Univerzitet Crne Gore
Prirodno-matematički fakultet

Džordža Vašingtona b.b.
1000 Podgorica, Crna Gora

tel: +382 (0)20 245 204
fax: +382 (0)20 245 204
www.pmf.ac.me

Broj: 2024/01-445/1
Datum: 23.02.2024.

UNIVERZITET CRNE GORE

SENATU

CENTRU ZA DOKTORSKE STUDIJE

U prilogu akta dostavljamo Predlog Odluke sa CX sjednice Vijeća PMF-a održane 20.02.2024. godine.

S poštovanjem

UNIVERZITET CRNE GORE Prirodno-matematički fakultet

Dekan,
Prof. dr Miljan Bigović

Univerzitet Crne Gore
Prirodno-matematički fakultet

Džordža Vašingtona b.b.
1000 Podgorica, Crna Gora

tel: +382 (0)20 245 204
fax: +382 (0)20 245 204
www.pmf.ac.me

Broj: 2024/01-445
Datum: 23.02.2024

Na osnovu člana 64 Statuta Univerziteta Crne Gore, a u vezi sa članom 41 stav 1 Pravila doktorskih studija, Predloga komisije za doktorske studije broj 2024/01-365/2 od 19.02.2024.godine, na CX sjednici održanoj dana 20.02.2024.godine, Vijeća je donijelo

O D L U K U

I

Utvrđuje se da su ispunjeni uslovi iz člana 38 Pravila doktorskih studija za doktoranda Msc Kostu Pavlovića.

II

Predlaže se Odboru za doktorske studije sastav komisije za ocjenu doktorske disertacije:

1. Prof. dr Milenko Mosurović, redovni profesor Prirodno-matematičkog fakulteta Univerziteta Crne Gore (naučna oblast: Vještačka inteligencija);
2. Prof. dr Igor Đurović, redovni profesor Elektrotehničkog fakulteta Univerziteta Crne Gore (naučna oblast: Obrada signala);
3. Prof. dr Goran Kvaščev, vanredni profesor Elektrotehničkog fakulteta Univerziteta u Beogradu (naučna oblast: Vještačka inteligencija);
4. Prof. dr Vesna Popović Bugarin, redovni profesor Elektrotehniočkog fakulteta Univerziteta u Beogradu (naučna oblast: Vještačka inteligencija);
5. Doc. dr Igor Jovančević, docent Prirodno-matematičkog fakulteta Univerziteta Crne Gore (naučna oblast. Računarska vizija).

III

Odluka se dostavlja Odboru za doktorske studije Univerziteta Crne Gore.

DEKAN

Prof. dr Miljan Bigović

Univerzitet Crne Gore Prirodno-matematički fakultet

Crna Gora
UNIVERZITET CRNE GORE
PRIRODNO-MATEMATIČKI FAKULTET
Broj 1024/01-247
Podgorica, 06.02. 2024 god.

UCG

UNIVERZITET CRNE GORE
ObrazacD2: Ispunjenost uslova doktoranda

ISPUNJENOST USLOVA DOKTORANDA

OPŠTI PODACI O DOKTORANDU			
Titula, ime, ime roditelja, prezime	MSc Kosta, Jovan, Pavlović		
Fakultet	Prirodno-matematički fakultet		
Studijski program	Računarske nauke		
Broj indeksa	1/2018		
NAZIV DOKTORSKE DISERTACIJE			
Na službenom jeziku	Umetanje vodenih žigova u digitalne audio signale korišćenjem dubokih neuronskih mreža		
Na engleskom jeziku	Digital audio watermarking using deep neural networks		
Naučna oblast	Vještačka inteligencija		
MENTOR/MENTORI			
Prvi mentor	Prof. dr Igor Đurović	Elektrotehnički fakultet, Univerzitet Crne Gore, Crna Gora	Obrada signala
KOMISIJA ZA PREGLED I OCJENU DOKTORSKE DISERTACIJE			
Prof. dr Milenko Mosurović		Prirodno-matematički fakultet, Univerzitet Crne Gore, Crna Gora	Vještačka inteligencija
Prof. dr Igor Đurović		Elektrotehnički fakultet, Univerzitet Crne Gore, Crna Gora	Obrada signala
Prof. dr Goran Kvaščev		Elektrotehnički fakultet, Univerziteta u Beogradu, Srbija	Vještačka inteligencija
Prof. dr Vesna Popović-Bugarin		Elektrotehnički fakultet, Univerzitet Crne Gore, Crna Gora	Vještačka inteligencija
Doc. dr Igor Jovančević		Prirodno-matematički fakultet, Univerzitet Crne Gore, Crna Gora	Računarska vizija
Datum značajni za ocjenu doktorske disertacije			
Sjednica Senata na kojoj je data saglasnost na ocjenu teme i kandidata		13.5.'21. g.	

UNIVERZITET CRNE GORE
ObrazacD2: Ispunjenost uslova doktoranda

Dostavljanja doktorske disertacije organizacionoj jedinici i saglasanost mentora	6. 2. 2024. g.
Sjednica Vijeća organizacione jedinice na kojoj je dat predlog za imenovanje komisija za pregled i ocjenu doktorske disertacije	20. 2. 2024. g.

ISPUNJENOST USLOVA DOKTORANDA

U skladu sa članom 38 pravila doktorskih studija kandidat **je** cjelokupna ili dio sopstvenih istraživanja vezanih za doktorsku disertaciju publikovao u časopisu sa **(SCI/SCIE)/(SSCI/A&HCI)** liste kao prvi autor.

Spisak radova doktoranda iz oblasti doktorskih studija koje je publikovao u časopisima sa (SCI/SCIE) liste

1. Kosta Pavlović, Slavko Kovačević, Igor Djurović, Adam Wojciechowski, Robust speech watermarking by a jointly trained embedder and detector using a DNN, Digital Signal Processing, Elsevier, Volume 122, 103381, 2022, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2021.103381
2. Kosta Pavlović, Slavko Kovačević, Igor Djurović, Adam Wojciechowski, DNN-based speech watermarking resistant to desynchronization attacks, International Journal of Wavelets, Multiresolution and Information Processing, World Scientific, Volume 21, Issue 5, 2350009, 2023, ISSN 1793-690X, https://doi.org/10.1142/S0219691323500091

Obrazloženje mentora o korišćenju doktorske disertacije u publikovanim radovima

U radu objavljenom u časopisu *Digital Signal Processing* kandidat je predložio sistem vodenog žiga za digitalne audio signale u okviru kojeg se zadaci umetanja i detekcije vodenih žigova obavljaju dubokim neuronskim mrežama. Predloženi sistem je obučen na korpusu govornih signala i demonstrirana je njegova otpornost na osnovne audio efekte.

U radu objavljenom u časopisu *International Journal of Wavelets, Multiresolution and Information Processing* prezentovana su proširenja prethodno kreiranog sistema vodenog žiga za digitalne audio signale, koja su ga učinila otpornim na grupu desinhronizujućih efekata.

Datum i ovjera (pečat i potpis odgovorne osobe)

U Podgorici,
07.02.2024

UNIVERZITET CRNE GORE Prirodno-matematički fakultet UCG

MP

za DEKAN

Prilog dokumenta sadrži:

1. Potvrdu o predaji doktorske disertacije organizacionoj jedinici
2. Odluku o imenovanju komisije za pregled i ocjenu doktorske disertacije
3. Kopiju rada publikovanog u časopisu sa odgovarajuće liste
4. Biografiju i bibliografiju kandidata

5. Biografiju i bibliografiju članova komisije za pregled i ocjenu doktorske disertacije sa potvrdom o izboru u odgovarajuće akademsko zvanje i potvrdom da barem jedan član komisije nije u radnom odnosu na Univerzitetu Crne Gore

Univerzitet Crne Gore
Prirodno-matematički fakultet

Džordža Vašingtona b.b.
1000 Podgorica, Crna Gora

tel: +382 (0)20 245 204
fax: +382 (0)20 245 204
www.pmf.ac.me

Broj: 2024/01-246
Datum: 06.02.2024god

Na osnovu člana 33 Zakona o upravnom postupku, nakon uvida u službenu evidenciju, Prirodno-matematički fakultet izdaje

POTVRDU

MSc Kosta Pavlović, student doktorskih studija na Prirodno-matematičkom fakultetu u Podgorici, dana 06.02.2024.godine dostavio je ovom fakultetu doktorsku disertaciju pod nazivom ***"Umetanje vodenih žigova u digitalne audio signale korišćenjem dubokih neuronskih mreža"*** na dalje postupanje.

UNIVERZITET CRNE GORE Prirodno-matematički fakultet UCG

DEKAN
Prof. dr Miljan Bigović

Crna Gora
UNIVERZITET CRNE GORE
PRIRODNO-MATEMATIČKI FAKULTET
Broj 2024/01-247/1
Podgorica, 06. 02. 2024 god.

Univerzitet Crne Gore

Prirodno-matematički fakultet

Na osnovu člana 37 Pravila doktorskih studija Univerziteta Crne Gore dajem

SAGLASNOST

da doktorska disertacija pod naslovom „Umetanje vodenih žigova u digitalne audio signale korišćenjem dubokih neuronskih mreža" kandidate MSc Koste Pavlovića zadovoljava kriterijume propisane Statutom Univerziteta Crne Gore i Pravilima doktorskih studija.

