. their ground truth label is unlikely to be true label.

Table II provides the results from the three previously defined metrics computed on predictions from joined testing sets from 10 folds for each CFCMC model. For each metric, the best performing models are highlighted in bold.

CFCMC models trained on features that were extracted from pre-trained CNN models generally outperformed models trained from scratch with a certainty rate c_r (13.99% against 3.41% in average). All of the models, however, achieved very low certainty error c_e; 7 models achieved zero error, while the highest value was 2.17%. It follows that when a classifier labels its prediction as *certain*, it is unlikely that this prediction will be incorrect.

Moreover, all of the models likewise reached a low value of *ground truth label certainty error* c_e^{GT} (lower than 5%). This implies that when a classifier labels its prediction as *uncertain*, it is very likely that ground truth label will appear among the potentially true labels.

D. Explanations examples

Fig. 3 shows five examples of explanations with different difficulty levels of classification of the input image. If an explanation offers tissue types with a high probability of misclassification, the input image is hard to classify. In case of only low probable misclassification offer, the input image is not easy to classify. Finally, in cases that it provides no offer, the input image is easy to classify, thus, the prediction is certain. To generate explanations, the CFCMC model trained of features extracted from *DenseNet121* CNN architecture was used.

Fig. 3a illustrates an example of a prediction explanation for the input image, which is easy to classify because it offers no other tissue types. Moreover, it can be seen that the input image is very similar to the training image responsible for the prediction (TRP). Therefore, the prediction is certain. The following semantic explanation was extracted: **The input image is for sure *Tumour epithelium*, because it could no be misclassified to any other tissue types.**

Fig. 3b shows a prediction explanation for an input image that is not so easy to classify but was correctly classified. The input image was classified as *Immune cells* tissue type. Although it offers three tissue types with a low probability of misclassification, namely *Tumour epithelium*, *Simple stroma* and *Complex stroma*, the input image is very similar to a TRP image. Therefore, the TRP image could gain trust in this prediction. The following semantic explanation was extracted: **The input image is *Immune cells* tissue type. However, there is a low possibility that in reality it could be *Tumour epithelium* or *Simple stroma* or *Complex stroma*.**
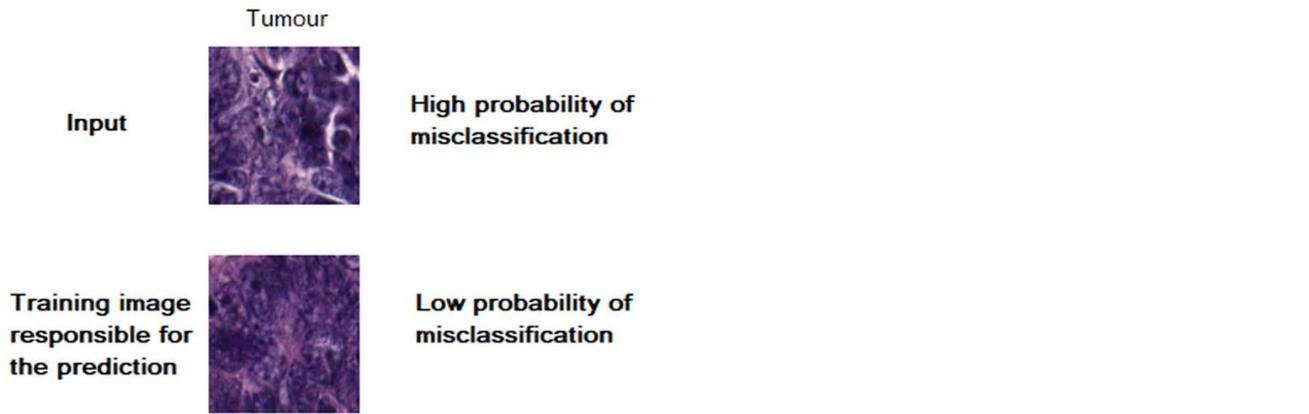
Fig. 3c shows a prediction explanation for an input image that is not easy to classify and was misclassified. The input image was predicted as *Adipose tissue type*. The explanation offers two tissue types with low probability, namely *Simple stroma* and *Debris or mucus*. It can be seen that the input image is more similar to low probable tissue types than the TRP image. This could lower trust in this particular decision. However, it offers the true tissue type of the input image, which is *Debris or mucus*. The following semantic explanation was extracted: **The input image is *Adipose* tissue type. However, there is a low possibility that in reality it could be *Simple stroma* or *Debris or mucus*.**

Fig. 3d illustrates a prediction explanation for an input image that is hard to classify, because in addition to three tissue type with a low probability for misclassification, it offers three highly probable tissue types. Therefore, the expert should investigate the input image more deeply. The following semantic explanation was extracted: **The input image is *Immune cells* tissue type. However, there is a high possibility that in reality it could be *Tumour epithelium* or *Complex stroma* or *Mucosal glands*. Moreover, there is a low possibility that it could be *Simple stroma* or *Debris or mucus* or *Adipose tissue*.**
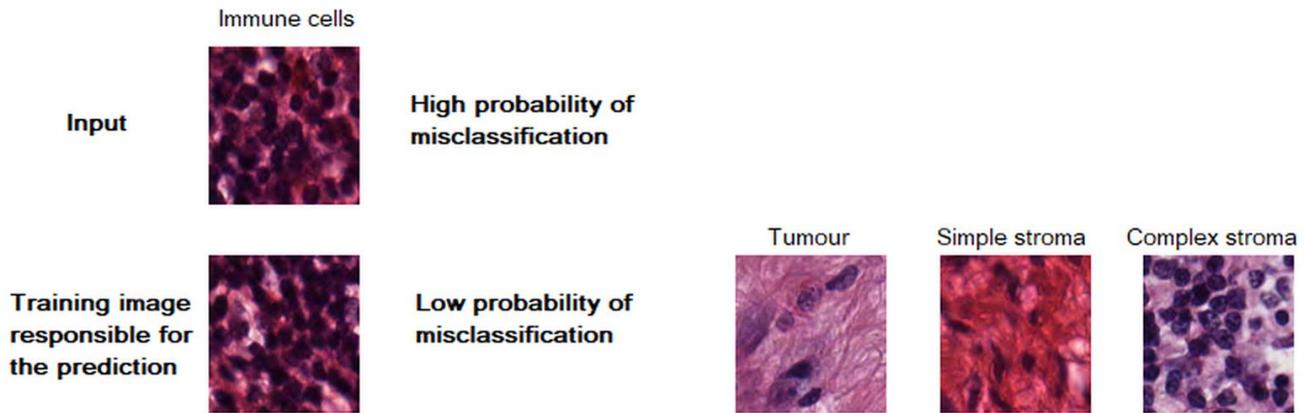
Fig. 3e shows a prediction explanation for an input image

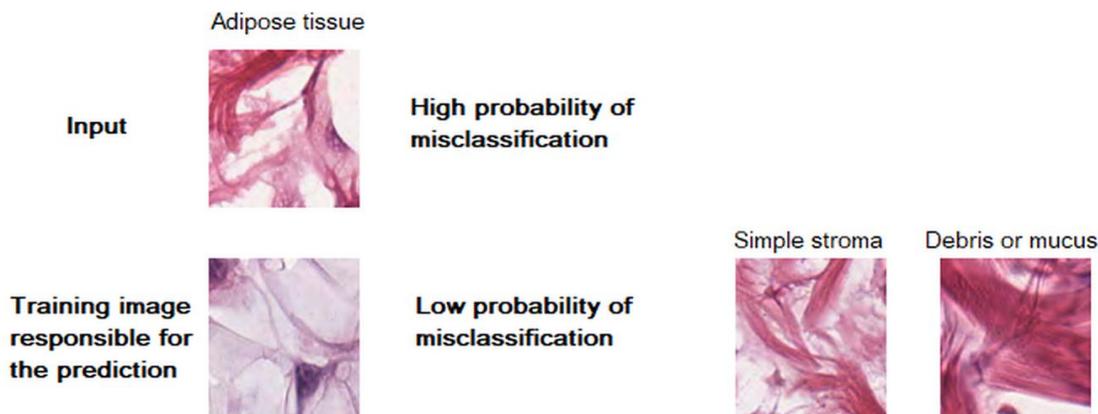| CFCMC models	$\#y^*$	$\#\widetilde{y}^*$	err	$\#y_c^*$	c_r	$\#\widetilde{y}_c^*$	c_e	$\#\widetilde{y}^*_{GT_c}$	c_e^{GT}
AlexNet	5000	734	14.68%	538	10.76%	**0**	**0.00%**	6	0.82%
VGG16	5000	379	7.58%	646	12.92%	1	0.15%	5	1.32%
Inception-v3	5000	460	9.20%	513	10.26%	**0**	**0.00%**	7	1.52%
ResNet-50	5000	436	8.72%	589	11.78%	**0**	**0.00%**	13	2.98%
Xception	5000	**361**	**7.22%**	**1368**	**27.36%**	2	0.20%	15	4.16%
DenseNet121	5000	397	7.94%	648	12.96%	**0**	**0.00%**	**4**	**1.01%**
Inception-ResNetV2	5000	428	8.56%	612	12.24%	**0**	**0.00%**	8	1.87%
EfficientNet0	5000	707	14.14%	680	13.60%	3	0.44%	**4**	**0.57%**
VGG-like	5000	990	19.80%	46	0.92%	1	2.17%	7	0.71%
Inception-like	5000	801	16.02%	291	5.82%	**0**	**0.00%**	6	0.75%
ResNet20	5000	800	16.00%	175	3.50%	**0**	**0.00%**	**4**	**0.50%**

TABLE II: The results from three metrics for validation of the certainty threshold, certainty rate c_r, certainty error c_e, ground truth label error c_e^{GT} with corresponding occurrences

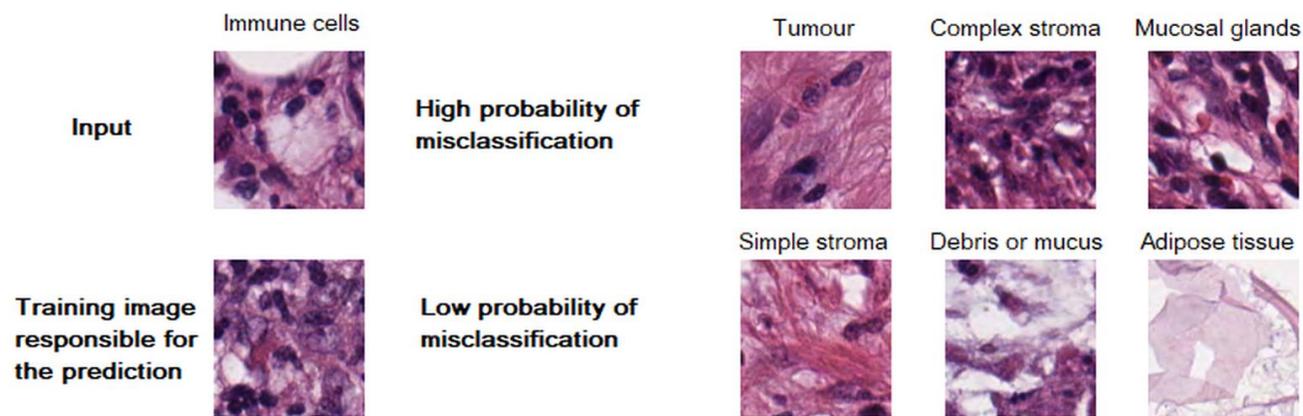(a) The input image is easy to classify - the prediction is certain.

(b) The input image is not so easy to classify, however, it was correctly classified.
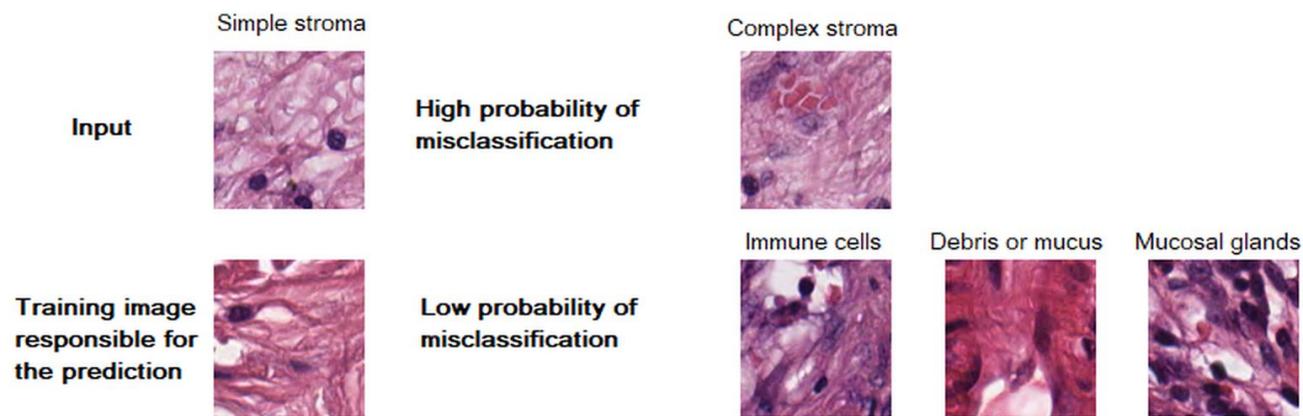

(c) The input image is not so easy to classify and was wrongly classified. The true tissue type is *Debris or mucus*

Fig. 3: Examples of the explanations generated for the testing dataset. The tissue types are displayed above the image.

(d) The input image is hard to classify, however, it was correctly classified

(e) The input image is hard to classify and was wrongly classified. The true tissue type is *Complex stroma*

Fig. 3: *(Continued)*

that is hard to classify and was misclassified. It was predicted as *Simple stroma* tissue type. The explanation offers three tissue types with low and one with high probability of misclassification. It can be seen that in this case, input image is the most similar to the high probable tissue type, *Complex stroma*, which is also the true tissue type. The following semantic explanation was extracted: **The input image is *Simple stroma* tissue type. However, there is a high possibility that in reality it could be *Complex stroma*. Moreover, there is a low possibility that it could be *Immune cells* or *Debris or mucus* or *Mucosal glands*.**

V. CLINICAL TRIALS RESULTS

To evaluate the influence of the explanation generated from X-CFCMC for human pathologists, we ran an acceptability test against the plain CNN. The objective of this experiment was to evaluate the acceptability of the explanation-generating X-CFCMC for human pathologists. We used within-subject experimental design, thus at the clinical trials both systems were shown to 14 pathologists (3 men and 11 women), with

an an average age of 40.7 years and an average length of service of 14.9 years in which the shortest length of service was 4 years and the longest was 45 years. At the end of the session, feedback from the pathologists was collected in the form of questionnaires.

A. Experiment setting

Prior to the experiments, each participant was informed about the dataset. None of them were familiar with machine learning concepts. Therefore, each of the participant was informed about the automatic classification of histopathological samples by machine learning. Afterwards, both interfaces were explained to the participants, including the controls and the means of presenting the predictions. Finally, it was explained to the participants that with the exception of asking for help with controls, dialogue with the interviewers was discouraged. One different classified WSI from dataset was shown to every participant, who they were asked to examine 20 arbitrary area for both interfaces and evaluate each prediction outcome. This takes approximately 30 minutes on average. At the end of

the experiment session, every participant was asked to fill out questionnaire.

B. Evaluation on users experiences

The users' experiences in using the stand alone CNN and X-CFCMC were evaluated using a questionnaire. The internal consistency of the questionnaire reached value of the Cronbach's alpha, $\alpha = 0.89$. Therefore, we can state that the participants sufficiently understood the objectives of the experiment. The questionnaire was divided into three parts.

The first and the second parts use a semantic differential scale, which presents respondents with a set of bipolar scales (useful/useless, reliable/unreliable). Respondents were asked to choose a number (from 1 to 6) that indicates the extent to which the adjectives relate to a characteristic evaluation of the stand alone CNN and X-CFCMC systems. While 1 represents a positive adjective (f.e. useful), 6 represents a negative adjective (f.e. useless). The adjectives were selected to cover four evaluation parameters of the systems' influences on trust and reliance to the pathologists:

1) **objectivity** - objective / subjective, useful / useless, relevant / irrelevant, serious / unserious, ethically / unethically
2) **details** - precise / imprecise, consistent / inconsistent, complicated / uncomplicated, complete / incomplete
3) **reliability** - accurate / inaccurate, faultless / faulty, straight / misleading, certain / uncertain, reliable / unreliable
4) **quality** - systematic / unsystematic, time saving / time consuming, clear / unclear, expert / inexpert, good quality / bad quality

Table III shows the average scores of the four evaluation parameters for both systems, which are computed as the arithmetic mean of the chosen number of each corresponding bipolar scale, while the total average is computed from the average scores of each evaluation parameter. Better values are highlighted in bold. Because the lower the score means the better the evaluation of a certain characteristic, the results reveal that the *X-CFCMC system* obtained better average scores in all parameters. The most significant differences between the scores were in the level of details (0.31) and the reliability (0.24).

The third part of the questionnaire used dichotomous scale. It consisted of closed-ended items that covered a subjective evaluation of both interfaces so that the participants express their agreement or disagreement with the statements. The statements were created to be focused on the evaluation the truthfulness and usefulness of both systems.

Analyzing the dichotomous scaled items of the third part of the questionnaires, we came to the following findings:

a) X-CFCMC system: The semantic explanation **influenced the trust** of 10 pathologists, while 9 of them **increased** and 1 **decreased** their trust in the prediction. *The visualization of the training image responsible for the prediction* **increases trust** of 10 of the pathologists while for the rest there was no influence. *The visualization of the other types of tissue*

Evaluation parameters	Average score	
	plain CNN	X-CFCMC
Objectivity	1.75	**1.65**
Level of Details	2.16	**1.85**
Reliability	2.81	**2.57**
Quality	2.10	**1.99**
Total average	2.21	**2.01**

TABLE III: A comparison of the average scores for each of the four evaluation ares for both systems, the stand alone CNN model without providing explanations and the X-CFCMC model that explains its decision

influenced only half of the pathologists, however, it increased their trust of the system.

b) Stand alone CNN system: The probability distribution of the prediction had no influence to 10 of the pathologists. For the rest, it increased their trust in the prediction.

c) Comparison of the systems: Analyzing five items devoted to a direct comparison of the usefulness of both systems, **the X-CFCMC system** achieved *a cumulative score of 50*, while **the plain CNN system** achieved *a cumulative score of 20*.

d) Credibility of the systems: Comparing two items about the credibility of both systems, **the plain CNN system** achieved *a cumulative score of 23*, while **the X-CFCMC system** achieved *score of 21*.

e) Usefulness of the whole-slide segmentation: Two items showed that, in general, all pathologists consider an automatic whole-slide segmentation of the histopathological samples useful.

C. Discussion

From the results above, key findings emerge. The comparison of the characteristic of both systems revealed that in the level of details, the pathologist consider the X-CFCMC system as more rigorous, more precise, more consistent, more complete. Moreover, the X-CFCMC system is considered as more accurate, reliable and confident regarding its predictions.

The statement evaluation indicates that the most useful means of explanations are *semantic explanation* and *a visualization of the training image responsible for the prediction*. *Visualization of the other types of tissue* was only appreciated by half of the pathologists. A direct comparison of both systems indicates that *the X-CFCMC system* is more acceptable than the *plain CNN system*.

VI. CONCLUSION

In this study, we extend the explainability of the explainable Cumulative Fuzzy Class Membership Criterion (CFCMC) classifier and used it for classification of eight tissue type from histopathological cancer image samples.

First, we boosted the performance of the CFCMC classifier on colorectal image data using a fine-tuned Convolutional Neural Network (CNN) as a feature extractor, which was pre-trained on different dataset.

Next, we defined the *factor of misclassification (FoM)*, which is able to estimate the possibility of the input sample being misclassified to a particular conflicting class. Moreover, we defined the *certainty threshold*, thanks to which we are able to say, whether the prediction is certain or uncertain. The proposed uncertainty measure is significantly different from many other uncertainty measures of models, e.g. neural networks model, firstly, because it is based on the classifiability of the data in the space, where the classifier makes its decision, and secondly, in the case of uncertain prediction, it is able to suggest the classes in which the input sample could be misclassified. Thus, it offers relevant classes to be further examined. The experiments clearly supported this ability.

Finally, we developed two systems for segmentation of the whole-slide images of histopathological cancer tissue. The first system used stand alone CNN and the second used an X-CFCMC classifier, which provides three means of explanations: a semantic explanation about the prediction and possible misclassification, a visualization of the training image responsible for the prediction and a visualization of the other types of tissue.

At the clinical trials with 14 pathologists, we measured the acceptability and the trust of the pathologists for the proposed system. The results indicate that the X-CFCMC system is more useful and more reliable than the plain CNN.

In conclusion, this paper discussed about the usability and the reliability of a explainable classifier in real world medical settings through clinical trials. We believe that our proposed system can contribute to the use of AI, especially by improving the usability and acceptability of AI systems in medical domains, where speed in making decisions, reliability and accountability are crucial. We are aware that the scale of our preliminary experiment was limited. The expansion of clinical trials to include more pathologists from various fields are of our immediate future interest. Moreover, in the medical settings we will be confronted with imbalanced, heterogeneous and inaccurate data sets. Therefore, our next challenge in our research is also to examine how our classifier will perform on imperfect data.

References

[1] Juanying Xie, Ran Liu, IV Luttrell, Chaoyang Zhang, et al. Deep learning based analysis of histopathological images of breast cancer. *Frontiers in genetics*, 10:80, 2019.

[2] Jakob Nikolas Kather, Cleo-Aron Weis, Francesco Bianconi, Susanne M Melchers, Lothar R Schad, Timo Gaiser, Alexander Marx, and Frank Gerrit Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6:27988, 2016.

[3] Babak Ehteshami Bejnordi, Guido Zuidhof, Maschenka Balkenhol, Meyke Hermsen, Peter Bult, Bram van Ginneken, Nico Karssemeijer, Geert Litjens, and Jeroen van der Laak. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *Journal of Medical Imaging*, 4(4):044504, 2017.

[4] Jonathan de Matos, Alceu de Souza Britto Jr, Luiz ES Oliveira, and Alessandro L Koerich. Histopathologic image processing: A review. *arXiv preprint arXiv:1904.07900*, 2019.

[5] Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16:34–42, 2018.

[6] Teresa Araújo, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. Classification of breast cancer histology images using convolutional neural networks. *PloS one*, 12(6):e0177544, 2017.

[7] Miriam Hägele, Philipp Seegerer, Sebastian Lapuschkin, Michael Bockmayr, Wojciech Samek, Frederick Klauschen, Klaus-Robert Müller, and Alexander Binder. Resolving challenges in deep learning-based analyses of histopathological images using explanation methods. *arXiv preprint arXiv:1908.06943*, 2019.

[8] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. What do we need to build explainable ai systems for the medical domain? *arXiv preprint arXiv:1712.09923*, 2017.

[9] A Holzinger, G Langs, H Denk, K Zatloukal, and H Mueller. Causability and explainability of ai in medicine. *Data Mining and Knowledge Discovery doi*, 10, 2019.

[10] Andreas Holzinger, Markus Plass, Michael Kickmeier-Rust, Katharina Holzinger, Gloria Cerasela Crişan, Camelia-M Pintea, and Vasile Palade. Interactive machine learning: experimental evidence for the human in the algorithmic loop. *Applied Intelligence*, 49(7):2401–2414, 2019.

[11] Gabriel R Vásquez-Morales, Sergio M Martínez-Monterrubio, Pablo Moreno-Ger, and Juan A Recio-García. Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning. *IEEE Access*, 7:152900–152910, 2019.

[12] James Mullenbach, Sarah Wiegreffe, Jon Duke, Jimeng Sun, and Jacob Eisenstein. Explainable prediction of medical codes from clinical text. *arXiv preprint arXiv:1802.05695*, 2018.

[13] Scott M Lundberg, Bala Nair, Monica S Vavilala, Mayumi Horibe, Michael J Eisses, Trevor Adams, David E Liston, Daniel King-Wai Low, Shu-Fang Newman, Jerry Kim, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature biomedical engineering*, 2(10):749, 2018.

[14] Avleen Malhi, Timotheus Kampik, Husanbir Pannu, Manik Madhikermi, and Kary Främling. Explaining machine learning-based classifications of in-vivo gastral images. In *2019 Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–7. IEEE, 2019.

[15] Pitoyo Hartono. A transparent cancer classifier. *Health Informatics Journal*, 26(1):190–204, 2020. PMID: 30596318.

[16] Łukasz Raczkowski, Marcin Możejko, Joanna Zambonelli, and Ewa Szczurek. Ara: accurate, reliable and active histopathological image classification framework with bayesian deep learning. *bioRxiv*, page 658138, 2019.

[17] Patrik Sabol, Peter Sinčák, Kana Ogawa, and Pitoyo Hartono. Explainable classifier supporting decision-making for breast cancer diagnosis from histopathological images. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2019.

[18] P. Sabol, P. Sinčák, J. Buša, and P. Hartono. Cumulative fuzzy class membership criterion decision-based classifier. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 334–339, Oct 2017.

[19] Patrik Sabol, Peter Sinčák, Jan Magyar, and Pitoyo Hartono. Semantically explainable fuzzy classifier. *International Journal of Pattern Recognition and Artificial Intelligence*, 33(12):2051006, 2019.

[20] Jakob Nikolas Kather, Johannes Krisam, Pornpimol Charoentong, Tom Luedde, Esther Herpel, Cleo-Aron Weis, Timo Gaiser, Alexander Marx, Nektarios A Valous, Dyke Ferber, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1):e1002730, 2019.

[21] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[22] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE*

conference on computer vision and pattern recognition, pages 248–255. Ieee, 2009.

[23] Sina Mohseni, Niloofar Zarei, and Eric D Ragan. A survey of evaluation methods and measures for interpretable machine learning. *arXiv preprint arXiv:1811.11839*, 2018.

[24] Robert Tibshirani, Guenther Walther, and Trevor Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.

[25] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.

[26] David L Davies and Donald W Bouldin. A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, (2):224–227, 1979.

[27] John Henry Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[28] Pang Wei Koh and Percy Liang. Understanding black-box predictions via influence functions. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1885–1894. JMLR. org, 2017.

[29] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[30] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[31] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[32] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.

[33] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[34] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*, 2017.

[35] Mingxing Tan and Quoc V Le. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*, 2019.