Mentor

Prof. dr Igor Đurović

Curriculum vitae

PERSONAL INFORMATION

Kosta Pavlović

Generala Anđelića 3/A, Kolašin, Montenegro

+382 69 528 960

kosta@ucg.ac.me

www.github.com/kosta-pmf/

www.linkedin.com/in/kosta-pavlovic/

ORCID 0000-0001-6838-6585

Gender Male | Date of birth 8 May 1994 | Nationality Montenegrin

WORK EXPERIENCE

October 2016 – Present

Teaching and research assistant

Faculty of Natural Sciences and Mathematics, University of Montenegro

Teaching assistant for the following subjects: Artificial Intelligence, Algorithms and Data Structures, Programming, Bioinformatics, Software Engineering, Information Systems, Distributed Computing

October 2016 – Present

Instructor

DoMEn d.o.o

Preparing high school students for the International Olympiad in Informatics and the Balkan Olympiad in Informatics

July 2022 – Present

Machine learning engineer

Bixbit d.o.o

- Developed deep learning based web application for automatic license plate recognition.
- Developed deep learning based mobile application for automatic face detection and recognition.
- Expanded automatic proofreading software with natural language processing tehniques.

October 2016 - December 2019

Consultant

Datum Solutions

- Participated in developing a system for automatic document classification.
- Participated in developing a system for automatic document indexing and developed several tools for extracting information from scanned documents.
- Performed consulting activities in the development of a system for license plates recognition.
- Performed consulting activities in several data science projects for healthcare and finance.

PROJECTS

January 2022 - July 2022

Technical lead

BonsAPPs Consortium

- Developed a deep learning based system for automatic detection and segmentation of defects on images of industrial parts.
- Compressed and optimized trained model for deployment on edge devices.

December 2020 - February 2021

Software developer

Institute of Public Health of Montenegro

Created data warehouse and developed a web application for storing and online analytical processing of records of citizens tested for the Covid-19 virus to track the development of the pandemic in Montenegro.

EDUCATION AND TRAINING

November 2018 – Present

PhD - Thesis Title: "Digital Audio Watermarking Using Deep Neural Networks"

Faculty of Natural Sciences and Mathematics, University of Montenegro

October 2016 - October 2018

Master of Computer Science

Faculty of Natural Sciences and Mathematics, University of Montenegro

September 2012 - July 2015

Bachelor of Computer Science

Faculty of Natural Sciences and Mathematics, University of Montenegro

PERSONAL SKILLS

Mother tongue: Serbian

Other languages

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C1	C2	B2	B2	C1

Levels: A1 and A2: Basic user – B1 and B2: Independent user – C1 and C2: Proficient user
Common European Framework of Reference for Languages

Organisational / managerial skills
- Led a team of 8 members in developing and implementing an automatic defect detection system for industrial parts.

Digital competences

SELF-ASSESSMENT				
Information Processing	Communication	Content creation	Safety	Problem solving
Proficient user	Proficient user	Proficient user	Proficient user	Proficient user

Digital competences - Self-assessment grid

Computer skills
- Programming languages: Python, Java, C#, PHP, C, C++, Matlab, R, Javascript, Prolog
- Frameworks: Tensorflow, Express, Flask, Django, Eclipse Modeling Framework
- Tools and libraries: OpenCV, scikit-learn, RabbitMQ, Pentaho tools
- SQL, HTML, CSS, Git, LaTeX, Microsoft Office Tools

Other skills
- Strong mathematical background.
- Analytical approach to problem solving.
- I find joy in regularly playing football and swimming, showcasing solid skills and unwavering dedication to both sports.
- Consistently ranked within the top 0,3% of Fantasy Premier League managers for three consecutive years.

Driving licence: B

PUBLICATIONS

[1] Kosta Pavlović, Slavko Kovačević, Igor Djurović, and Adam Wojciechowski. "DNN-based speech watermarking resistant to desynchronization attacks". In: *International Journal of Wavelets, Multiresolution and Information Processing* (Mar. 2023). URL: https://doi.org/10.1142/s0219691323500091.

[2] Kosta Pavlović, Slavko Kovačević, Igor Djurović, and Adam Wojciechowski. "Robust speech watermarking by a jointly trained embedder and detector using a DNN". In: *Digital Signal Processing* 122 (2022), p. 103381. URL: https://www.sciencedirect.com/science/article/pii/S1051200421004206.

[3] Kosta Pavlović, Slavko Kovačević, and Igor Djurović. "Speech watermarking using Deep Neural Networks". In: *Proceedings of 28th Telecommunications Forum (TELFOR)*. 2020, pp. 292–295.

[4] Savo Tomovic, Kosta Pavlovic, and Milija Bajceta. "Aligning document layouts extracted with different OCR engines with clustering approach". In: *Egyptian Informatics Journal* (2020). URL: https://www.sciencedirect.com/science/article/pii/S1110866520301638.

[5] Savo Tomovic and Kosta Pavlovic. *Long life learning system for document understanding: Document understanding in cognitive manner*. Scholars' Press, 2020.

[6] Savo Tomovic and Kosta Pavlovic. "Cognitive Approach in Document Indexing". In: *Eastern European Journal for Regional Studies (EEJRS)* 4.1 (2018), pp. 67–75.

[7] Kosta Pavlovic and Aleksandar Popović. "Jezik MetaR". In: *Informacione Tehnologije, Žabljak*. 2018.

[8] Kosta Pavlović and Goran Šuković. "Deep Learning Techniques for Classification of Handrwitten Digits". In: *Informacione Tehnologije, Žabljak*. 2017.

HONOURS AND AWARDS

- Plaque of University of Montenegro in the field of technical and natural sciences (2016)
- Award for the best students, awarded by capital city (Podgorica) (2015)
- Annual student award for best students of the University of Montenegro (2014)
- Award at the national competition in computer science (2011)
- Award at the knowledge Olympiad, (2011)
- Member of Montenegrin team at one International Olympiad in Informatics held in Thailand in 2011.
- Member of Montenegrin team at three Balkan Olympiads in Informatics held in Montenegro (2010), in Romania (2011), in Serbia (2012)

Digital Signal Processing 122 (2022) 103381

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Robust speech watermarking by a jointly trained embedder and detector using a DNN

Kosta Pavlović [a,*], Slavko Kovačević [b], Igor Djurović [b,c,d], Adam Wojciechowski [c]

[a] University of Montenegro, Faculty of Natural Sciences and Mathematics, Podgorica, Montenegro
[b] University of Montenegro, Faculty of Electrical Engineering, Podgorica, Montenegro
[c] Lodz University of Technology, Institute of Information Technology, Lodz, Poland
[d] Montenegrin Academy of Sciences and Arts, Podgorica, Montenegro

ARTICLE INFO

Article history:
Available online 30 December 2021

Keywords:
Speech watermarking
Blind watermarking
Watermark attacks
Deep neural networks
Autoencoders

ABSTRACT

This paper utilizes deep neural networks for robust digital watermarking and authentication of speech signals. We present two adversarial neural networks, called embedder and detector. The embedder network tends to achieve imperceptible watermark embedding by minimizing the differences between the original and watermarked signals. The detector strives towards errorless watermark detection. We have proposed joint optimization of these two networks to achieve a trade-off between these requirements. Two models are trained, one with raw speech signals and another with STFT signal representations. Proposed techniques are tested on attacked audio recordings with measures such as SNR, PESQ, BER, and bps. The proposed system achieves high robustness to common watermarking attacks with BERs well below 1%. SNR values greater than 38 dB and PESQ values of 4.33 demonstrate that the difference in original and watermarked signals is negligible. In comparison with other approaches, our scheme shows good overall performance in terms of imperceptibility, robustness, and capacity.

1. Introduction

Digitalization of multimedia has brought many improvements to digital data storing, reusing, rewriting, and to the overall quality of data perseverance in general. From analogue videotapes and digital compact discs to hard drives and Internet clouds, data storage has grown and become much more accessible and affordable to the end consumer. As the accessibility of digital data increased, data editing computer programs became more available, less complex, and significantly more capable of altering original recordings. These tools are not and should not be labeled as a threat, as they are the backbone of the digital revolution, but their misuse can lead to severe damage not only to intellectual property. For example, in the case of forgery of speech recordings, it can bring lasting damage to the reputation of public personalities. The expansion of the Internet has made online data piracy an almost unstoppable trend, and there does not seem to be a significant movement to stop it. Torrenting sites and clients have made data easily exposed to the public. Authors and rights owners are unable to prevent malicious individuals from exploiting their intellectual property, which results in significant financial losses for the entertainment industry. In this paper, we try to preserve the intellectual property and integrity of speech recordings. Speech is chosen from all audio data, as it is a basic form of expression, and freedom of expression is not just a cornerstone of democracy but a fundamental human right enshrined in Article 19 of the Universal Declaration of Human Rights.[1]

Watermarking is a process of marking digital data with a watermark to preserve copyright and authenticity. A watermark that becomes embedded into a signal carrier can be a message of any kind, and it does not need to have any expressive meaning. Additionally, watermarking techniques should not decrease the overall quality of the information, as by doing so, we would substitute one problem for the other. Therefore, watermarking techniques should achieve a trade-off between data quality and watermark robustness where, ideally, the result should represent an inaudible and indelible watermark embedded into an original signal.

Digital watermarking emerged during the 1990s with a cornerstone publication [1] reviewing some of the most important spread spectrum watermarking techniques. Since then, a vast number of techniques have been developed for watermarking all types of digital multimedia [2,3], such as images [4–10] and videos [11]. Digital

* Corresponding author.
E-mail address: kosta@ucg.ac.me (K. Pavlović).