[36] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Štúdium štruktúrnej relaxácie polymérnych materiálov metódami NMR

Alojz Šoltýs

Katedra fyziky
FEI TUKE
Košice, Slovenská republika
alojz.soltys@tuke.sk

Jana Tóthová

Katedra fyziky
FEI TUKE
Košice, Slovenská republika
jana.tothova@tuke.sk

Abstrakt—**Biodegradovateľné polymérne materiály na báze škrobu sú jednou z alternatív riešenia problému znečisťovania životného prostredia plastovým odpadom. Výhodami uvedených materiálov je dostupnosť, cena a obnoviteľnosť základného materiálu. Príprava termoplastických škrobov zahrňuje termo-mechanickú úpravu v prítomnosti plastifikátora, pri ktorej sa rozpadne pôvodná štruktúra škrobových granúl a vzniká homogénny materiál s vlastnosťami podobnými termoplastom. Dôsledkom tohto procesu je systém zároveň vychýlený zo stavu termodynamickej rovnováhy a začína štruktúrne relaxovať. Dynamika relaxačných procesov závisí od množstva faktorov, akými sú druh a množstvo plastifikátora, prímesi v materiáli či chemická modifikácia škrobu. Štúdium relaxačných procesov na kukuričnom škrobe plastifikovanom zmesou s rôznym podielom glycerínu a močoviny bola pozorovaná pomocou ^{1}H a ^{13}C NMR v tuhej fáze a metódou WAXS.**

Kľúčové slová—škrob, glycerín, močovina, štruktúrna relaxácia, NMR, WAXS

Abstract—**Starch-based biodegradable polymers are promising alternatives to deal with environmental pollution by plastic waste. The advantages of such materials are availability, price, and renewability of the raw material. Thermoplastic starches preparation requires thermo-mechanical treatment in the presence of a plasticizer where starch granules disruption occurs in order to create homogenous thermoplastic-like material. Due to this process, thermodynamic equilibrium is disrupted and the system starts to relax. The dynamics of relaxation processes depends on factors like selection and amount of plasticizer, additives in material or chemical modification of starch. Study of relaxation processes in native corn starch plasticized with different mixtures of glycerol and urea were examined by Solid-State ^{1}H, ^{13}C NMR and WAXS.**

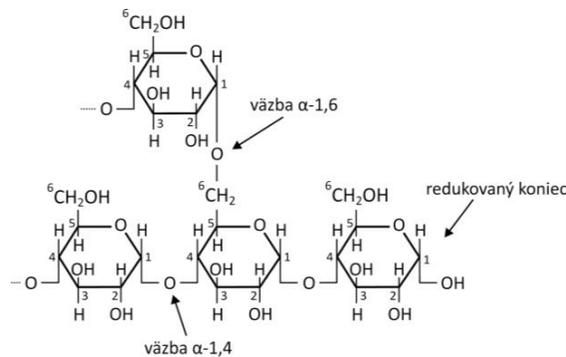*Keywords—starch, glycerol, urea, structural relaxation, NMR, WAXS*

I. ÚVOD

Dôležitosť ochrany životného prostredia sa stále viac dostáva do popredia. Vážnou hrozbou je rýchlo rastúci objem plastového odpadu v prírode, ktorého schopnosť degradovať je obmedzená a navyše v mnohých prípadoch uvoľňujú do svojho okolia toxické látky [1]. Riešením by mohla byť náhrada takýchto materiálov za biodegradovateľné polyméry, ktorých

uplatnenie nájdeme v aplikáciách s krátkou dobou životnosti, resp. v prípadoch, kde je recyklačný proces technologicky náročný alebo ekonomicky nevýhodný [2]. Ich nedostatkom je však oproti súčasne používaným polymérom vo všeobecnosti vysoká cena a nepostačujúce mechanické vlastnosti. Kvôli tomu sa časť výskumu venuje biodegradovateľným polymérom na báze škrobu, ktorých základný materiál je globálne dostupný, obnoviteľný a cenovo výhodný [3].

Škrob je prírodný polysacharid produkovaný v telách vyšších rastlín ako zásoba energie vo forme semikryštalických granúl. Základnou stavebnou jednotkou škrobu je α-D-gluko-pyranóza, najstabilnejšia forma glukózy, ktorá je primárnym produktom fotosyntézy rastlín [3, 4]. Polykondenzačnou reakciou molekúl α-D-glukopyranózy dochádza ku vzájom-nému vzniku glykozidických väzieb [5], ktoré v prípade väzieb α-1,4 tvoria reťazce a väzbami α-1,6 vznikajú body vetvenia makromolekúl škrobu (Obr. 1) [3].

Rozlišujeme dva druhy makromolekúl škrobu. Lineárnu alebo iba málo vetvenú amylózu a silne rozvetvený amylo-pektín obsahujúci veľké množstvo krátkych reťazcov. Uvedené zložky môžu byť v amorfnej alebo usporiadanej štruktúre, pričom usporiadanie makromolekúl je dôsledkom konformácie cis väzby α-1,4, ktorá umožňuje tvorbu ľavotočivých závitni-cových štruktúr. Pre amylózu sú typické jednozávitnice obsahujúce vo svojom vnútri komplexujúci prvok (najčastejšie tuky). Amylóza, ale predovšetkým amylopektín sú typické tvorbou dvojzávitníc. Jedno- aj dvojzávitnice sa môžu ďalej usporadúvať do väčších celkov prostredníctvom vodíkových
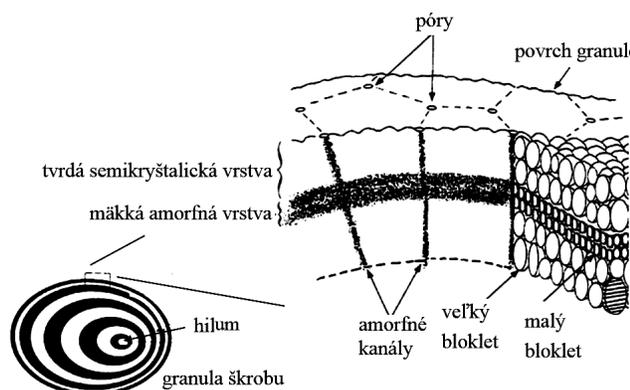

Obr. 1 Štruktúra makromolekuly škrobu.

väzieb, tzv. polymorfizmov. Dvojzávitnice sa usporadúvajú do monoklinickej (typ A), hexagonálnej (typ B) mriežky, prípadne ich kombinácie (typ C). Jednozávitnice môžu vytvárať usporiadanie typu V [6, 7]. Striedanie takto vzniknutých usporiadaných a amorfných štruktúr v rámci jednej alebo niekoľkých makromolekúl amylopektínu tvorí bloklet, vyšší štruktúrny celok charakteristický zvýšenou odolnosťou voči chemickej degradácii. Ich veľkosť odráža dokonalosť usporiadanosti dvojzávitníc, pričom organizácia je v prípade objemnejších blokletov dokonalejšia ako v prípade menších [8]. Radiálne usporiadanie veľkých a malých blokletov tvorí „amorfné" a „semikryštalické" vrstvy granúl škrobu (Obr. 2) [9].

Natívny škrob (NS) nie je termoplastickým materiálom dôsledkom silných inter- a intra-molekulových interakcií, kvôli ktorým sa teplota topenia nachádza nad teplotou termického rozkladu. Termo-mechanickou úpravou NS v prítomnosti nízkomolekulového plastifikátora, ktorý interaguje so škrobom prostredníctvom hydroxylových skupín, môžeme nahradiť interakcie škrob-škrob interakciami škrob-platifikátor, pričom dochádza k rozpadu pôvodnej štruktúry granuly škrobu [10]. Uvedený proces nazývame gélovatenie a dochádza v špecifickom rozsahu teplôt a pôsobením šmykového napätia. Medzi vhodné plastifikátory patrí glycerín, močovina, sorbitol, kyselina citrónová a mnoho ďalších. Homogenizáciou škrobového gélu získavame tzv. termoplastický škrob (TPS), ktorého vlastnosti sú podobné termoplastom [11].

Nemodifikovaný TPS neposkytuje postačujúce mechanické vlastnosti, je citlivý na vzdušnú vlhkosť a podstupuje relatívne rýchle relaxačné procesy [2]. Zlepšiť vlastnosti TPS je možné chemickým nahradením hydrofilných OH skupín hydrofóbnymi, čím sa dá výrazne znížiť citlivosť materiálu na vlhkosť. Ďalšou možnosťou je zosieťovanie makromolekúl škrobu, vďaka ktorej je možné okrem poklesu citlivosti na vlhkosť navyše výrazne zvýšiť mechanické vlastnosti TPS. Najväčším obmedzením takýchto prístupov je však znížená schopnosť biodegradability výsledného materiálu [11]. Alternatívny prístup zahŕňa vytváranie polymérnych zmesí s inými biodegradovateľnými alebo nebiodegradovateľnými polymérmi, či výroba kompozitných materiálov a nano-kompozitov [2, 12 – 14].

Obr. 2 Štruktúra granúl škrobu na úrovni blokletov usporiadaných do koncentricky sa striedajúcich mäkkých amorfných a tvrdých semikryštalických vrstiev od centrálnej dutiny zvanej hilum [6].

Štruktúra a relaxačné procesy na prírodnom kukuričnom škrobe plastifikovanom glycerínom a močovinou boli pozorované metódami NMR a WAXS v časovom intervale 52 týždňov od prípravy vzoriek.

II. EXPERIMENTÁLNA ČASŤ

A. Vzorky

Vzorky TPS (Tab.1) boli vyrobené z kukuričného škrobu Meritena® 100 (Brenntag Slovakia) v spolupráci s Ústavom polymérov SAV v Bratislave.

Príprava pozostávala namiešaním suspenzie kukuričného škrobu, plastifikátorov a vody v hmotnostnom pomere 1:0,7:2,3, ktorá podstúpila proces gélovatenia za neustáleho miešania pri teplote 70 °C po dobu 5 – 10 minút. Voda, ktorá pôvodne slúžila na rozpustenie kryštálov močoviny a umožnila úplný rozpad štruktúry granúl, je v TPS nežiadúca, preto sa vzorky sušili po dobu 5 hodín pri teplote 100 °C. Homogenizácia vzniknutého gélu sa uskutočnila prostredníctvom laboratórnej miešačky Plastograph Barbender PLE 331 pri teplote 130 °C po dobu 10 minút. Laboratórnym lisom LabEcon300 boli následne vyformované doštičky TPS pri teplote 130 °C a tlaku 100 kPa [15]. Vzorky boli uskladnené vo vreckách z PE na tmavom mieste pri teplote 22 °C.

B. NMR

NMR merania boli uskutočnené na spektrometri Varian Solid-State NMR (VNMRS 400, Palo Alto, CA, USA) vybaveným supravodivým magnetom so širokou dutinou pracujúcim pri rezonančnej frekvencii jadier vodíka 400 MHz. Vzorky boli uložené do ZrO_2 rotorov s priemerom 4 mm s možnosťou rotácie pod magickým uhlom (MAS) pri frekvencii do 16 kHz. Teplota vzorky sa stanovila na základe výsledkov práce [16] a stupnica chemického posunu bola kalibrovaná pre tetrametylsilán (TMS) s využitím adamantanu ako externej referencie. Vyhodnocovanie sa robilo za pomoci softvérov Varian VnmrJ 3.2, Mestrelab Research Mnova 9.0 a Origin Pro.

Trvanie π/2 impulzu na jadrách vodíka bolo 2,9 µs. Čakacia doba ^{1}H NMR meraní bola 10 s, doba akvizície 20 – 80 ms. Merania boli uskutočnené po 1, 3, 7, 20, 50, 190 a 365 dňoch od výroby vzoriek pri teplote 22 °C pre širokočiarové (BL) a 30 °C pre MAS ^{1}H NMR. Po 52 týždňoch bola urobená teplotná analýza v rozsahu 30 – 110 °C pre U-TPS, UG(1:1)-TPS a G-TPS využitím MAS ^{1}H NMR techniky.

TAB. 1 VZORKY TPS S HMOTNOSTNÝM PODIELOM MOČOVINY A GLYCERÍNU VZHĽADOM NA JEDEN HMOTNOSTNÝ DIEL ŠKROBU.

Vzorka:	m_u/m_s	m_g/m_s
U-TPS	0,7	0
UG(3:1)-TPS	0,525	0,175
UG(1:1)-TPS	0,35	0,35
UG(1:3)-TPS	0,175	0,525
G-TPS	0	0,7

Merania MAS ^{13}C NMR sa uskutočnili za aplikácie výkonného heteronukleárneho dipólového dekaplingu (HDD) a krížovej polarizácie (CP). Kontaktný impulz bol nastavený na 2 ms, dobe akvizície 20 – 40 ms, čakacia doba 8 s a výkon HDD bol 110 kHz.
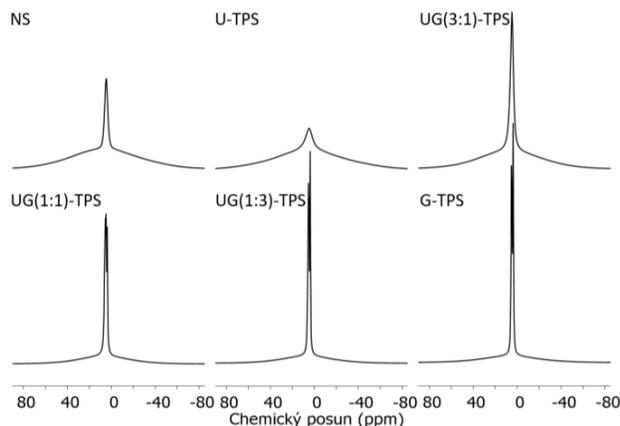
C. WAXS

Merania WAXS boli uskutočnené pomocou Rigaku MiniFlex 600 X-Ray práškového difraktometra. Generátor röntgenového žiarenia s vlnovou dĺžkou 0,154 nm (Cu anóda) pracoval pri napätí 40 kV a využíval Ni filter na elimináciu K_β žiarenia. Rozsah difraktogramov bol v rozmedzí 2θ od 5 – 40° pri rýchlosti merania 5°/min a s krokom 0,01°.
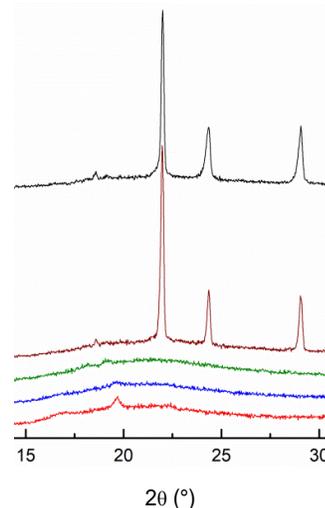
III. VÝSLEDKY A DISKUSIA

A. BL ^{1}H NMR

Inter- a intra-molekulové interakcie určujú tvar a šírku spektier NMR polykryštalických polymérov, pričom v istých prípadoch je možné pozorovať oblasti rôznej pohyblivosti [17, 18]. Podobné charakteristiky majú na prvý pohľad aj BL ^{1}H NMR spektrá NS a vzoriek TPS (Obr. 3). Je zrejmé, že pozorovaný systém disponuje oblasťami s výrazne odlišnou pohyblivosťou. Úzke zložky spektra predstavujú pohyblivé molekuly vody, močoviny a glycerínu. Široký, nepravidelný signál popisuje makromolekuly škrobu a molekuly vody či močoviny, ktorej pohyblivosť bola zásadným spôsobom obmedzená [19].

Procesom plastifikácie sa predpokladá nárast pohyblivosti makromolekúl škrobu kvôli rozrušeniu pôvodne silných vodíkových interakcií a vzniku amorfného materiálu. Takáto zmena je pozorovaná v spektrách NMR zúžením širokej zložky. Pri pohľade na výsledky sa však ukazuje, že v prípade vzorky U-TPS nedošlo k zvýšeniu pohyblivosti reťazcov škrobu, pretože šírka zložky popisujúca makromolekuly škrobu zostala prakticky nezmenená (30,2 kHz pre U-TPS a 30,5 kHz pre NS). Ďalšou charakteristickou črtou vzorky U-TPS je značný pokles intenzity signálu pohyblivých zložiek oproti

Obr. 3 BL ^{1}H NMR spektrá pre NS a vzorky TPS namerané 52 týždňov po príprave. Spektrá sú normované podľa obsahu pričom amplitúda spektier v hornom riadku je z vizuálnych príčin dvojnásobkom amplitúd spektier v dolnom riadku.

Obr. 4 WAXS difraktogramy pre vzorky TPS namerané 52 týždňov po príprave vzoriek.

spektrám ostatných vzoriek, čo môže byť dôsledkom predpokladanej fázovej separácie molekúl močoviny, pričom pohyblivosť jednej fázy bude výrazným spôsobom potlačená. Uvedený jav bol taktiež pozorovaný pri TPS filmoch s obsahom močoviny nad 10 %. Vznikali pritom oblasti bohaté na močovinu s vyššou usporiadanosťou [20]. Tento predpoklad bol potvrdený meraniami WAXS, ktoré ukázali silné reflexie pri uhloch 2Θ ~ 22°, 24° a 29° v prípade vzoriek U-TPS a UG(3:1)-TPS popisujúce kryštalickú formu močoviny (Obr. 4).

Taktiež šírka signálu popisujúcich pohyblivé zložky je v porovnaní s ostatnými spektrami výrazne väčšia. Tento fakt indikuje prítomnosť dodatočných interakcií vo vzorke U-TPS spomaľujúcich pohyblivosť molekúl močoviny v amorfnom stave.

Prítomnosť už malého množstva glycerínu má výrazný vplyv na pohyblivosť skúmaných systémov bez ohľadu na množstvo močoviny nachádzajúcej sa vo vzorkách TPS. Pozorujeme pritom zúženie všetkých signálov v spektrách. Zúženie širokých zložiek je o 15,6 % pre UG(3:1)-TPS a približne 31 % pre vzorky s vyšším obsahom glycerínu. Fakt, že ďalší nárast množstva glycerínu ovplyvňuje šírky iba nepatrne, svedčí o možnej fázovej separácii glycerínu vo vzorkách UG(1:3)-TPS a G-TPS.

Pohyblivosť reťazcov škrobu závisí od výberu a množstva plastifikátora avšak BL ^{1}H NMR spektrá nepreukázali žiaden viditeľný vplyv relaxačných procesov na ich výslednú pohyblivosť. Všeobecný trend pohyblivých oblastí je postupný pokles ich šírky a nárast amplitúdy so zvyšujúcim sa podielom glycerínu.

B. MAS ^{1}H NMR

Metóda MAS potlačuje vplyv nežiaducich magnetických interakcií jadrového systému, čím dochádza k zúženiu rezonančných čiar v spektrách NMR. Veľmi silné nežiaduce interakcie reťazcov škrobu a nízka rýchlosť rotácie vzorky sa v spektrách prejavili vznikom série rotačných signálov, ktoré
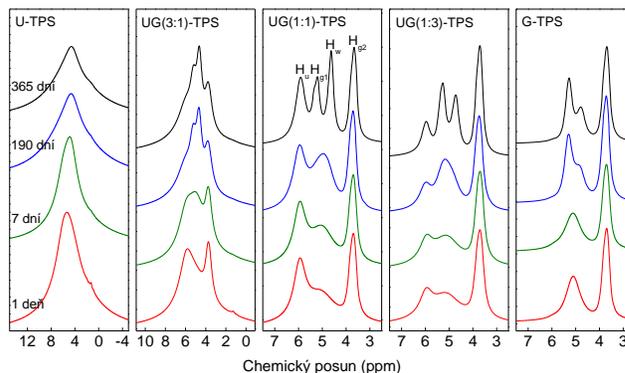
nie sú vhodné na interpretáciu, preto sa pri analýze budeme venovať iba úzkym zložkám spektier. MAS ^{1}H NMR spektrá vzoriek TPS namerané pri vybraných časových intervaloch štruktúrnej relaxácie sú znázornené na Obr. 5.

Podobne ako v prípade širokočiarových spektier sú spektrá U-TPS aj v tomto prípade najmenej rozlíšené. Tu je však možné pozorovať postupné rozšírenie rezonančnej čiary ako aj postupný pokles amplitúdy. Ukázalo sa, že uvedený proces je charakteristický exponenciálnym poklesom intenzity s časovou konštantou 15 dní, ktorý sa ustálil po približne 100 dňoch. Počas tejto doby došlo k postupnej kryštalizácii značného množstva močoviny vo vzorke. Rozširovanie signálu v priebehu starnutia môže byť vysvetlené postupným vznikom vodíkových interakcií medzi molekulami vody a pohyblivej močoviny. Vzniknutý komplex je mohutnejší a za rovnakých podmienok je charakteristický nižšou schopnosťou reorientácie, ktorá vysvetľuje viditeľné rozšírenie signálu v spektre.

Uvedenú teóriu podporujú aj spektrá vzorky UG(3:1)-TPS, v ktorej je možné pozorovať signál močoviny H_u približne na pozícii 6,0 ppm. V priebehu štruktúrnej relaxácie je možné pozorovať pokles amplitúdy a nárast šírky signálu. Tento fakt bol potvrdený aj dekonvolúciou spektra. Z výsledkov však vyplynulo, že intenzita signálu močoviny sa v priebehu starnutia zmenila iba nepatrne. Keďže WAXS merania ukázali prítomnosť kryštalickej formy močoviny, ale zmena intenzity signálu NMR nebola pozorovaná, znamená to, že podstatná časť kryštalizácie musela prebehnúť už krátko po príprave vzorky. Vzorky s nižším obsahom močoviny nevykazujú pozorovateľné zmeny šírky ani intenzity signálu H_u. V týchto prípadoch teda nedochádza ani k fázovej separácii, ani k vzniku interakcií molekúl močoviny a vody, resp. prebiehajú iba minimálne.

Na rozdiel od signálu močoviny, signál voľnej vody H_w ($\approx$4,7 ppm), signály OH skupín glycerínu H_{g1} ($\approx$5,35 ppm) a vodíkov priamo viazaných na uhlíky glycerínu H_{g2} ($\approx$3,72 ppm) sa v priebehu starnutia zužujú, pričom po 190 dňoch v prípade UG(3:1)-TPS a G-TPS a po 52 týždňoch pre UG(1:1)-TPS a UG(1:3)-TPS môžeme všetky signály v spektrách jasne rozlíšiť. Spektrum namerané po 1 a 7 dňoch sú pre vzorku UG(3:1)-TPS charakteristické superpozíciou signálov H_u, H_{g1} a H_w. Zrejmý je aj nárast intenzity spôsobený možnou adsorpciou vody vzorkou. Pravdepodobnejším vysvetlením môže byť však nárast množstva pohyblivej vody (tzv. voľnej vody), ktorá bola vytlačená zo štruktúry škrobu pri jeho reusporiadaní, pričom zanikli vodíkové interakcie s makromolekulami škrobu.

Podobné správanie sa signálov voľnej vody a glycerínu v priebehu starnutia pozorujeme aj v prípadoch UG(1:1)-TPS, UG(1:3)-TPS a G-TPS. Zúženie rezonančných čiar nám v tomto prípade umožňuje rozlíšiť signál močoviny od prvého dňa. Taktiež môžeme vidieť nárast intenzity signálu na pozícii 5,1 ppm v prípade vzoriek UG(1:1)-TPS, UG(1:3)-TPS, ktoré potvrdzujú nárast množstva voľnej vody vo vzorkách. Vzorka

Obr. 5 MAS ^{1}H NMR spektrá vzoriek TPS meraných vo vybraných intervaloch starnutia.

G-TPS nezaznamenala výrazný nárast množstva voľnej vody a teda predpokladáme, že adsorpcia alebo relaxačné procesy v tomto prípade museli prebehnúť veľmi rýchlo.
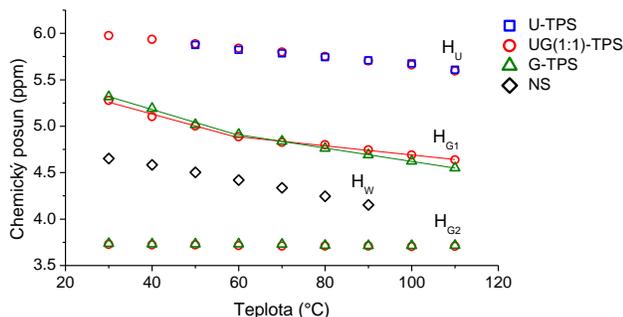
C. Určenie množstva voľnej vody

Ako sa ukázalo v predchádzajúcej časti, množstvo voľnej vody môže v budúcnosti poskytnúť veľa užitočných informácií o štruktúre a molekulovom pohybe vzoriek. Na základe toho bol vytvorený teoretický koncept stanovenia množstva voľnej vody vo vzorkách TPS pomocou ^{1}H NMR spektroskopie.

Teória vychádza z predpokladu kvantitatívneho charakteru jednoimpuzových techník NMR a známej chemickej štruktúry všetkých komponentov vzorky. Hmotnostný podiel zložky neznámeho množstva môžeme vypočítať na základe podielu jej intenzity s intenzitou zložky, ktorej množstvo je vo vzorke pevne dané. Vďaka tomu môžeme teoreticky určiť hmotnostný pomer voľnej vody vzhľadom na hociktorý dostatočne rozlíšený signál plastifikátora v spektre. Na ilustráciu uvádzame v Tab. 2 výsledky výpočtu množstva voľnej vody pre BL ^{1}H NMR spektrum vzorky UG(1:1)-TPS namerané po 52 týždňoch od jej výroby.

Výsledky ukazujú, že v prípade dobre rozlíšeného spektra môžeme určiť množstvo voľnej vody dostatočne presne pre akýkoľvek signál plastifikátora. Častokrát v spektre rozlišujeme iba signál glycerínu H_{g2}, na základe ktorého vieme stanoviť množstvo voľnej vody aj pre spektrá nižšieho rozlíšenia. Platí to však iba v prípadoch, že množstvo všetkých plastifikátorov prispievajúcich úzkym zložkám je známy. To je jedným z dôvodov, prečo stanovenie množstva voľnej vody nebolo možné v prípade vzorky UG(3:1)-TPS. Vzorka U-TPS navyše neposkytuje postačujúce rozlíšenie spektier.

TAB. 2 Intenzity signálov a vypočítané hmotnostné podiely voľnej vody vo vzorke UG(1:1)-TPS nameranej 52 týždňov po jej príprave.

	H_u	H_{g1}	H_w	H_{g2}
Intenzita (a.u.)	0,335	0,165	0,226	0,274
Podiel voľnej vody	0,142	0,141	-	0,141

Obr. 6 Teplotná závislosť chemického posunu signálov voľnej vody (H_w), glycerínu (H_{g1} a H_{g2}) a močoviny (H_u) v spektrách pre NS, U-TPS, UG(1:1)-TPS a G-TPS v teplotnom rozsahu 30 – 110 °C namerané 52 týždňov po príprave vzoriek. Plnými čiarami je zvýraznená nelinearita signálu H_{g1}.
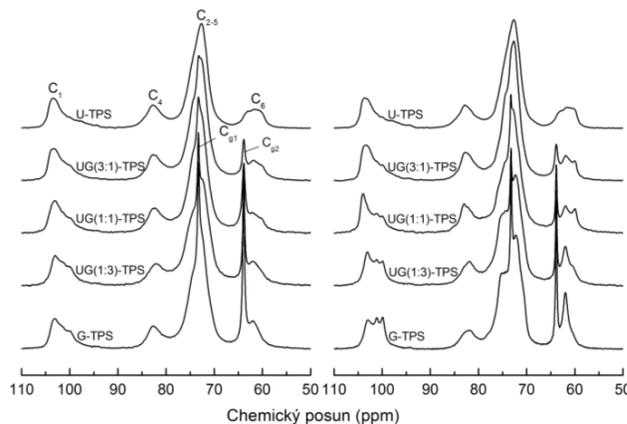
D. Teplotná analýza MAS ^{1}H NMR spektier

^{1}H NMR spektrá merané po dobe dlhšej ako 52 týždňov nevykazovali výrazné zmeny, čím môžeme predpokladať stabilizáciu štruktúry vzoriek. Kvôli tomu sme sa rozhodli pre teplotnú analýzu vzoriek NS, U-TPS, UG(1:1)-TPS a G-TPS v rozsahu 30 – 110 °C. Výrazným efektom v spektrách NMR bol posun polohy signálov polárnych skupín so zvyšujúcou sa teplotou. Dôvodom je nárast pohyblivosti spojený s nárastom strednej dĺžky vodíkových väzieb. Dôsledkom toho sa oslabí magnetické tienenie v okolí takéhoto jadra, ktoré je sprevádzané následným posunom polohy rezonančnej čiary v spektre [21, 22].