[1] United Nations General Assembly, General Assembly resolution 217 A.

https://doi.org/10.1016/j.dsp.2021.103381
1051-2004/

image watermarking remains the most researched part of the field, with many ideas being borrowed from it and applied elsewhere.

Audio watermarking also emerged more than 20 years ago. During that period, numerous techniques for embedding watermarks into audio signals were proposed. We adopted terminology and classification of audio watermarking techniques from [12], as it is the most recent and comprehensive examination of existing works. There are several criteria used to categorize audio watermarking methods, but the fundamental is the type of domain in which the watermark is embedded. Time domain has been used as the straightforward candidate in which the process of watermarking is being done with the majority of techniques classified as echo-based [13–15] or time-aligned [16–19] methods as per [12].

Transform domain methods have been used more extensively for audio watermarking. Most of these approaches have primarily been designed for other types of media, such as images, and then adapted for application to audio signals. The spread spectrum technique was introduced in [1] for general multimedia. This technique was further developed and used for audio watermarking in [20–24]. Quantization index modulation (QIM) is another class of transfer domain methods. This method consists of two main steps, index modulation and quantization, and was originally presented in [25] with a case study for image watermarking. The idea from this paper was later included in numerous audio watermarking techniques proposed in [26–31]. A well-known patchwork algorithm from [32] was proposed for audio watermarking in [33]. The modified patchwork algorithm [34] improves performance with respect to watermark embedding into audio signals in terms of robustness and imperceptibility. Many transform domain methods perform the aforementioned techniques on coefficients of orthogonal transformations, such as discrete Fourier transform (DFT) [35], discrete wavelet transform (DWT) [26,31], or discrete cosine transform (DCT) [24,29,36]. However, a significant number of approaches simultaneously use two or more of these transformations [27,28,37] or could be equally applied with any of these transformations [34]. For these reasons, it is impractical to divide the approaches by the type of the transform domain.

Several recent prominent audio watermarking techniques include [29,31,35,37]. The DCT domain watermarking technique was proposed in [29] with careful dispersion of the watermark energy to achieve a trade-off between watermark strength and inaudibility. A couple of DWT-based methods have been proposed in [31]. Bipolar synchronization codes and watermark bits are embedded into different DWT subbands in frames selected by intensity thresholding. DFT, fractional Fourier transform (FrFT) and quaternion DFT were developed and compared in [35]. No significant difference in performance was observed in imperceptibility or robustness related to the employed transformation domain. Watermarking with a combination of the DWT and DCT domains was proposed in [37]. The DWT is used to filter out the low-frequency signal components in which the watermark is embedded using the DCT.

Machine learning has also been used in the field of watermarking, primarily for watermark detection [38]. More significantly, a surge in deep neural network (DNN) popularity has led to an increase in the attention given to these techniques. Several recent research papers employ DNNs for both watermark embedding and watermark detection [39–43]. DNN-based techniques show promising results but currently lag classical digital watermarking approaches. We are still not observing a big leap forward as in other digital signal processing disciplines. In particular, DNN-based watermarking methods will be important in the future since DNN techniques are already used on the other side of the equation. In particular, the appearance of so-called "deep fakes" further facilitates the generation of fake or misleading information by using DNNs. These attacks are not used only to spread disinformation but also to damage the reputation and trust of a person or entities or to perform identity theft. As these implications can have severe consequences, it is important to counter them. Therefore, to address the issues raised by the application of DNNs to counterfeiting, it is necessary to develop novel watermarking and authentication techniques using the same means in the future. The ability of DNN generalization could eventually lead to a simpler design of the audio watermarking system as not every part of the system will have to be the result of human effort.

The technique proposed in this paper utilizes DNNs to perform watermark embedding and detection, and it is based on our previous work [44], which introduced a naive watermark embedding system that was not trained against any watermark attacks. This paper presents a robust speech watermarking system that is resilient to attacks, a significant step forward that required engineering advanced training procedures to achieve both imperceptible watermark embedding and errorless watermark detection. The goals of watermark embedding and detection are opposite and must be balanced. The watermark should be robust and resilient to attacks, but the quality of the signal needs to be preserved. Therefore, we propose a deep learning system that can be divided into two parts. The first part of the system, the embedder, has a task to embed the watermark into the original signal. Rather than using traditional domains, the embedder network generates a transformation domain. Latent space representation of the signal is input-dependent. Choosing an adequate input domain was also a subject of this research. Unlike in [44], where we used spectrogram, we concluded that the input domain must be reversible to the time domain as a prerequisite for preserving signal quality. Using an irreversible transform domain would render the system useless in practical applications. The short-time Fourier transform (STFT) and the time domain were compared as input domains. We selected the STFT as a signal representation in the time-frequency domain due to its invertibility property. This could not be achieved with other traditional signal representations commonly used with DNNs, such as spectrograms or Mel coefficients. Since DNNs are often thought to be able to perform feature extraction on their own, time-domain representation has also been considered. The second part of the system, the detector, has the task of detecting the watermark in the signal carrier. It behaves as a classifier, as there are a limited number of possible watermarking messages. These two neural networks are adversaries, as their goals are completely opposing. They must be balanced to fulfill the task of the watermarking system. Joint and carefully conducted training procedure enables the convergence of both networks. It can be expected that the message will be hidden well enough so that only the detector network can find it, and its presence should not influence the quality of the information. During the design, the possibility that a signal is attacked is considered and the influence of the percentage of expected attacks on the watermark detection process. Ultimately, we tested our technique on a set of speech signals against the techniques from [29,31,35,37] in terms of signal quality, system robustness, and capacity. The ability of the proposed technique to avoid reverse engineering attacks is an additional benefit concerning the increase in the overall complexity of the system. Embedder performance is not an issue, as its results are not required in real time. However, the detector is reduced to a simple classifier that does not require intensive calculations since it can be implemented on mobile phones and similar platforms.

The paper is organized as follows. Section 2 gives a formal definition of the main watermarking methods in the time and transformation domains, as well as a brief description of the STFT transform as a particular transform/domain used for watermarking in this paper. The proposed DNNs (the embedder and the detector) and the corresponding training procedure are covered in Section 3. The proposed technique is designed to be robust against

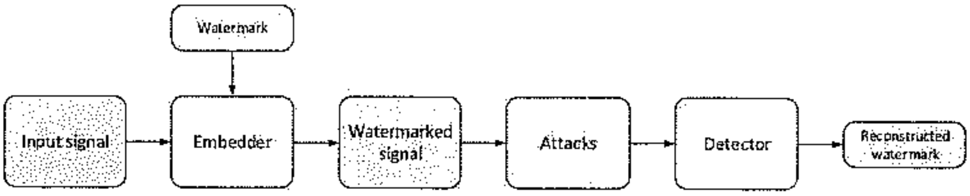


Fig. 1. Architecture.

the most common attacks on the watermarking system. As characteristic examples, we have considered noise embedding, filtering, and sample suppression. These attacks are given in Section 4. The utilized dataset is described in Section 5. Measures used to evaluate the performance of the proposed and comparative methods are presented in Section 6. The results of the study and simulations are given in Section 7. Concluding remarks and further potential research steps are given in Section 8.

2. Watermarking domains

The generic formula for audio watermarking in the time domain can be expressed as:

$$y(n) = x(n) + \alpha w(n) \tag{1}$$

where y represents a watermarked signal, while x and w are the original signal and watermark, respectively. Parameter α determines the strength of the watermark having an impact on the watermark audibility and detectability. Here, it is given as constant, but in almost all approaches, it varies and adapts to the signal contents so that the watermark can be less audible but easier to detect. Time-domain watermarking is commonly considered efficient for realization, while the robustness to attacks is often considered inadequate, leading to the development of various transform domain watermarking techniques. The generic formula for audio watermarking in the transform domain can be expressed as:

$$Y(k) = X(k) + \alpha W(k) \tag{2}$$

where Y represents the watermarked signal in a transform domain, while X and W are the original signal and watermark in the same domain, respectively. Transform domain can be DFT, DWT, FrFT, etc. However, in this research, we consider the STFT domain as a sort of generalization of the DFT technique. The STFT can be defined as:

$$STFT(n,k) = \sum_{m=-N/2}^{N/2-1} g(m)\, x(n+m)\, e^{-j2\pi mk/N} \tag{3}$$

where g is a window that localizes the signal in the time-frequency plane. N is the length of the window. The quality of the STFT depends on the window type and its length, but the STFT is a good representation of the signal if these parameters are well chosen. The squared magnitude of the STFT is defined as:

$$SPEC(n,k) = |STFT(n,k)|^2 \tag{4}$$

and is called a spectrogram. It can be defined as an energetic version of the STFT. The disadvantage of the spectrogram is that it is difficult to perform an inverse transformation on it and retrieve the original audio signal, while in the case of the STFT, its linearity makes this operation effortless. Mel spectrogram is a popular representation of audio signals in the field of neural networks. It is often used in audio classification and speech-to-text systems. It is irreversible like the spectrogram. The spectrogram and Mel spectrogram have advantages over the STFT when used in neural networks, as they are far less spatially complex. However, the considered problem requires invertible transformation.