Teplotná závislosť chemického posunu je znázornená na Obr. 6. Poloha signálu H_{g2} sa mení iba nepatrne v pozorovanom teplotnom rozsahu, pretože nepolárne väzby medzi atómami uhlíka a vodíka ako jediné neumožňujú vznik vodíkových väzieb.

Signál voľnej vody bolo možné dostatočne rozlíšiť iba v prípade NS, ktorý s nárastom teploty klesal lineárne rýchlosťou $8{,}4.10^{-3}$ ppm/K. Miernejší lineárny pokles pozorujeme aj v prípade signálu vodíkov močoviny H_u s rýchlosťou $4{,}5.10^{-3}$ ppm/K. V tomto prípade je zrejmé, že teplotná závislosť chemického posunu močoviny vzoriek UG(1:1)-TPS a U-TPS nevykazuje výrazné rozdiely, vďaka čomu môžeme vylúčiť interakcie medzi glycerínom a amidovými skupinami močoviny. Rozdielne správanie však môžeme pozorovať v prípade signálu H_{g1}, ktorý je vo vyšetrovanom teplotnom rozsahu charakterizovaný dvojicou lineárnych oblastí so zlomom pri teplote okolo 60 °C. Príčina nelinearity nie je úplne známa, avšak odlišné chovanie pre vzorky UG(1:1)-TPS a G-TPS môže byť vysvetlené vznikom interakcií medzi kyslíkom karbonylovej skupiny v molekule močoviny a vodíkom v hydroxylovej skupine glycerínu [23].

E. CP MAS ^{13}C NMR

^{13}C NMR spektrá vďaka absencii homonukleárnych dipólových interakcií nám umožňujú rozlíšiť signály jadier uhlíka makromolekúl škrobu rozdelených do štyroch celkov. Rezonancie uhlíka C_1 sa nachádzajú v rozmedzí 90 – 110 ppm. Signál v oblasti 78 – 90 ppm popisuje jadrá uhlíka C_4 nachádzajúceho sa v amorfných oblastiach [24]. Oblasť 65 –

Obr. 7 CP MAS ^{13}C NMR spektrá vzoriek TPS namerané 3 dni (vľavo) a 52 týždňov (vpravo) po príprave vzoriek.

78 ppm predstavuje rezonancie uhlíkov C_2, C_3, C_5 ako aj uhlík C_4 nachádzajúci sa v kryštalických doménach. Bočná skupina s uhlíkom C_6 sa nachádza v oblasti 55 – 65 ppm [25]. Úzke signály na pozíciách 63,8 a 73,2 ppm patria uhlíkom glycerínu [26].

Obr. 7 znázorňuje CP MAS ^{13}C NMR spektrá vzoriek TPS namerané 3 dni a 52 týždňov po ich príprave. Vzorka U-TPS podstúpila najmenšie štruktúrne zmeny v priebehu starnutia. V oblasti C_1 sa nachádza dominantný signál na pozícii 103 ppm so širokým plecom po pravej strane, indikujúci amorfnú štruktúru molekúl amylopektínu. Na základe výsledkov však nie je možné vylúčiť prítomnosť usporiadania typu V charakteristického pre amylózu. S rastúcim podielom glycerínu vo vzorkách sa ukazuje v spektrách dublet na pozícii 100 a 101 ppm potvrdzujúci prítomnosť usporiadania typu B [27], pričom podiel usporiadanosti rastie spolu s množstvom glycerínu. V procese starnutia došlo k výraznému zúženiu signálov a nárastu rozlíšenia spektier.

Taktiež je pozorované štiepenie signálu v oblasti uhlíka C_6 na dvojicu signálov na pozíciách 60,0 popisujúci usporiadanie typu V a 61,9 ppm popisujúci usporiadanie typu B. Z výsledkov sa ukázalo, že najvýhodnejšie podmienky pre vznik usporiadania typu V boli v prípade vzorky UG(1:1)-TPS, kde v spektrách bolo možné dodatočne pozorovať posun signálu z pozície 103 na 104 ppm a vznik signálu na pozícii 82,9 ppm potvrdzujúce toto usporiadanie.

IV. Záver

Vzorky TPS plastifikované glycerínom a močovinou boli pozorované metódami NMR. Výskum sa zameriaval na štruktúrne zmeny, pohyblivosť systémov a vznik/zánik interakcií medzi zložkami pozorovaných systémov v priebehu 52 týždňov ich štruktúrnej relaxácie a pri zvýšenej teplote.

V prípade vzorky U-TPS bola pozorovaná transformácia štruktúry NS v priebehu plastifikácie, ktorá však nemala pozorovateľný vplyv na pohyblivosť reťazcov škrobu. Tie v priebehu starnutia podstúpili najmenšie štruktúrne zmeny a nebol pozorovaný nárast usporiadanosti. Vysoký podiel močoviny mal však za následok jej fázovú separáciu do oblastí

s výrazne odlišnou pohyblivosťou a stupňom usporiadanosti. Proces reorganizácie molekúl močoviny pokračoval v priebehu relaxácie, pričom sa ustálil približne po 100 dňoch. Pozorované postupné znižovanie pohyblivosti molekúl močoviny v amorfných oblastiach môže byť dôsledkom vzniku vodíkových väzieb s molekulami voľnej vody.

Podobné správanie molekúl močoviny, ale v menšej miere, bolo pozorované aj v prípade UG(3:1)-TPS. Výsledky však ukázali zásadný vplyv glycerínu na nárast pohyblivosti a tým aj urýchlenie relaxačných procesov. V priebehu starnutia bol pozorovaný nárast pohyblivosti molekúl voľnej vody a glycerínu ako dôsledok vzniku usporiadanosti typu B. Tým došlo k uvoľneniu interakcií medzi molekulami škrobu a vodou či glycerínom. Otvorenou otázkou je možnosť adsorpcie vody vzorkou z okolitého prostredia, keďže podmienky skladovania výrazne obmedzujú túto možnosť.

Vzorky s nižším obsahom močoviny neobsahujú usporiadanú formu močoviny, taktiež nebol pozorovaný vznik komplexov močoviny s voľnou vodou. Teplotná analýza však preukázala možnosť vzniku interakcií medzi molekulami glycerínu a močoviny, ktorá bola pozorovaná na vzorke UG(1:1)-TPS. Ich prítomnosť sa nedá vylúčiť ani v prípade vzoriek UG(3:1)-TPS a UG(1:3)-TPS.

Prítomnosť glycerínu sa v najväčšej miere podpísala na zvýšení pohyblivosti makromolekúl škrobu a rýchlosti relaxačných procesov. Merania ukázali najpriaznivejšie podmienky pre vznik usporiadania typu V pre UG(1:1)-TPS. S rastúcim podielom glycerínu bol pozorovaný predovšetkým nárast usporiadania typu B.

POĎAKOVANIE

POUŽITÁ LITERATÚRA

[1] R. L. Shorgen, Starch Polymers as Advanced Material for Industrial and Consumer Products. in: Starch-Based Polymeric Materials and Nanocomposites Boca Raton: CRC Press, 2012, pp. 396. ISBN 978113819623.

[2] M. Kaseem, K. Hamad, F. Deri, Thermoplastic starch Blends: A review of recent works, in Polymer Science, 2012, vol. 54 (2), pp. 165-176.

[3] A. Carvalho, Monomers, Polymers and Composites from Renewable Resources, First Edition, Kidlington: Elsevier Ltd., 2008, pp. 321-343.

[4] D.-J. Thomas, W.-A. Atwell, Starches, The American Association of Cereal Chemists, Inc., St. Paul, Minnesota: Eagan Press, 1997.

[5] S. Pérez, E. Bertoft, The molecular structures of starch components and their contributions ot the architecture of starch granules: A comprehensive review. in Starch/Stärke, vol. 62, 2010, pp. 389-420.

[6] H. Tang, Y. Wang, High-Resolution Solid-State NMR Spectroscopy of Starch Polysacharides. in Modern Magnetic Resonance, Food Science, 2008, pp. 1783-1791.

[7] H. Thérein-Aubin, X.-X. Zhu, NMR Spectroscopy and imaging studies of pharmaceutical tablets made of starch. in Carbohydrate Polymers, vol. 75, 2009, pp. 369-379.

[8] N. P. Badenhuizen, Die Struktur des Stärkekorns. in: Protoplasma, Volume 28, 1937, pp. 293-326.

[9] W.-S. Ratnayake, D.-S. Jackson, Gelitinization and Solubility of Corn Starch during Heating in Exces of water: New Insight. in Journal of Agricultural and Food Chemistry, American Chemical society, vol. 54, 2006, pp. 3712-3716.

[10] Y. Zhang, C. Rempel, D. McLaren, Innovation in Food Packaging, Chapter 16, Thermoplastic Starch, 2014, pp. 391-408.

[11] A.-M. Nafchi, M. Moradpour,M. Saeidi,A.-K. Alias, Thermoplastic starches: Properties, challengies, and prospects, in Starch/Stärke, 2013, vol. 65, pp. 61-72.

[12] F. Xie, E. Pollet, P.-J. Halley, L. Avérous, Starch-based nano-biocomposites, Progress in Polymer Science, Elsevier Ltd., 2013, vol. 38, pp. 1590-1628.

[13] L. Avérous, P.-J. Halley, Biocomposites based on plasticized starch, Review, in Biofuels, Bioprod. Bioref., 2009, vol. 3, pp. 329-343.

[14] A. Maurizio et al. Biodegradable starch/clay nanocomposite films for food packaging applications, in Food Chemistry, 2005, vol.93, pp. 467-474.

[15] F. Ivanič, D. Jochec Mošková, I. Janigová, I. Chodák, Physical properties of starch plasticized by a mixture of plasticizers. in: European Polymer Journal 93, 2017, pp. 843 – 849.

[16] V. Hronský, Temerature calibration of solid-state MAS NMR probe, in Physics of Materials, 2012, pp.217 – 220.

[17] U. Scheler, Solid Polymers, in: M. J. Duer (Ed.), Solid-State NMR Spectroscopy: Principles and aplications, Wiley, London, 2002, pp. 483 – 511.

[18] D. Olčák, V. Hronský, O. Fričová, M. Kovaľaková, P. Duranka, I. Chodák, Solid and metl-state ^{1}H NMR studies of relaxation processes in isotactic polypropylene. in: Journal of Polymer Research 20, 2013.

[19] J. Spěváček, J.Brus, T. Divers, Y. Grohens, Solid-state NMR study of biodegradable starch/polycaprolactone blends. in: European Polymer Journal 43, 2007, pp. 1866–1875.

[20] J. Wang, F, Cheng, P. Zhu, Structure and properties of urea-plasticized starch films with different urea contents. in: Carbohydrate Polymers 101, 2014, pp. 1109 – 1115.

[21] N. Muller, R. C. Reiter, Temperature Dependence od Chemical Shifts of Protons in Hydrogen Bonds. in: The Journal of Chemical Physics 42, 1965, pp. 3265-3269.

[22] V. Balevicius, K. Aidas, Temperature Dependence of ^{1}H and ^{17}O NMR Shifts of Water: Entropy Effect. in: Applied Magnetic Resonance 32, 2007, pp. 363-376.

[23] A. Šoltýs, V. Hronský, N. Šmídová, D. Olčák, F. Ivanič, I. Chodák, Solid-state 1H and 13C NMR of corn starch plasticized with glycerol and urea, European Polymer Journal 117, 2019, pp. 19-27.

[24] M. Paris, H. Bizot, J. Emery, J.-Y. Buzaré, A. Buléon, Crystallinity and structuring role of water in native and recrystallized starches by CP 13C CP-MAS NMR spectroscopy 1: Spectral decomposition, in Carbohydrate Polymers, Elsevier Ltd., 1999, vol. 39, pp. 327-339.

[25] D. Fan a kol., Determining the effects of microwave heating on the ordered structures of rice be NMR. in: Carbohydrate Polymers 92, 2007, pp. 1525 – 1529.

[26] H. Liu, Q. Guo, R. Adhikari, Preparation and characterization of glycerol plasticized (high-amylose) starch-chitosan films, in Journal of Food Engineering, 2013, vol. 116, pp. 588-597.

[27] C. Mutungi, L. Passauer, C. Onyango, D. Jaros, H. Rohm, Debranched cassava starch crystallinity determination by Raman spectroscopy: Correlation of features in Raman spectra with X-ray diffraction and 13C CP/MAS NMR spectroscopy, in Carbohydrate Polymers, 2012, vol. 87, pp. 598-606.

Sentiment Analysis in Human-robot Interaction

Martina Szabóová

Department of Cybernetics and Artificial Intelligence
Faculty of Electrical Engineering and Informatics
Technical university of Košice
Email: martina.szaboova@tuke.sk

Kristna Machová

Department of Cybernetics and Artificial Intelligence
Faculty of Electrical Engineering and Informatics
Technical university of Košice
Email: kristina.machova@tuke.sk

Abstract—The main goal of this work is to interconnect two large research areas such as sentiment analysis and human-machine interaction. Intersection of these two areas is affective computing, more precisely the analysis of emotions. Emotion analysis, as a subfield of sentiment analysis, explores text data and evaluates what emotion is presented in it. In the human-robot interaction it aims to evaluate the emotional state of the human being and on this basis to decide how the robot should adapt its behavior to the human being. Combined method of dictionary approach with machine learning algorithms was applied. We created a dictionary of Slovak emotional words within Aesop's Fables, where we proved the need to personalize the evaluation of emotions. Since the amount of text was above the possibility of manual annotation, we utilized selected annotated sentences, to extend annotating on unannotated data. For this purpose we used *k-means* clustering method. Because we were solving a multi-label problem, we decided to transform the problem to solve several simple binary classifiers. Based on the overview of the problem, we chose the Naive Bayes (NB) algorithm and Support Vector Machine (SVM). In second part of the work we have designed and conducted several experiments to test our hypotheses of what one expects in good interaction with a robot.

Keywords—emotion detection, sentiment analysis, human-robot interaction, machine learning, dictionary approach

Abstrakt—Cieľom predkladanej práce je prepojenie dvoch výskumných oblastí: analýza sentimentu a interakcia človek-robot. Prienikom týchto dvoch oblastí je analýza emócií. Analýza emócií v oblasti analýzy sentimentu skúma text a vyhodnocuje, aká emócia sa v danom texte nachádza. Analýza emócií v interakcii človek-robot má za úlohu vyhodnotiť emočný stav človeka a na základe toho rozhodnúť, ako má robot svoje správanie prispôsobiť človeku. K vyhodnoteniu tohto stavu je potrebných viac vstupov: výraz tváre, chrakteristika hlasu, pohybový prejav a hovorený text. Tieto vstupy sa spoločne nazývajú multimodality a v predostrenej práci pojednávame práve o jednej z týchto modalít - hovorený text. Na detekciu emócií v texte existuje niekoľko prístupov. Rrozhodli sme sa aplikovať kombinovanú metódu slovníkového prístupu s algoritmami strojového učenia. V práci sme vytvorili slovník slovenských emočných slov v rámci Ezopových bájok, kde sme dokázali potrebu personalizácie ohodnotenia emócií. Ďalej sme pokračovali analýzou anglických Ezopových bájok. Jednotlivé vety bolo možné zaradiť do viacerých tried. Pre množstvo textu sme z pár vybraných anotovaných viet použitím metódy zhlukovania k-means, označkovali aj neoznačené vety. Problém klasifikácie do viacerých tried sme reansformovali na niekoľko (podľa počtu detegovaných emócií) jednoduchých binárnych klasifikátorov. Na základe prehľadu problematiky sme zvolili algoritmy Naivného Bayesa a podporných vektorov. Výsledky sme aplikovali v druhej časti práce, pri interakcii človek-robot. Navrhli a vykonali sme niekoľko experimentov na overenie našich hypotéz toho, čo človek pri interakcii očakáva aby boli sčasti naplnené jeho potreby pri komunikácii s robotom. Z experimentov je zrejmé, že prejavovanie emócii na základe slov, ktoré hovoríme, je dôležité.

Kľúčové slová—detekcia emócií, analýza sentimentu, interakcia človek-robot, strojové učenie, slovníkový prístup

I. Introduction

Sentiment analysis is not a new field, however, only recently has it enjoyed a huge burst of research activity. The reasons may include:

- the rise of machine learning methods in natural language processing and information retrieval;
- the availability of datasets for machine learning algorithms for training, due to a blossoming of the World Wide Web and, specifically, the development of review-aggregation web-sites and social networks;
- the recognition of fascinating intellectual challenges and the commercial and intelligence applications the area offers;

Most projects use sentiment analysis for analyzing internet content such as posts from social networks (Twitter, Facebook), various public discussions and blogs. The data acquired from these sources are subsequently used to analyze user opinions on specific products (aggregated from discussions about products) and current circumstances (from twitter posts and news). This data can be then used for targeted ads, or to influence public opinion.

Apart from using sentiment analysis on problems mentioned above we propose using sentiment analysis in human-robot interactions. Human-robot interactions are a hot topic recently. Studies indicate that one of the principal potential applications for robots may be for therapy affecting children with autism.

The next area of research connected with humans and robots is caring for the elderly. Its estimated that by 2050, the elderly will account for 16 percent of thr global population. That's 2.1 billion people over the age of 60. Caring for these seniors - physically, emotionally, and mentally - will be an enormous undertaking, and experts say there will be a shortage of trained professionals and those willing to take on the job. Robots may fill the gap, taking care of older people.

Accomplishing the task will require complex systems, able to communicate effectively without rejection. Robots must be able to adapt and behave according to specific situations. They must behave differently around children (as even individual

children are different) than around elderly. Regarding this we use term personalization, that means every individual will require specific approaches to become relaxed around robots.

II. SENTIMENT ANALYSIS

Sentiment analysis challenges projects in the field of natural language processing, computational linguistic and text mining. As we can see from the number of papers published by reputable conferences and journal papers in natural language process- ing and computational linguistics, it is an admittedly hot topic. The vital role is to deal with opinion, sentiment and subjectivity in text. [1] [2]

Sentiment analysis involves various research tasks [3], such as:

- **subjectivity detection** aiming to discover subjective or neutral terms, phrases or sentences and is frequently used as a prior step in polarity and intensity classifications, with the aim of separating subjective information from the objective.
- **polarity classification** attempts to classify texts into positive, negative or neutral terms. It forms the basis for determining the polarity of the text as a whole.
- **intensity classification** goes a step further and attempts to identify the different degrees of positivity and negativity; e.g. strongly-negative, negative, fair, positive, and strongly positive. It can be best described by numbers or words.
- **opinion spam** is another problem inhibiting accurate sentiment analysis. In recent years we noticed an increased demand for opinion classification, but almost no attention has been paid to examining the credibility of opinions in reviews. We distinguish three types of opinion spam: the first is a misguided opinion, the second an opinion that does not relate directly to the subject and the last is distortion, as in text not relevant to opinion analysis [4]
- **emotion detection** seeks to identify if a text conveys any type of emotion or not. It is similar to subjectivity detection. Within the scope of emotion detection we discriminate *emotion classification* - fine-grained classification of existing emotion in a text into one (or more) of a set of specific emotions (e.g. anger, fear, etc.), *emotion intensity* - degree or amount of an emotion (such as anger - very angry or sadness - slightly sad, etc.) [5], and finally *emotion cause detection* - extracting potential causes that lead to emotion expressions in text [6].

Sentiment analysis attempts to analyze and take advantage of extensive quantities of user-generated content. It enables the computer to understand text. Sentiment classification should be done on varied levels such as word/phrase, sentence and document categories.

Polarity classification can be applied to both opinion and emotion. Techniques and methods used for polarity classification can be useful for emotion classification as well for two reasons [7]:

- opinion and emotion are semantically related concepts. Generally, having an opinion towards an entity can cause

the person to feel an emotion in the same direction (positive or negative),

- these techniques often do not have any opinion-specific characteristics, and, they can directly be applied to emotionally labeled problems, too.

Considering this, we believe it is worth reviewing polarity classification methods before entering emotion research. There are two approaches towards sentiment analysis:

- the lexicon-based approach (lexicon-based methods, corpus-based methods)
- the text classification approach (machine learning methods, statistical methods)

III. EMOTION ANALYSIS

Emotion analysis can be viewed as a natural evolution of sentiment analysis and its more fine-grained model.

Digging deeper into psychology, we have to differentiate between terms *emotion, mood, feeling. Emotion* is an instantaneous perception of a feeling. They can be over in a matter of seconds to minutes, at most [8]. *Mood* is considered as a group of persisting feelings associated with evaluative and cognitive states which influence all the future evaluations, feelings and actions [9]. Unlike emotions, moods are non-intentional, though they may be elicited by a particular event or things. It is difficult to identify triggers causing mood however while in the state of a certain mood, the threshold is lowered for arousing related emotion. *Feeling* is mental associations and reactions to an emotion that are personal and acquired through experience.

A. Emotions Models

According to [10], three major directions in affect computing are recognized: categorical/discrete, dimensional and appraisals-based approaches.

1) **Basic emotion model** - The categorical approach - claims there are a small number of basic emotions that are hard-wired in our brain, and recognized across the world. Each affective state is classified into a single category. However, a couple of researchers proved that people show non-basic, subtle and rather complex affective states that could be impossible to handle, such as thinking, embarrassment or depression. Assigning text to a specific category can be done either manually or using learning-based techniques.

2) **Dimensional feeling model** - The dimensional approach - based on Wundts proposal that feelings (which he distinguishes from emotions) can be described as pleasantness/unpleasantness, excitement/inhibition and tension/relaxation, as well as Osgoods work on the dimensions of affective meaning (arousal, valence, and potency). Most recent models concentrate on only two dimensions, valence and arousal. Valence (pleasure/displeasure) depicts how positive or negative an emotion may be. Arousal (activation/deactivation) depicts how excited or apathetic an emotion is

3) **Componential appraisal models** - proposes that emotions are extracted from our appraisals (i.e., our evaluations, interpretations, and explanations) of events. These appraisals lead to different specific reactions in different people. OCC model is presumably the most widely accepted cognitive appraisal model for emotions [11] propose three aspects of the environment to which humans react emotionally: events of concern to oneself, agents that one considers responsible for such events, and objects of concern. It defines emotions as a valanced reaction to events, agents, and objects, and considers valenced reactions as a means to differentiate between emotions and nonemotions. This approach is very suitable for affect sensing from the text.

Despite the existence of various other models, the categorical and dimensional approaches are the most commonly used models for automatic analysis and prediction of affect in continuous input.

B. Approaches towards Emotion Analysis

Recent advances in the field of sentiment analysis and computational linguistics in general allow us to accomplish more advanced tasks such as emotion detection in documents. To detect emotion, researchers use generally known algorithms created for sentiment analysis. There are four major approaches to detecting emotions in text:

- **Keyword-based methods** - is the most intuitive approach. The main goal is to find out patterns similar to emotion keywords and match them.
- **Lexicon based methods** - classifies text using a lexicon (a knowledge-base with text labeled according to emotions) appropriate for the input dataset. This method is similar to the first one but instead of keywords list we are using emotion lexicon
- **Machine Learning methods** - Both supervised and unsupervised methods are used for emotion classification. For supervised methods, an annotated emotions dataset is used from which one learns which features are most salient to distinguish between classes. The dataset is divided into training and testing sets. Nave Bayes classifier, Support Vector Machine, MaxEntropy and Decision tree are the most used algorithms.
- **Hybrid methods** - combine any two or all three methods defined to achieve the benefit of multiple methods and reach the maximum level of accuracy

IV. HUMAN-ROBOT INTERACTION

Human-Robot Interaction (HRI) is a study of interaction dynamics between humans and robots, a multidisciplinary field that includes engineering (electrical, mechanical, industrial and design), computer science (human-computer interaction, artificial intelligence, robotics, natural language understanding, computer vision and speech recognition), social sciences (psychology, cognitive science, communications, anthropology, and human factors) and the humanities (ethics and philosophy).

Robots are poised to fill a growing number of roles in todays society, from factory automation to service applications, medical care and entertainment. While robots were initially used for repetitive tasks where all human direction is given a priori, they are becoming involved in increasingly more complex and less structured tasks and activities, including interaction with the humans required to complete those tasks. The fundamental goal of HRI is to develop the principles and algorithms for robot systems that enable safe and effective interaction with humans [12].

Interaction, by definition, means *"communication with each other or reacting to each another"* [1]. There are several possibilities for robots to communicate with humans. The way of communication is largely influenced by whether the human and robot are in close proximity to each other or not. Therefore, the interaction can be categorized into remote and proximate interaction. Within these two general categories we can differentiate applications that require mobility, physical manipulation, and social interaction [13].

- remote interaction - interaction where human and the robot are not co-located and are separated spatially or even temporally. Within this category we can distinguish:
  - teleoperation or supervisory control - remote interaction with mobile robots
  - telemanipulation - remote interaction with physical manipulators
- proximate interaction - interaction where the human and robot are co-located. Within proximate interaction we can according to [14] distinguish:
  - physical interaction - involves the research areas of manipulation and haptics, among others, and is used in medical and rehabilitation robotics
  - social/emotional interaction - includes social, emotive, and cognitive aspects of interaction. Involves research areas of assistive robotics, social robotics, and socially assistive robotics

A. Socialy Assistive Robotics

Social assistive robotics (SAR) is defined as the intersection of assistive robotics and socially interactive robotics. It is a comparatively new field of robotics that focuses on developing robots capable of assisting users through social rather than physical interaction. Social robots have to be able to perceive, interpret and respond appropriately to verbal and nonverbal cues from human. SAR compared with social robots, focuses on the challenges of providing motivation, education, therapy, coaching, training, and rehabilitation through nonphysical interaction. An effective socially assistive robot must understand and interact with its environment, exhibit social behaviour, focus its attention and communication on the user, sustain engagement with the user, and achieve specific assistive goals. The robot must do all of this in a way that is safe, ethical and effective for the potentially vulnerable user. SAR has been shown to have promise as a therapeutic tool for children, the

---

[1] https://dictionary.cambridge.org/dictionary/english/interaction

elderly, stroke patients, and other special-needs populations requiring personalized care.