3. Proposed DNN architecture and training

The architecture of the DNN presented in this paper is partially based on [39] and [44]. It can be divided into two parts: the embedder and the detector. The considered architecture is shown in Fig. 1. The proposed architecture is broad enough to cover two models called here A and B. They have different inputs, so their realizations differ at some points, but the backbone idea is mostly the same. Model A uses STFT as an input, while Model B uses raw audio samples. Both models are well argumented. While DNNs have long been championed by their success in dealing with raw data, STFT is a proven tool in digital audio processing. Feeding DNNs with raw data can reduce dependence on preprocessing steps but, in turn, forces the network to learn feature extraction on the data presented in the training dataset. This can slow the process of learning and even lead to worse results. However, by using common transformation domains, the network loses the chance of potentially discovering features that could be lost during the extraction. This is more pronounced in irreversible transformations, such as standard or Mel spectrograms. Model A has one obvious advantage. Model B simultaneously learns feature extraction and watermark embedding, whilst in the case of Model A, feature extraction is partially done by STFT. Embedder and detector networks are described in Sections 3.1 and 3.2. The DNN is trained against the attacks described in Section 4. We assume that these attacks are an integral part of the DNN. They are not learnable layers, but their position in the architecture design makes them part of the DNN as the gradient passes through them from the detector to the embedder. Since they introduce constant gradient flow and their parameters are not learnable, we do not observe them as an independent part of the network. As these two networks have opposing tasks, each of them has its own loss function. However, gradients produced by the detector loss function are being backpropagated towards the embedder. Therefore, these two networks are dependent on each other.

3.1. Embedder

The embedder network architecture is based on the U-net design concept [45], structurally similar to autoencoder networks. Autoencoder networks can be summarized as in Fig. 2. This network transforms input x into some latent space representation r, followed by expansion to $\hat{x}$. Ideally, the autoencoder should satisfy $x = \hat{x}$.

Autoencoders are mostly used in feature extraction and dimensionality reduction. In our case, we use an autoencoder to learn a representation that is unique for our dataset in which the watermark can be hidden. In other words, a watermark is added to r, but the overall goal of the autoencoder remains the same. It should learn to ignore the watermark and eliminate any difference between x and $\hat{x}$. However, leaving the possibility for watermark detection is still the goal of the entire system. This indicates that the embedder cannot be trained independently but jointly with the detector. The encoding space is dependent on the dataset but also the training process and, more precisely, initialization. Watermark represents a one-dimensional binary message that has a fixed size.

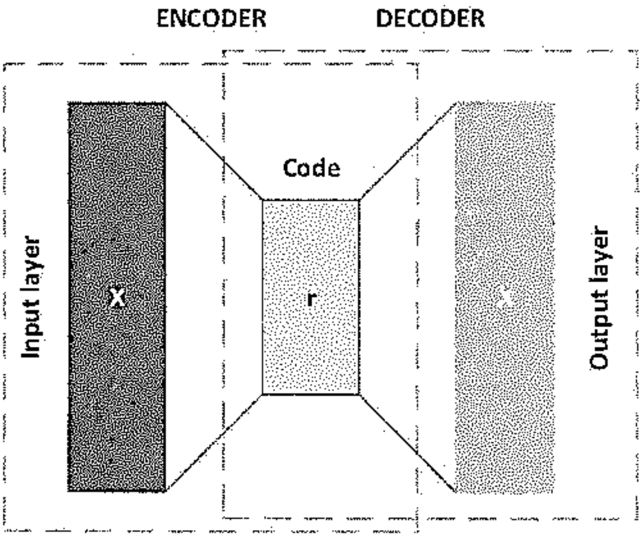


Fig. 2. Autoencoder.

This message could also contain meaningful information, but we have chosen to generate a pool of random messages. As data are shuffled during the training, getting to the same autoencoder parameters, i.e., layer weights, is virtually impossible. To analyze the impact of the input used, we observed two models. Embedder architectural parameters for Model A and Model B are presented in Table 1. The loss function used in both models is the mean absolute error calculated between the output of the embedder and the original signal.

3.1.1. Model A

Model A architecture is shown in Fig. 3. The input of Model A is an STFT representation of the original signal. A Hanning window of 1024 samples is used, giving the STFT with the same number of frequency bins. The STFT is applied with partially overlapped windows where time instants are separated for 512 samples. The resulting matrix has a size of $512 \times 64 \times 2$, where 512 is the height, 64 is the width, and 2 represents the number of channels, i.e., real and imaginary parts. The downsampling part of the embedder is made of 5 blocks. The reduction of the spatial dimension and increase in the number of filters is a feature of the U-net architecture. After the convolution, the signal passes through batch normalization layer [46] and leaky ReLU [47] activation function whose α parameter is set to 0.2. Batch normalization is used to normalize the data and is helpful in the overall stability of the training process at the expense of slowing it. The resulting latent space representation of the signal has a size of $16 \times 2 \times 256$. Here, in the latent representation of the signal, watermarking message, which is a 1-dimensional array, is embedded by concatenating it at the end of the channel dimension. To increase the chances of preserving the watermark, it is repeated 16×2 times before embedding. Latent space representation of the signal is a 3D tensor, a rectangular cuboid, and it consists of 16×2 arrays of length 256. To further expand this shape by concatenation, we would require another tensor to match at least two of its dimensions. This newly formed tensor infringes on the symmetry of the U-net design pattern, which affects the second part of the embedder that performs upsampling and reconstruction. We solve this problem by introducing another 2D convolutional layer. The upsampling part of the embedder also consists of 5 blocks. After transposed convolution, the signal passes through batch normalization layer, dropout layer with a dropout probability of 50% and ReLU activation [48]. As in the downsampling part of the embedder, batch normalization normalizes the data, while dropout is an attempt of reaching robustness. By dropping half of each layer output, the network is forced to adapt to attacks. The main problem of using batch normalization and dropout layers together is described in [49]. Here, robustness is as important as reconstruction, so using these two layers together is appropriate. The last two blocks of the upsampling part of the embedder do not have dropout layers. Outputs of the corresponding downsampling blocks are concatenated after the corresponding transposed convolutional layers in the upsampling blocks. This helps the embedder reconstruct the original signal. The last transposed convolutional layer has 8 filters, while the STFT input has only 2 channels. Equalizing embedded input and output dimensions is achieved with another 2D convolutional layer. After leaving the embedder, the reconstructed STFT passes through a series of attacks. In Section 4, attacks are presented in the time domain. The reconstructed STFT signal is transformed into its time-domain pair using inverse STFT, attacked, and then again transformed back to the STFT.

3.1.2. Model B

The input of Model B are time-domain representations of audio signals. As previously stated, this approach is more common for DNNs, as it avoids the data preprocessing (transform evaluation) stage. The general design pattern of Model B is similar to that of Model A. 1D convolution is used due to signal dimensions. As there are many samples available, we used a large kernel with a significant stride in the first block to reduce dimensionality. The signal around point t in time can be declared constant if the time offset is sufficiently small. For a sample rate of 16000, there are 16 samples for each millisecond. It can be expected that the signal will not sufficiently change in just over two milliseconds. Therefore, using a kernel size of 41 is justified. The kernel size decreases as the spatial dimension does. Otherwise, the network could omit important features. The number of filters in convolutional layers is the same as in Model A. However, there is no batch normalization here. Instead, a convolutional layer is followed by a scaled exponential linear unit (SELU), which is an activation function that induces normalizing properties [50]. Normal LeCun initialization must be used [50,51] for SELU to work properly. This activation function is defined as:

$$f(\alpha, x) = \lambda \begin{cases} \alpha\left(e^{x} - 1\right) & x < 0 \\ x & x \geq 0, \end{cases} \tag{5}$$

where λ is a constant scaling parameter while α can vary. However, we have adopted here 1.0507 for λ and 1.67326324 for α as in [50]. The resulting latent space has a size of 128×256, which is still a large number of channels despite a drastic reduction in spatial dimensions. Therefore, there is an exact number of samples in the raw signal and its latent space representation, which differs from Model A, where one latent space sample represents 4 samples of the STFT input. Watermark becomes embedded as in Model A by concatenating it at the end of the channel dimension while repeating it 128 times to match the spatial dimension. Symmetry infringement occurs again, so another convolutional layer is introduced, this time with 256 filters, stride 1, and kernel of size 9. The activation function is SELU, and the layer is initialized with LeCun initialization, like all other downsampling layers.

The upsampling part of the embedder has 5 blocks. In the first three blocks, after transposed convolution with SELU activation and normal LeCun initialization, the signal passes through the alpha dropout layer with a probability of 40%. Alpha dropout is introduced along with the SELU activation function in [50], as it maintains the self-normalizing property. Therefore, there is no need for batch normalization.

The last two blocks of the upsampling part of the embedder are not using alpha dropout. Outputs of the corresponding downsam-

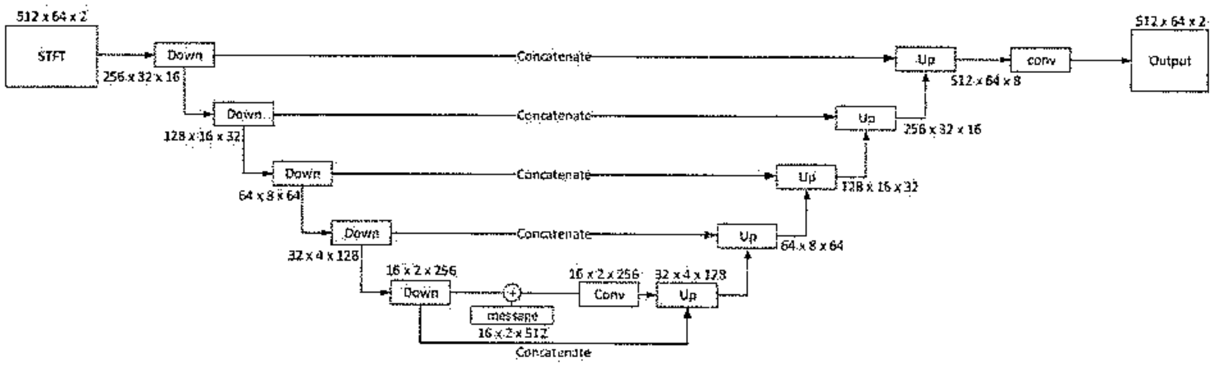


Fig. 3. Embedder architecture of Model A. Downsampling blocks are convolutional layers that decrease signal dimensions. Upsampling blocks are transposed convolutional layers that increase signal dimensions.

Table 1
Embedder architectural parameters for Model A and Model B. Additionally, Model A utilizes batch normalization layers, whilst Model B evades this by using the SELU activation function which induces normalizing properties.

Type	Model A Filters	Size/Stride	Output	Model B Filters	Size/Stride	Output
Convolutional	16	5×5 / 2	256×32	16	41 / 8	4096
Convolutional	32	5×5 / 2	128×16	32	21 / 4	1024
Convolutional	64	5×5 / 2	64×8	64	21 / 2	512
Convolutional	128	5×5 / 2	32×4	128	21 / 2	256
Convolutional	256	5×5 / 2	16×2	256	21 / 2	128
	-watermark embedding-					
Convolutional	256	5×5 / 2	16×2	256	21 / 2	128
Transposed Convolutional	128	5×5 / 2	32×4	128	21 / 2	256
Transposed Convolutional	64	5×5 / 2	64×8	64	21 / 2	512
Transposed Convolutional	32	5×5 / 2	128×16	32	21 / 2	1024
Transposed Convolutional	16	5×5 / 2	256×32	16	21 / 4	4096
Transposed Convolutional	8	5×5 / 2	512×64	8	21 / 8	32768
Convolutional	2	5×5 / 1	512×64	1	41 / 1	32768

pling blocks are concatenated after the corresponding transposed convolutional layers in the upsampling blocks, as in Model A. The dimensionality of input and output do not match, as in Model A. This problem is easily solved by another 1D convolutional layer. The second problem lies in the input signal, which is bounded. Additionally, audio signals can have both negative and positive values. Therefore, a hyperbolic tangent activation function is used with Xavier or Glorot initialization [53]. As in Model A, embedder output passes through a series of attacks which are presented in the time domain. Unlike in Model A, the output of the Model B embedder is already in the time domain, so no inversion has to be performed to apply the attacks. This is an obvious advantage for Model B over Model A.

3.2. Detector

An overview of the detector architecture is shown in Fig. 4. Detector architectural parameters for Model A and Model B are presented in Table 2. The output of the embedder, a reconstructed version of the signal, is the input of the detector. The detector architecture is similar to an image classifier. Strictly speaking, the detector is not a classifier, but from the DNN point of view, it can be considered in that role. The goal of the detector is to accurately detect every bit of the watermark, i.e., classify it as 0 or 1. The loss function used in both models is binary cross-entropy calculated between the output of the detector and the original watermarking message.

3.2.1. Model A

Model A detector consists of 6 convolutional downsampling blocks. These blocks are similar to those from the embedder network. After the convolution, the signal passes through batch normalization. Leaky ReLU is used as an activation function, and the α parameter is the same as in the embedder, 0.2. Strides are set so that the output of the detector has dimensions $L_w \times 1$, where L_w is the length of the watermark. To reshape the output, a fully connected layer is used after flattening the output of the last block.

3.2.2. Model B

Model B detector is slightly different from Model A for the reason that the SELU activation function is used with normalized LeCun initialization for the convolutional layers, as in the Model B variation of the embedder.

3.3. Training procedure

The training procedure of this system must be carefully conducted because the embedder and the detector have opposite tasks. One network could prevail and prevent the other from learning. The set goal of the embedder network is the reconstruction of the original signal after appending the watermark. In this process, the embedder would certainly try to obliterate the watermark because that would best fulfill his goal. By restraining the embedder during the training, we prevent the watermark from being erased. The detector network expects the message to be present in the reconstructed signal. If the reconstructed signal does not carry a message, then there is no way for the detector to learn anything out of it. On the other hand, by restraining the embedder too much, the detector network can take the upper hand and thus avert the embedder from performing high-quality reconstruction, a task equally as important as watermark detection.

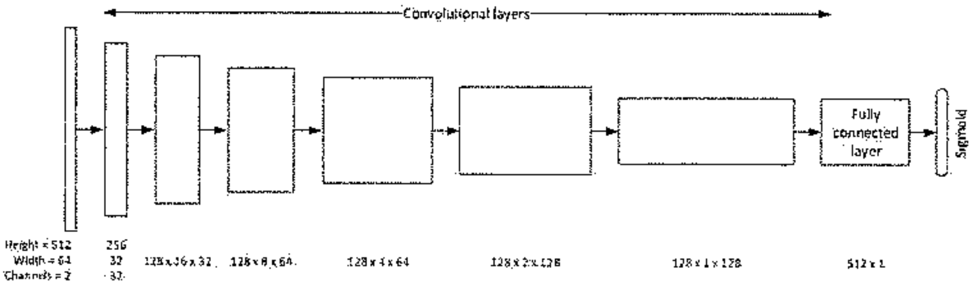


Fig. 4. Model A detector architecture. The first block represents the input signal.

Table 2
Detector architectural parameters for Model A and Model B.

Type	Model A			Model B		
	Filters	Size/Stride	Output	Filters	Size/Stride	Output
Convolutional	32	5 × 5 / 2	256 × 32	32	41 / 8	4096
Convolutional	32	5 × 5 / 2	128 × 16	32	21 / 4	1024
Convolutional	64	5 × 5 / 1 × 2	128 × 8	64	21 / 2	512
Convolutional	64	5 × 5 / 1 × 2	128 × 4	64	11 / 2	256
Convolutional	128	5 × 5 / 1 × 2	128 × 2	128	11 / 1	128
Convolutional	128	5 × 5 / 1 × 2	128 × 1	128	9 / 1	64
Fully Connected		16384 × 512	512		8192 × 128	128

Table 3
Parameter values for the two versions of Model A and Model B.

Parameter name	Value	
	Model A-V1 and Model B	Model A-V2
Υ_{w_e}	1	1
Φ_{w_e}	$\emptyset$	3.5
Δ_{w_e}	0.2	0.2
Υ_{w_d}	2	2
Φ_{w_d}	$\emptyset$	0.5
Δ_{w_d}	0.1	0.1
K	10	8
M	10	10

For these reasons, the weights are introduced in the learning process to balance the described requirements and ensure the convergence of both networks. These weights vary in epochs of learning as a function of the epoch index:

$$w_e(i) = \begin{cases} \Upsilon_{w_e} & i \le 1 \\ \Upsilon_{w_e} + (i-1)\cdot\Delta_{w_e} & 1 < i \le K \\ \Phi_{w_e} & K < i \le M, \end{cases} \tag{6}$$

$$w_d(i) = \begin{cases} \Upsilon_{w_d} & i \le 1 \\ \Upsilon_{w_d} - (i-1)\cdot\Delta_{w_d} & 1 < i \le K \\ \Phi_{w_d} & K < i \le M \end{cases} \tag{7}$$

where $w_e(i)$ and $w_d(i)$ are the respective weights of the embedder and detector networks as a function of epoch index i and Υ_{w_e}, Φ_{w_e}, Δ_{w_e}, Υ_{w_d}, Φ_{w_d}, Δ_{w_d}, K, and M are design parameters. Υ and Φ denote the initial and final values of the weights, while Δ is the step size. Parameter K is an arbitrary number of epochs, whilst M is the total number of epochs. Here, we considered two sets of parameter values for Model A training. This resulted in two versions of that model, called Model A-V1 and Model A-V2. Models A-V1 and B were trained with identical parameter values. In the case of Model A-V2, a slight advantage is given to the embedder by increasing its loss weight for the last two epochs. This led to better signal quality results with insignificant deterioration in robustness, which will be shown in Section 7. The assignment of parameter values is given in Table 3.

These weight rules are empirically determined following simple reasoning. In the beginning, the embedder dominates over the detector, so we assign larger weights to the detector network. The initial weight values according to our experiments are of great importance. Excessively high values can introduce instability and lead to divergence. We then gradually decrease the weight of the detector while increasing the weight of the embedder to ensure an acceptable level of disturbance in the output signal of our system.

Part of the detector input samples must be intentionally attacked during training for the system to be able to learn to overcome the attacks in Section 4. It is necessary to determine the optimal size of this portion of the samples to ensure the best performance of the model. In every attack layer, we introduced an additional hyperparameter to represent the probability that a given attack will occur. We chose three values (specifically 15%, 25%, and 35%) for this hyperparameter and compared the performance of our system depending on those values. Fig. 5 shows trends of performance evaluation measures for Model A-V2 and three different percentages of attacks. Model A-V1 and Model B exhibited similar behavior in this regard. Performance evaluation measures will be described in Section 6. As shown in Fig. 5a, the detector converges after the first couple of epochs. In subsequent epochs, only the embedder network effectively learns, while the detector maintains its accuracy. Additionally, Fig. 5a shows that the model has significantly worse watermark detection accuracy when it is trained with 35% of attacked samples. Since the remaining two models had similar performance in watermark detection, we chose the model with 25% attack probability because it performs slightly better in terms of preserving signal quality. It could seem confusing that the model with the highest percentage of attacked samples gave the best results in signal reconstruction, but this is completely expected. Both networks are trying to converge, but the best result for the system is that they balance out. Thus, when both networks have good results, it is fair to estimate that these results are not the best that they can achieve individually. When the detector network fails to detect the watermark, the training process becomes unbalanced, and the embedder converges completely. The training process lasted approximately 4 hours for each of the model versions.