1) Long-Term Interaction: Many applications with social robots involve only short-term interactions. However short-term interaction is not enough. Many real-world applications (e.g. education, therapy, companionship, and elderly care) call for keeping people interested for longer. We have to maintain human engagement and build relationship and trust between human and robot through adaptation and personalization.

Talking about long-term interaction we have to take into account *novelty effect*. Novelty effect, in the context of HRI, can be explained in such a way that interaction with robot can be initially highly triggering and engaging but after a couple of interactions, the newness wears off and people can lose an interest in interaction with the robot. To avoid such behaviour the challenge is to keep people engaged in the interaction and motivate them to interact longer (weeks, months or even years).

2) Personalization: Personalization is closely associated with long-term interaction mention above. It is another important research area in SAR. Personalization is an ability of the robot to adapt its behaviour to specific human, context, environment, and tasks.

We implemented personalized robot behaviour in our user-case scenario.

3) Artifical Companionship: So far, robot companions lack many important social and emotional abilities (e.g. recognising social affective expressions and states, understanding intentions, and accounting for the context of the situation, expressing appropriate social, affective behaviour) to engage with humans in natural interaction.

An artificial companion should be capable of evaluating how humans feel about the interaction, and how they interpret the agents actions and use this information to adapt its behaviour accordingly [15]. For instance, a robotic companion (Fig. 1) should act empathically towards a user if it detects that she is sad or not willing to engage in an interaction, e.g., it would not disturb them trying to engage them in some activity if they do not approach it.

4) Affective Loop: Another challenging research task in SAR is endowing the robot with an emotion intelligence. It is important that the interaction between human and robot would be affective, thus it must have the ability to perceive, interpret, express and regulate emotions.

Understanding human emotions by robot and at the same time having the option to express emotion back to human was defined by [16] as *affective loop* (AL). AL (see Fig. 2) is the interactive process in which the user [of the system] first expresses her emotions through some physical interaction involving her body, for example, through gestures or manipulations; and the system then responds by generating affective expression, using for example, colours, animations, and haptics which in turn affects the user (mind and body) making the user respond and step-by-step feel more and more involved with the system. [17]

Fig. 1. Robot companions; Humanoids in top row - from left to right 1) Zeno (Hanson Robotics), 2) Nao (Aldebaran Robotics), 3) Pepper (Aldebaran Robotics), 4) iCub (Italian Institute of Technology); Middle row - from left to right 1) Leonardo (MIT), 2) Kismet (MIT), iCat (Philips), 4) Buddy (Blue Frog Robotics); Bottom row - from left to right 1) Paro (AIST), 2) TEGA (MIT) 3) New AIBO (Sony)
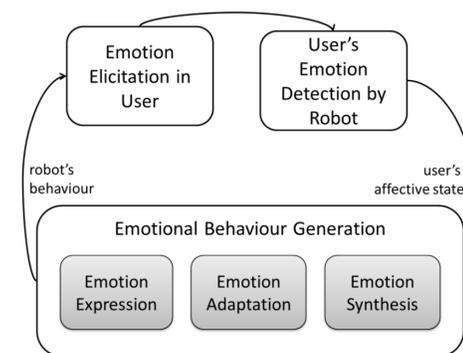

Fig. 2. Affective Loop adopted from [17]

Emotion detection is part of the broader area of Affective Computing (AC) with aims to enable computers recognize and express emotions [18]. AC defines emotion as playing an essential role in decision making and learning. Emotions influence the mechanisms of rational thinking. [18] highlighted several results from neurological literature that indicate emotions play a necessary role in human creativity and intelligence, as well as rational human thinking and decision-making.

Computers that interact naturally and intelligently with humans need at least the ability to recognize and express affect. Affect plays a key role in understanding such phenomena as attention, memory, and aesthetics. Emotion is necessary for creative behavior in humans. Neurological studies indicate that decision-making without emotion can be as impaired as that made with too much emotion. [18] argues affective computers should not only provide better performance in assisting humans, but also might enhance computers abilities to make decisions.

Therefore, one of the main goals of AC is enabling computers to understand human emotional state and adjust its response accordingly. Human emotional state can be expressed

either non-verbaly, verbaly, or both. Pioneer researcher in body language [19] found that within the realm of interpreting the affect or emotional state of others, we perceive 55% non-verbally (facial expression), 45% verbally out of which 38% by speech (tone of voice, inflection, and other sounds) and 7% by words.

On the other hand, how and when, machines should exhibit emotions is also an important research question. Herewith with this is closely linked *synthetic emotion*. Synthetic emotion is an emotion produced by robot. Integration of different modalities, when they are congruent and synchronous, leads to a significant increase in human emotion recognition accuracy [20]. However, when information is incongruent across different sensory modalities, integration may lead to a biased percept, and emotion recognition accuracy is impaired [20].

V. Proposing a new learning approach for emotion detection in texts

We propose a new learning algorithm based on lexicon methods and machine learning methods.

A. Data

We build our own English and Slovak corpus consisting of Aesop's fable. Fables were downloaded [2] [3] [4], cleaned and saved into .txt documents. Each document represented one fable. In total we have 740 English fables and 35 Slovak fables.

Fable is a short story with moral truth, using animals as main characters. We wanted these stories to be read in the human-robot experiment scenario. To keep the audience interested and to stay focused, the text should be neither long nor very short and interesting.

In the first stage, fables were annotated by human annotators. We programmed the web-based application to create dictionaries with information about polarity and emotions. We are considering the wisdom of the crowd to be the most straightforward way of obtaining data from users. Besides the fables, users were allowed to input any other text they liked into the system.

Forty subjects above age 23 took part in our experiments. Everybody was asked to mark words in Aesop fables that, in their opinion, carry polarity or emotion. The polarity ranged from -5 (the most negative word) to 5 (the most positive word). For emotion, we were using Ekmans six emotions: happiness, sadness, surprise, fear, disgust, and anger. The emotion of every word could differ from one emotion to all six emotions.

Analysis of Obtained Slovak Dictionaries

We obtained a total of 3975 entries from our subjects. We then created two dictionaries: the first dictionary contained all words (general dictionary); the second dictionary contained emotional words only (emotional dictionary). Emotional words have to have at least one emotion out of six marked. The number of unique words in the first dictionary was 1017. The number of words in the emotional dictionary was 747.

---

[2] http://www.aesopfables.com

[3] http://read.gov/aesop/

[4] https://www.klasici.sk/node/146

Consecutive we calculated an agreement between annotators on each word. Because annotated words were not the same for every annotator and we had multiple annotators, we decided to use *Krippendorff's alpha α*. Krippendorff's alpha (KA) is a reliability coefficient developed to measure the agreement among observers [21].

It is customary to require $\alpha \geq 0.800$. Where tentative conclusions are still acceptable, $\alpha \geq 0.667$ is the lowest conceivable limit [22].

Applying KA to our dictionaries we obtained:

- General dictionary: Words in agreement between annotators: 67 (α above 0.667 threshold). Words in disagreement between annotators: 385. Words in systematic disagreement between annotators: 155 (α below 0). Word with only one annotator: 410

- Emotional dictionary: Words in agreement between annotators: 138. Words in disagreement between annotators: 217. Words in systematic disagreement between annotators: 80. Emotional words with one annotator: 312. Entries with only one emotion: 1065 (number of entries appertaining to each emotion are shown in Tab. **??**); with two emotions: 323; with three emotions: 69; with four: 10; and lastly with six: 1.

We calculated emotion for every Aesops fable individually for every subject.

As we can see from Fig. 3 our subjects do not have the same opinion about the same Aesops fable. The number of emotions in every fable differs from one to all six emotions. The reason for such variation can lay in many aspects such as different points of view about read text, uncertainty which word to mark, etc. On the other hand, we can see that sometimes one emotion outweighs others remarkably (e.g. fear in first and ninth fable and sadness in tenth fable). Another considerable thing to say is that every fable consists mostly of negative emotion (fear and sadness). That may be caused by the nature of fable animals have bad characteristics of people with the intention to enlighten readers.
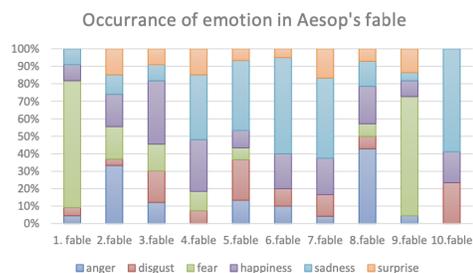

Fig. 3. Percentage of emotions occurrence in fables

For our first experiment it was not crucial to get that align (high KA) rather we invite diversity. We could therefore argue that we justified building personalised emotion dictionaries. Based on this result we can conclude that **making emotional behavior on the robot will have an impact on the human-robot interaction**.

English Fables

Corpus of English fables consisted of 393 annotated sentences and 999 unannotated. Further we will discuss only annotated sentences. Sentences were annotated into 9 categories (Plutchik's eight emotions: joy, trust, sadness, fear, disgust, anger, anticipation, surprise and "No emotion" category). The number of emotions chosen for each sentence was arbitrary. Having more than one emotion for sentence means, that we are dealing with **multi-label** classification problem. There is no evidence of positive/negative relationship between emotion's classes (Fig. 4).
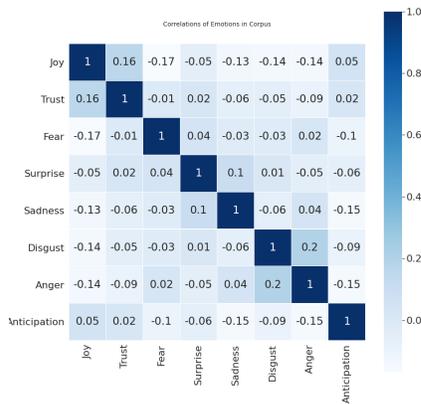

Fig. 4. Correlation of emotion's classes in dataset

B. Processing of the Data

English Fables

Fables were formatted as follows: one sentence = one row in a document. Firstly we lowered every character; applied function for dividing shortened forms of words into two words (grammatical contractions - $we're \rightarrow we\ are$); and cleaned from interpunction (sign of question mark, colon, and an exclamation mark was used as features). Every sentence was tokenized into words. Afterward Part-Of-Speech (POS) tagger was applied. Next, we applied the NRC dictionary to find out if any given word is a word from the vocabulary if yes we assign emotion to the word. Following by stopwords removal and lemmatization of the words (keeping words in their root form).

C. Feature extraction and Word Embeddings

Every sentence was represented by a vector of features. Vectors were constructed as follows:

- POS tagging (4 features): we used *pos_tag* function from *nltk* library. Every POS tag was represented as a number of occurrences in a given sentence. We chose 14 POS labels grouped into four categories:
  - noun: NN noun, singular, NNS noun plural
  - adjective: JJ adjective, JJR adjective, comparative, JJS adjective, superlative

  - verb: VB verb, base form, VBD verb, past tense, VBG verb, gerund/present participle, VBN verb, past participle, VBP verb, sing. present, non-3d, VBZ verb, 3rd person sing. present
  - adverb: RB adverb, RBR adverb, comparative, RBS adverb, superlative

- emotion (8 features): we used the NRC dictionary to obtain counts of emotions occurring in each sentence - anger, fear, sadness, disgust, surprise, anticipation, trust, and joy
- punctuation (3 features): we extracted presence of colons " : ", exclamation marks "!", and question marks "?"
- numerical feature vector: we extracted all words from pre-processed sentences to create vocabulary:
  - Bag-of-Words representation (number of features was dependent on thresholding occurrence of tokens in input): each sentence was represented as a number of occurrence of given words in the vocabulary. Vocabulary was generated from all tokens in sentences.
  - TF/IDF (number of features was dependent on thresholding occurrence of tokens in input): similar to BoW, but instead of the number of occurrences, each token was represented as a proportion between the number of occurrence in given sentence and occurrence in the whole corpus
  - sentence embeddings (300 features): every word (token) in a sentence is represented by its vector obtained from pretrained ConceptNet Numberbatch model. We used word embeddings to create sentence embeddings. Sentence embeddings are basically averaged sum of word embeddings vectors appertaining to the sentence.

We used *k-means* clustering algorithm. It is well-known algorithm that approximates the maximum-likelihood solution for determining the locations of the means of a mixture density of component densities. It is special case of self-organizing-maps. The algorithm follows equation:

$$E(em_1,....em_K) = \frac{1}{S} \sum_{k=1}^{K} \sum_{w_n \in EM} ||w_n - em_k||^2 \quad (1)$$

where:
- S - number of sentences in dataset
- W - vector representation of sentences
- K - number of emotion categories
- EM_k - clusters of sentences
- S_k - number of sentences in each cluster
- em_k - center of clusters

The outcome of the algorithm are clustered data annotated according to the centroid where they belong.

Because annotation of sentences is exhausting and time-consuming we decided to utilize this approach to annotate additional data based on k-means. We randomly chose 5 representatives of each class (e.g. in class joy - 5 representatives for 0 category and 5 representatives for 1 category) and calculated

centroid. Centroids were calculated as an average of the sum of vectors (our data had vector representation). We ended up with 18 centroids. Before every pair of centroid was fed into the k-means algorithm, we calculated the distance of every sentence from given centroids and removed the furthest and closest one. After that, labels for every class were predicted.

D. Classification

Our approach to the classification lies in transforming our problem into 9 separate problems (8 emotion classes and one class without emotion). Based on the fact that emotions are not dependent on each other we trained classifiers separately. When a new sample comes into classification, all of them estimate the probability for each class. Every classificator has only one vote. The threshold is set to be 50% for accepting label.