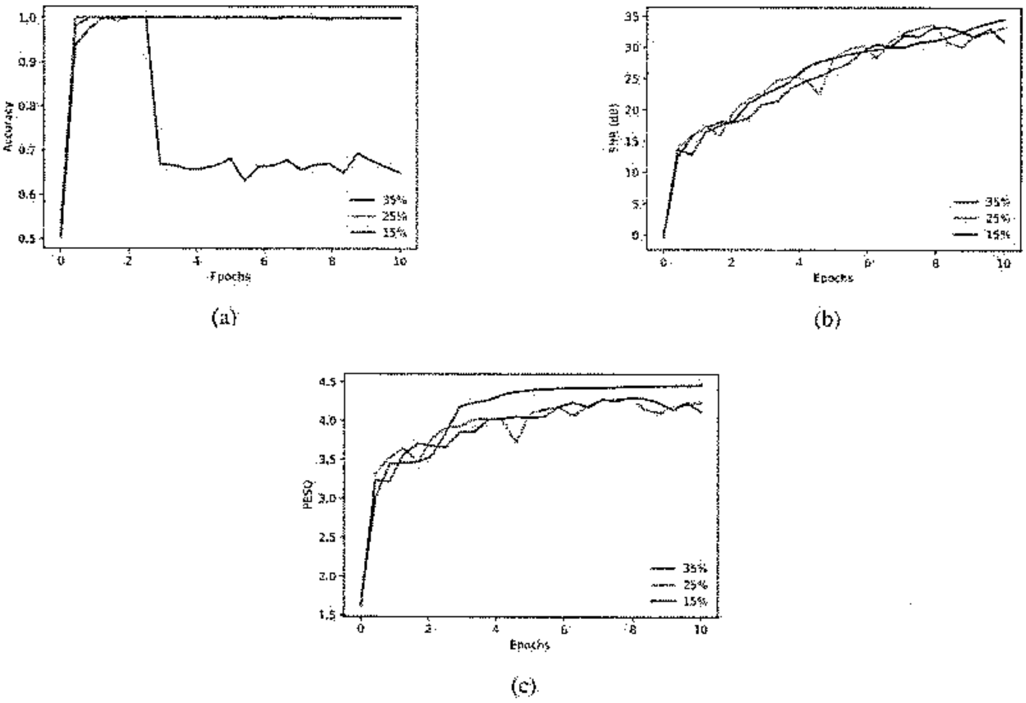


Fig. 5. Trajectories of performance measures for Model A-V2 for different probabilities of attacks: (a) Watermark detection accuracy; (b) SNR of the watermarked signal; and (c) PESQ of the watermarked signal.

4. Audio watermarking attacks

An audio watermarking system is often required to be able to detect a watermark when the watermarked signal has been distorted or damaged, either due to unprofessional or malicious signal handling or even sometimes due to completely legitimate operations performed on signals. All these actions, whether they are malicious or not, are considered attacks from the point of view of the audio watermarking system. If the audio watermarking system fulfills this requirement, it is labeled robust. Robustness presents the greatest challenge for watermarking systems due to the abundance of attacks and their sophistication.

A significant number of audio watermarking attacks have been proposed thus far. Review papers on audio watermarking attacks are [53], [54], [55], but no document covers all of them. In reference [53], the authors propose a benchmark set of attacks to be used to test audio watermarking systems. Additionally, designing new attacks is an active area of research. Attacks are constantly in development with new technologies such as DNNs and other emerging machine learning techniques, giving fertile ground for further development in this field over time. In these circumstances, designers of audio watermarking systems need to select a limited set of attacks suitable for the intended application of their system and try to make the system resistant at least to those attacks. The authors in [31] tested the proposed algorithm on numerous attacks. Unfortunately, there will certainly be attacks to which the designed watermarking system is not resistant.

The fight against the attacks cannot be left to the detector alone because if the embedder does not embed the watermark properly, it can easily be destroyed by some of these attacks, and the detector cannot do anything about it. For example, if the embedder inserts watermark bits exclusively into high-frequency bands, a low-pass filter would completely erase it. Therefore, robustness is a task for both the embedder and the detector, so they must be jointly trained. The error in watermark detection caused by the attacks must be propagated to the embedder. This means that watermarking attacks must be implemented as neural network layers.

In this paper, we have trained our system against three types of audio watermarking attacks, i.e., Gaussian additive noise, Butterworth filter, and sample suppression. All the attacks are parametrized, and the values of the parameters are set so that the attacks do not affect the signal perceptual quality. We believe that attacks lose their meaning if it is obvious that the signal has been altered by certain audio processing operations.

4.1. Random noise

The most common attack on audio watermarking systems is adding noise to the watermarked signal. This attack could also cover all kinds of attacks caused by compression losses. Adding noise attempts to destroy the information added during the embedding procedure. We selected a noise model from Stirmark Benchmark [53]: $y(n) = x(n) + \alpha \cdot \epsilon(n)$, where x, ϵ and y are the original signal, noise, and output signal, respectively, and α is a parameter which defines the relative strength of the noise to the original signal. The selected value for α is 0.009 so that adding noise to the original signal would not significantly degrade signal quality. The resulting signal-to-noise ratio with these settings is 30 dB on average.

Noise strength and distribution parameters are predefined and nontrainable. This means that only the gradient with respect to the input signal x needs to be propagated to the preceding layers during network training. Partial derivatives of y with respect to x are given with:

$$\frac{\partial y(n)}{\partial x(m)} = \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases} \tag{8}$$

By the chain rule, the total portion of the error (loss) E to be propagated to the input $x(m)$ is calculated as:

$$\frac{\partial E}{\partial x(m)} = \sum_{n=0}^{L_s} \frac{\partial E}{\partial y(n)} \cdot \frac{\partial y(n)}{\partial x(m)} = \frac{\partial E}{\partial y(m)} \tag{9}$$

where L_s is the length of the signal.

4.2. Low-pass filter

Butterworth filters are used as an example of low-pass filters, which are also a common type or part of attacks on watermark systems because they can erase the watermark from the signal or disrupt its detection. When embedding a watermark into audio or speech signals, high imperceptibility can be achieved by altering high-frequency components. The human auditory system is less sensitive to these frequencies, so the changes in them will not be audible. However, the watermark embedded with such a technique would be completely erased by passing the signal through a low-pass filter. This is why it is important to design an audio watermarking system resilient to low-pass filtering. We achieve this by adding a layer to our model to simulate this attack.

A Butterworth filter is a type of digital filter designed to minimize the ripple in the frequency response of the filter. It is also called a maximally flat filter because it has the narrowest possible roll-off for a given order without causing ripples in other parts of the frequency response. The amplitude response of the Butterworth filter is given by the following equation:

$$|H(j\omega)| = \frac{1}{\sqrt{1+\left(\frac{\omega}{\omega_c}\right)^{2N}}} \tag{10}$$

where ω is the angular frequency, ω_c and N are the two defining parameters of this filter, ω_c is the cut-off frequency, and N is the filter order. The amplitude response of the 16th-order Butterworth filter with a cut-off frequency of 4 kHz is given in Fig. 6a.

Poles of the Butterworth filter transfer function in the s-domain $H(s)$ are located on the circle of radius ω_c and are given with:

$$\begin{aligned} s_k &= \omega_c \cdot e^{\frac{j\pi(2k-1+N)}{2N}} \\ &= \omega_c \cdot \left[\cos\left(\frac{\pi(2k-1+N)}{2N}\right) + j\sin\left(\frac{\pi(2k-1+N)}{2N}\right)\right] \end{aligned} \tag{11}$$

for $k \in \{1, 2, \ldots, N\}$. After determining the poles, we define the transfer function of the Butterworth filter as follows:

$$H(s) = \frac{1}{(s-s_1)\cdot(s-s_2)\cdots(s-s_N)} = \sum_{k=1}^{N} \frac{r_k}{s-s_k} \tag{12}$$

where r_k are coefficients obtained by partial decomposition of the transfer function $H(s)$. To apply this filter in the time domain, we need to calculate its impulse response. Fig. 6b shows impulse responses of Butterworth filters of different orders. The impulse response of the Butterworth filter is given by (13):

$$h(t) = \sum_{k=1}^{N} r_k \cdot e^{s_k t} \cdot u(t). \tag{13}$$

where s_k are poles from (11), $u(t)$ is the unit step function, and r_k are coefficients from (12).

After calculating the impulse response, the filter can be applied by convoluting it with the signal.

$$y(t) = (x * h)(t). \tag{14}$$

It follows that the filter can be implemented as a convolutional layer with fixed and nontrainable weights calculated with Equation (13).

In the discrete-time domain, convolution from (14) has the following form:

$$y(n) = \sum_{m=-\infty}^{+\infty} x(m) \cdot h(n-m). \tag{15}$$

Partial derivatives with respect to the input signal x needed to propagate the error through this layer are calculated as:

$$\frac{\partial y(n)}{\partial x(m)} = h(n-m) \tag{16}$$

and the derivative of the loss function E with respect to the input signal x is:

$$\frac{\partial E}{\partial x(m)} = \sum_{n=0}^{L} \frac{\partial E}{\partial y(n)} \cdot \frac{\partial y(n)}{\partial x(m)} = \sum_{n=0}^{L} \frac{\partial E}{\partial y(n)} \cdot h(n-m). \tag{17}$$

We used a Butterworth filter with a cut-off frequency at 4 kHz since a lower value can cause audible alterations for speech signals.

4.3. Sample suppression

The sample suppression attack sets a random set of input signal samples to zero. Although this attack has a fairly simple model, it proves to be quite problematic for existing watermarking techniques. It can be described by the following equation:

$$y(n) = \text{mask}(n) \cdot x(n) \tag{18}$$

where the mask is a randomly generated vector of 0 and 1 that determines which samples will be suppressed. Partial derivatives for the input x are:

$$\frac{\partial y(n)}{\partial x(m)} = \begin{cases} \text{mask}(n) & n = m \\ 0 & n \neq m. \end{cases} \tag{19}$$

Applying the chain rule gives us:

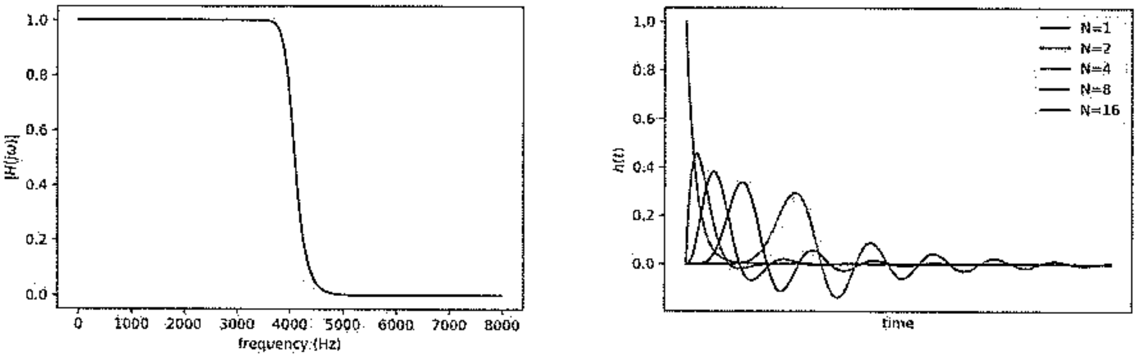
$$\frac{\partial E}{\partial x(m)} = \sum_{n=0}^{L} \frac{\partial E}{\partial y(n)} \cdot \frac{\partial y(n)}{\partial x(m)} = \frac{\partial E}{\partial y(m)} \cdot \text{mask}(m). \tag{20}$$

To preserve the quality of the signal and at the same time make the attack effective, 1000 samples are randomly suppressed, which makes up approximately 3% of the input signal.

5. Dataset

A standardized and widely adopted evaluation dataset for speech and audio watermarking does not exist. Hence, objective comparison of different watermarking algorithms is difficult. Various audio files are used in the literature. For example, SQAM files [56] are used in [37], while a small portion of the TIMIT database [57] is used in [31].

One particular problem that the audio watermarking system tries to solve is to protect statements of politicians and other public figures from possible changes. There is also a need to identify and authenticate the creators of audio content. Therefore, we decided to use a realistic dataset that contains the speeches of politicians in the Parliament of Montenegro. It can be used in the future to benchmark similar watermarking and authentication applications. It is available upon request with the permission of the owner (Parliament of Montenegro). We hope that it will soon be publicly available on a cloud service for scientific purposes.



(a) (b)

Fig. 6. Butterworth filters (a) Amplitude response of the 16th-order Butterworth filter with cut-off frequency at $f_c = 4$ kHz ($\omega_c = 2 \cdot \pi \cdot f_c$), (b) impulse responses of Butterworth filters of different order.

The dataset contains 6199 10-minute long audio recordings from the parliamentary sessions between 2016 and 2019. There are 280 different speakers, with speech times ranging from several seconds to 3-4 hours. All subjects were adults, aged from 22 to 73. They represent various ethnic and regional groups in Montenegro with different voice tonalities, rhythms, and dynamics of speech, and they take pauses of widely varying durations. The signal quality is generally good, with a sampling frequency of 44.1 kHz. Some of the session recordings are disturbed with ambient noise, grumbling, and other sounds common for heated parliamentary debates.

While the audible frequency for humans ranges from 20 Hz to 20 kHz, the voice frequency reaches a maximum of 8 kHz for some people [58]. Mostly, this frequency is anywhere from 1 to 5 kHz [59]. Narrowband telephone infrastructure limits the audio signal to just 3.4 kHz, which is enough to preserve information. For this reason, we can downsample speech signals from 44.1 kHz and still retain high signal quality. By the Nyquist-Shannon sampling theorem, we must choose a sampling frequency at least two times higher than the maximum signal frequency, so we downsampled our audio signals to 16 kHz ($2 \cdot 8$ kHz). Downsampling can have drawbacks by reducing quality, but it drastically increases the performance of neural networks, as more lengthy sequences require more hardware resources and much more time. In addition, we removed silence intervals longer than 1 s with an experimentally determined threshold of −35 dBFS. This reduced the size of our dataset from 1033 hours to approximately 868 hours.

The remaining speech was divided into segments of 32768 samples (approximately 2 s). There are several reasons we have chosen to divide our speeches into segments of this length. The first reason is the length of the watermark. If the watermark is too short, the embedder will eventually learn to erase it while trying to perfectly reconstruct the original signal. Furthermore, the length of the watermark should not be of the same order of magnitude as the length of the segment, as this would lead to a significant amount of additive noise, which would reduce the overall quality. Embedder network would not be able to converge. Additionally, we wanted our watermark length to be comparable with the length of the STFT window that we used to create the dataset. If the lengths are not the same, they should at least exhibit the same order of magnitude, which are the best practices of the signal processing community. The last constraint we considered is the segment length in existing audio watermarking schemes. We wanted to have a comparable bit rate to create similar conditions in which a comparison between our and other results would be fair.

6. Performance evaluation measures

The performance of audio watermarking systems can be assessed according to several measures. The two main tasks of audio watermarking systems are, as mentioned before, preserving signal quality after embedding the watermark and achieving resistance to the attacks when detecting the watermark. Over the years, various measures have been designed to assess how well the system has responded to these tasks. Designing novel measures is also an active area of research [60]. Experimentally achieved values for our and comparative methods regarding each of the measures described in this section will be presented in Section 7.

6.1. Imperceptibility

Imperceptibility represents the ability of the watermarking system to embed the watermark so that it is indistinguishable from the carrier signal. Objective signal quality measures [61] are used to assess the quality of the proposed system in terms of imperceptibility. These measures assess the quality of the processed signal by its numerical comparison with the original signal or in certain cases even without using the original recording.

In this paper, we have used the signal-to-noise ratio (SNR) as it is the most used objective measure for audio signal quality. Additionally, as this paper deals exclusively with speech signals, we decided to evaluate a special measure for this type of signal. Perceptual evaluation of speech quality (PESQ) [62] is an objective quality measure that is considered to be the most suitable to assess different types of distortions in speech signals and is recommended by the ITU Telecommunication Standardization Sector (ITU-T). Values for the PESQ measure range from −0.5 to 4.5. PESQ gives a better assessment of speech quality, i.e., imperceptibility, than SNR, which does not consider the specifics of the human auditory system. However, we used both PESQ and SNR measures for comparison with other approaches in Table 4.

6.2. Robustness

Robustness is a fundamental requirement for audio watermarking systems. Especially in the case of speech watermarking, it becomes even more important in relation to imperceptibility because, with speech signals, it is much more important to preserve the information they carry than the quality itself.

To assess the robustness of the system, it is necessary to test the possibility of watermark detection when the signal is degraded by attacks. The bit error rate (BER) is a generally accepted measure of robustness among all watermarking approaches. It represents the portion of incorrectly detected watermark bits.

Table 4
Comparison of the proposed DNNs with the state-of-the-art techniques in terms of imperceptibility. Results in brackets are declared in references on an alternative dataset.

Watermarking scheme	PESQ	SNR
Model A-V1	4.21	35.19
Model A-V2	4.33	38.75
Model B	4.09	25.02
DCT-b1 [29]	4.45	27.59 (20.38)
DWT-IAMM [31]	-	(21.35)
QDFT-V1 [35]	-	(37.95)
[37]	3.17	11.04 (30.18)

6.3. Capacity

The watermark capacity or embedding rate is the third measure considered within this study. Although it is regarded as less important than imperceptibility and robustness, capacity is always measured as it describes the system's ability to embed as many watermark bits as possible in a unit of time. Capacity values are given in bits per second (bps).

6.4. Space and runtime

Time and spatial complexity are important aspects of every computational system. The same goes for audio watermarking systems. Besides time complexity, which is a theoretical measure, for practical considerations, the runtime is more important. Runtime is determined experimentally by measuring the time it took the system to perform the necessary operations. The spatial complexity of a DNN-based system is given in the total number of neural network parameters.

The DNN-based system proposed in this paper is a set of layers that are applied sequentially. From a strictly theoretical standpoint, it could be argued that the time complexity of our system is $O(1)$ since all design parameters have constant values. If we let parameters such as input size and the number of layers be variable, the time complexity of such a system would be equal to the time complexity of its most complex layer times the number of layers. The most critical operations in our systems are STFT and 1D and 2D convolution. The time complexity of the STFT is $O(T \cdot N\log N)$, where T is the number of segments and N is the number of FFT coefficients. Convolution could be calculated as a product in Fourier space, which means that its time complexity could be reduced to the time complexity of the FFT. However, in our case, convolution is performed directly, resulting in time complexity of $O(D_1 \cdot k_1)$ for 1D convolution and $O(D_{21} \cdot D_{22} \cdot k_2^2)$ for 2D convolution, where D_1, D_{21} and D_{22} are input dimensions, and k_1 and k_2 are kernel sizes.