E. Evaluation of Results

To evaluate results, we used statistical metrics usually used in text documentation: precision, recall, F1 score, Matthews correlation coefficient, and subset accuracy.

1) Baseline: Our baseline model consisted of NRC dictionary and 393 annotated sentences. We matched every word against the dictionary and assigned the number of appertaining occurrences to each emotion. Later we transformed the number of occurrences to binary representation (0 if an emotion is not present, 1 if emotion has more than one occurrence). Tab. I shows that out of eight emotion, *Joy* is classified most accurately and *Disgust* with *Trust* the worst. The reason for it lies in our data. Looking back at Fig. **??** we can see that trust and disgust are the least represented classes.

TABLE I
ACCURACY OF EMOTION DICTIONARY, LEXICON APPROACH

	Precision	Recall	F1 score	MCC
Joy	0.32	0.78	0.45	0.33
Trust	0.03	0.55	0.06	0.01
Fear	0.15	0.77	0.25	0.1
Surprise	0.20	0.38	0.27	0.07
Sadness	0.18	0.69	0.28	0.16
Disgust	0.08	0.54	0.14	0.04
Anger	0.22	0.65	0.33	0.11
Anticipation	0.2	0.69	0.31	0.15
No emotion	0.59	0.21	0.31	0.23
F_1 micro	0.18	0.55	0.27	
F_1 macro	0.22	0.58	0.27	
Exact Accuracy	0.22			

2) Experiment 1: We firstly begin by testing our data against the Bag-of-Words representation (Tab. II). As we can see, precision is low. Above-average results gives only *No emotion* class. *Trust and Disgust* got 0. We experimented with several model's setups such as:

- changing threshold for minimal/maximal count of the word to be excluded from vocabulary when creating BoW representation
- uni-grams, bi-grams
- stopwords removing/not removing

- changing the number of additional features for sentence representation (NRC, POS, Punctuation)
- changing classifiers Multinomial Naive Bayes, Bernoulli Naive Bayes, Gaussian Naive Bayes, SVM
- expanding training set for data annotated by K-means

TABLE II
ACCURACY OF BAG-OF-WORDS REPRESENTATION, MULTINOMIAL NAIVE BAYES CLASSIFIER

	Precision	Recall	F1 score	MCC
Joy	0.36	0.25	0.29	0.18
Trust	0.00	0.00	0.00	0.00
Fear	0.10	0.08	0.09	-0.01
Surprise	0.20	0.15	0.17	0.03
Sadness	0.25	0.07	0.11	0.07
Disgust	0.00	0.00	0.00	-0.04
Anger	0.14	0.09	0.11	-0.04
Anticipation	0.20	0.06	0.09	0.03
No emotion	0.67	0.29	0.41	0.32
F_1 micro	0.29	0.15	0.20	
F_1 macro	0.21	0.11	0.14	
Exact Accuracy	0.15			

Fine-tuning with different setting such as stopwords removing/not removing, uni-grams/bi-grams, and the threshold for minimal/maximal count of a word to be excluded from vocabulary we improved *Joy* and *Anticipation* precision.

Next we added our proposed features - emotions from NRC dictionary, POS taggs, punctuation and continued tuning our model. We saw improvement on *No emotion* and *Anger* classes.

We tried also every feature individually. We noticed increase in accuracy of *Fear* to 25%, *Surprise* to 50%, *Sadness* to 50%, and *Disgust* to 25%. To increase the accuracy, we needed to use different setup for every class.

Adding k-means annotated data to the training set we observe *Disgust* accuracy to rise to 67%. All other accuracies stayed at the same level.

3) Experiment 2: Foundation of experiment 2 was the TfIDF representation of sentences. The highest score was obtained in *Joy* class. The lowest in *Trust* and *Disgust* class.

After fine-tuning the parameters of our model, we trained the model and compare results. We see improvement on *Joy, Sadness and Anger* classes in comparison with BoW approach.

Adding proposed features raised accuracy on *Joy, Trust, Fear, Anger Anticipation and No emotion* classes.

Adding additional semiautomatically labeled data, further raised the accuracy of *Sadness* to 67%.

4) Experiment 3: The base of experiment 3 were sentence embeddings. On top of that, we add emotional dictionary (NRC), punctuation, POS.

Lastly we used word embeddings - ConceptNet Numberbatch and converted them to *sentence embeddings*. We can see from tab.III, that accuracy in classes is low but it covers all classes except one - *Trust*.

Adding features to the model did not help to raise its accuracy markantly. Adding data labelled by k-means helped raise accuracy in class *No emotion* to 68% by using SVM

	Precision	Recall	F1 score	MCC
Joy	0.19	0.55	0.28	0.05
Trust	0.00	0.00	0.00	-0.04
Fear	0.15	0.46	0.23	0.09
Surprise	0.24	0.65	0.35	0.17
Sadness	0.20	0.67	0.31	0.19
Disgust	0.14	0.38	0.20	0.12
Anger	0.23	0.23	0.23	0.04
Anticipation	0.19	0.53	0.28	0.10
No emotion	0.31	0.38	0.34	0.03
F_1 micro	0.20	0.46	0.28	
F_1 macro	0.18	0.43	0.25	
Exact Accuracy	0.20			

classifier. The average accuracy for the rest of the classes was 20%. Which is not sufficient.

5) Ensemble of Classifier: We combined the best obtained result for each class and made ensemble classifier. We can see increase in exact accuracy to 31% which is the most strict metric and expresses how many completely correct rows (all labels are correct) we obtained from classifier.

VI. THE APPLICATION AND TESTING ENVIRONMENT FOR HRI

We propose scenarios with humanoid robot Nao and humans (either kids, or adults).

The controlled group was the same for each experiment. ThegGroup consisted of 8 participants (7 adults and 1 child). The age of the participants ranged from 3 to 50 years. The participant was interacting with a robot alone thus it was one on one interaction. During the experiments we paid attention to two variables: length of the interaction, number of fables red.

Throughout the experiments we used NAO robot v.5. NAO is a humanoid robot often utilized in HRI experiments. He can move with hands, walk, talk, listen. Taking into account its' very limited facial expression, he can make use of his eye' s led lights to signal blinking, even changing color can suggest different emotional states (e.g. red led = anger).

A. Experiment 1A

Setup of the first experiments is straightforward (Fig. 5). Nao is presented as "Narrator". He greets the participant of the experiment and asked him to sit down, facing him. Subsequently, he offers a telling story. He starts narrating as soon as he hears "yes". Input to the Nao is fable without any emotional markup, thus Nao is reading fable without any expression (either movement or vocal). Recipient is facing Nao and listening to the story. After telling the whole story, Nao gives the option either to continue with another story or to finish. The number of stories is fully dependent on the participant. At the end of the experiment, we give every participant with questions shown in tab. IV.

We can break down our system to the following parts:

- Fable - Nao randomly picks one fable from a given set of 145 fables. Fables are preprocessed to the sentences.

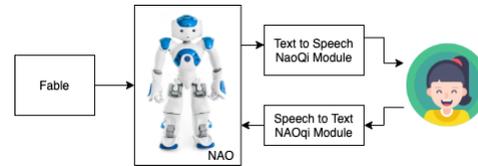| Questions	
Q1.	I like robot Nao narrating the story.
Q2.	I like to hear the story again.
Q3.	Nao was believable narrator.
Q4.	Experiment was not interessting.
Q5.	I was not lacking anything in the robot perfomance.

Fig. 5. Setup for the experiment 1A

- Text to Speech NAOqi Module - converts fable to speech. Input are sentences. Robot is not tracking human while telling a story.
- Speech to Text NAOqi Module - user can communicate short commands via this interface. It is used when Nao asks whether he can start telling the story or at the end of the story if the participant wants another story.
- Human - can request more stories

B. Experiment 1B

Setup for the second experiment was similar to the first experiment with three exceptions. Number one: The input to the Nao is Aesop's fable marked with emotion. Second is closely connected to the first: Nao is narrating the story with movements and changes in pitch. The third difference is in case the participant wants to hear another story. After requested second story, Nao is telling that he is tired and asks if the participator really wants to hear another story. If he gets a positive response he continues, otherwise he thanks and the experiment is finished. At the end of the experiment, participant fills in the survey with the same question as before (tab. IV).

C. Experiment 1C

We took setup from experiment 2, removed classification block and modified block *Generating script for Nao text-to-speech and emotional gestures* to generate any gestures, incongruent to the emotions in written text.

D. Results of the Experiment 1

For responses we used a five-point Likert scale with options: 5 - I agree extremely; 4- I agree very; 3 - I agree moderately; 2 - I agree slightly; 1- I do not agree. We took an average of scores for each question. The average length of the interaction was measured from the point where Nao robot greeted the person until he finished narrating his last fable rounded to the minutes. The average number of fables read indicates how many fables were read during one session.

From the results above we can conclude, that robot with emotional/random cues (experiment *1B,1C*) achieved better

overall rating in comparison to the robot without emotional cues (experiment 1A). We demonstrate that there is a difference in perceiving text from robot to human by adding emotional/random manners to robot.

VII. CONCLUSION

Presented work connects two big areas of research namely Sentiment analysis and Human-robot interaction. We saw a gap in HRI years ago that SA could fulfill.

To demonstrate our claim for emotion detection in text within HRI we conducted experiments with humanoid robot NAO. We proposed quantitative research with surveys and trackable variables during the experiment (length of interaction and number of fables read) and qualitative research by asking our participants about the experiment to measure improved robot to human interaction. The results of the experiments show, there is indeed positive feedback on the human side. From the questionnaire results, it's obvious adding gestures to robot increase positivity in interaction.

The results from emotion detection in text implies an increase in precision and accuracy for each label. Adding additional features from emotional dictionary raised accuracy in some classes more, in some classes less. The biggest increase of accuracy can be seen in class *Disgust* followed by *Joy, Anger, No emotion, ANticipation, Sadness*. The rest of the classes have accuracy equals to or lower than 50%. Lastly we observed a change in testing precision and accuracy when we added new data, annotated by k-means algorithm. Despite the sawing increase in accuracy it did not reach the level that would be acceptable for us (at least 50%).

We see potential based on the obtained results, in utilization automatic emotion detection from text in human-robot interaction. As experiment 1C showed, the system did not have to be 100% percent accurate to arouse a positive response from human. We can take a look from another angle as well: not showing happy gestures when the perceived emotion should be sad and vice-versa. That can transform into a classification, as a problem where not occurrence of emotion should be observed.

ACKNOWLEDGMENT

REFERENCES

[1] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 12, pp. 1–135, jan 2008.

[2] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics*, vol. 37, no. September 2010, pp. 1–41, jun 2011.

[3] J. C. de Albornoz, L. Plaza, and P. Gervás, "SentiSense: An easily scalable concept-based affective lexicon for sentiment analysis," in *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. European Language Resources Association (ELRA), 2012, pp. 23–25.

[4] P. Szabo and K. Machova, "Various approaches to the opinion classification problems solving," in *2012 IEEE 10th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*. IEEE, jan 2012, pp. 59–62.

[5] S. Mohammad and F. Bravo-Marquez, "WASSA-2017 Shared Task on Emotion Intensity," 2018, pp. 34–49.

[6] L. Gui, R. Xu, Q. Lu, D. Wu, and Y. Zhou, "Emotion cause extraction, a challenging task with corpus construction," in *Communications in Computer and Information Science*, vol. 669. Springer Verlag, 2016, pp. 98–109.

[7] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining," *ACM Computing Surveys*, vol. 50, no. 2, may 2017.

[8] P. Ekman, "Expression and the nature of emotion," 1984.

[9] I. Amado-Boccara, D. Donnet, and J. P. Olié, "[The concept of mood in psychology]." *L'Encephale*, vol. 19, no. 2, pp. 117–22.

[10] D. Grandjean, D. Sander, and K. R. Scherer, "Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization." *Consciousness and cognition*, vol. 17, no. 2, pp. 484–95, jun 2008.

[11] A. Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge University Press, jul 1988.

[12] D. Feil-seifer and M. J. Mataric, "Human-Robot Interaction," pp. 1–20.

[13] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," pp. 203–275, 2007.

[14] M. J. Matarić and B. Scassellati, "Socially assistive robotics," in *Springer Handbook of Robotics*. Cham: Springer International Publishing, jan 2016, pp. 1973–1993.

[15] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, "Affect recognition for interactive companions: Challenges and design in real world scenarios," *Journal on Multimodal User Interfaces*, vol. 3, no. 1, pp. 89–98, mar 2010.

[16] K. Höök, "Affective loop experiences: designing for interactional embodiment," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3585–3595, dec 2009.

[17] A. Paiva, I. Leite, and T. Ribeiro, "Emotion Modelling for Social Robots," in *The Oxford Handbook of Affective Computing*. Oxford University Press, 2014.

[18] R. W. Picard, *Affective Computing*, R. W. Picard, Ed. MIT Press, 1997, vol. 73, no. 321.

[19] A. Mehrabian, *Nonverbal Communication*. New Brunswick NJ: Aldine Transaction, 1972.

[20] M. A. N. Mahani, S. Sheybani, K. M. Bausenhart, R. Ulrich, and M. N. Ahmadabadi, "Multisensory Perception of Contradictory Information in an Environment of Varying Reliability: Evidence for Conscious Perception and Optimal Causal Inference," *Scientific Reports*, vol. 7, no. 1, dec 2017.

[21] K. Krippendorff, "Computing Krippendorff ' s Alpha-Reliability," *Departmental Papers (ASC)*, p. 12, 2011.

[22] ——, "Reliability in Content Analysis." *Human Communication Research*, vol. 30, no. 3, pp. 411–433, jul 2004.

QuoVadis Research @ FEI