7. Experimental results

This section presents the results according to the four previously described performance evaluation criteria. The proposed DNNs are compared with some of the best approaches in the field in terms of overall achievements in high imperceptibility, robustness and capacity. For comparison purposes, we have realized techniques described in [29] and [37]. These techniques, which are currently not available in the public domain, are dependent on design parameters that affect the results. We adjusted those parameters to the best of our abilities. Our realizations are available at https://github.com/kosta-pmf/audio-watermarking. In addition to our realizations and dataset, we have compared the proposed DNN techniques with the results declared in references [29,31,35,37] with other datasets. These results are given in brackets in Tables 4 and 5.

To provide the fairest possible comparison between all methods, we tested their imperceptibility and robustness for a fixed watermark size of 512 bits, with the exception of Model B. The results for Model B are laid out for a capacity of 64 bps, i.e., 128-bit watermarks. Obtained results in terms of imperceptibility are given in Table 4. Table 5 shows the results of watermark detection accuracy in different attack scenarios. Embedding capacities of different approaches are listed in Table 6. All these experiments and comparisons are performed on the part of the dataset that was not used for training.

Both versions of Model A have achieved excellent results in terms of imperceptibility, with PESQ values well above 4.0. Model B also reached a PESQ value above 4.0, suggesting that the watermark is almost completely inaudible when embedded into the carrier signal. Fig. 7 visually confirms a small difference between the original and watermarked signal. This difference between the two signals can be considered a watermark estimation in the time domain. A detailed view of these signals can be seen in Fig. 8. In Fig. 9, we present the STFT magnitude of the watermark (i.e., watermark estimation) to examine in which areas of the spectrum the model embedded the watermark. STFT magnitude in Fig. 9 has a clear shape. The watermark is present in those time intervals in which the actual speech occurs, while the watermark is missing when the speaker pauses. The watermark is also present in the entire spectrum, which means it would be difficult to erase it. It could be argued that this difference is nothing else than the reconstruction noise or, more precisely, reconstruction error, but we think that this is only partially true.

Model A-V2 reached values higher than all comparative approaches in terms of SNR, with method [29] slightly outperforming it in terms of PESQ. The reason for somewhat higher values of SNR for our models could lie in the fact that the embedder network has been trained with mean absolute error as a loss function that is more similar to SNR than PESQ. This discrepancy in values of SNR and PESQ also supports the previously stated claim that SNR is not an appropriate measure of watermark imperceptibility. The SNR treats all parts of the frequency range equally, although they do not affect audibility accordingly.

We tested our models against attacks described in Section 4. The attacks we considered are particularly important because they are representatives of the majority of existing classes of attacks. Both models have bit error rates of less than 1%, which represents high resistance to set attacks. Model A-V1 exhibits barely any incorrectly detected bits and performs at the level above the best techniques in this regard. Apart from considering every attack separately, we also observed a combination of all attacks, a scenario that has not been considered in the available literature. Most realistic attack scenarios would use a set of attacks rather than one specific attack, which is why it is important to evaluate the performance of watermarking systems in such cases. Our experiment confirms this because the obtained BER values for all attacks combined were greater than the sum of BERs for individual attacks for both ours and the comparative methods.

Comparative methods show inferior performance when dealing with attacks. As expected, sample suppression considerably disrupts the ability of these techniques to successfully detect a watermark. Existing audio watermarking techniques embed a particular watermark bit only at one point in the signal. So, when samples responsible for preserving that watermark bit are lost or suppressed, retrieving the watermark becomes unfeasible. DNNs are able to learn to hide watermark bits at multiple locations in the signal, resulting in a lower embedding rate but improved robustness. Overall results are better with random noise and low-pass filtering attacks. By tuning the design parameters of the comparative methods, we make a compromise between watermark imperceptibility and robustness. These parameters tend to reflect the strength

Table 5
Comparison of the proposed DNNs with state-of-the-art techniques in terms of robustness. Results in brackets are declared in references on an alternative dataset.

Watermarking scheme	BER %				
	No attacks	RN	LF	SS	RN+LF+SS
Model A-V1 (25% attack probability)	0.00	0.00	0.00	0.01	0.02
Model A-V2 (25% attack probability)	0.00	0.01	0.00	0.27	0.33
Model B (25% attack probability)	0.01	0.01	0.01	0.01	0.05
DCT-b1 [29]	0.96 (0.00)	4.31 (0.21)	0.98 (0.00)	11.95	20.70
DWT-IAMM [31]	(0.00)	(0.00)	(0.00)	–	–
QDFT-V1 [35]	–	(7.13)	(6.94)	–	–
[37]	0.01 (0.00)	0.01 (0.00)	1.41	1.93 (0.00)	3.69

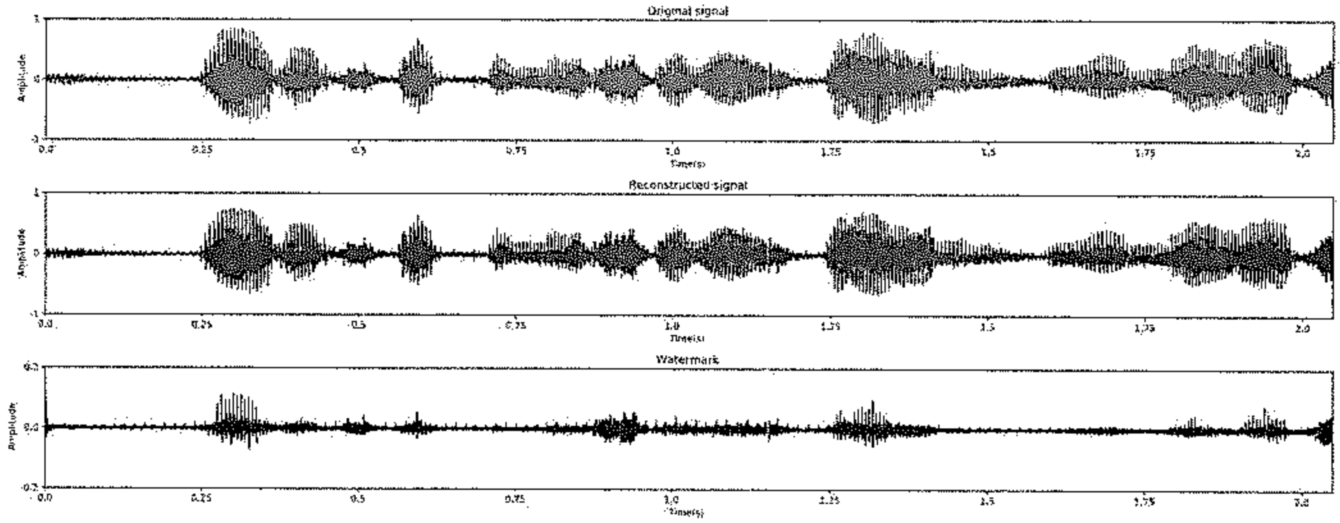


Fig. 7. Original signal, reconstructed signal and their difference or "watermark".

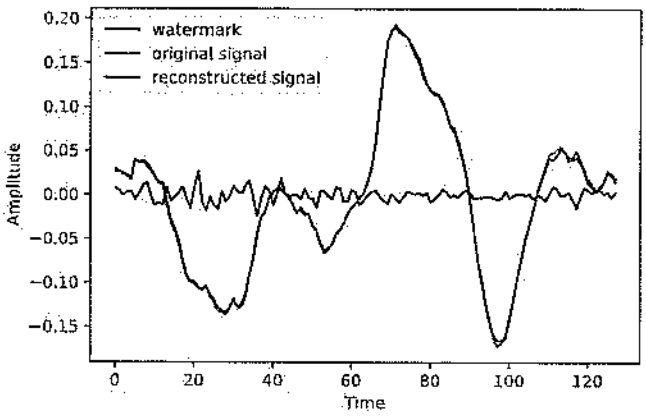


Fig. 8. Zoomed image of the original signal, reconstructed signal and their difference or "watermark".

Table 6
Comparison of the proposed DNNs with state-of-the-art techniques in terms of capacity.

Watermarking scheme	Capacity (bps)
Model A	256
Model B	64
DCT-b1 [29]	508.85
DWT-IAMM [31]	200
QDFT-V1 [35]	8.951
[37]	51.32

of the watermark, which means that they mostly affect resilience to random noise attacks. Performing watermark embedding in the low-frequency range, well below the cut-off frequency of the low-pass filter, should make these techniques resistant to this attack.

Results and comparisons in terms of capacity are given in Table 6. Model A achieved a higher capacity of 256 bps than Model B with 64 bps. From counterparts, [29,35] outperform our technique, while [31,37] are at similar capacity levels as Model A and Model B, respectively.

In order to facilitate the comparison of the proposed techniques with any other existing and future efforts, we have put the implementation details on our techniques into the public domain at https://github.com/kosta-pmf/dnn-audio-watermarking.

Both training and testing of our system were performed on NVIDIA Tesla P100 GPU with 16 GB of RAM. Table 7 provides information on embedding and detection runtimes for audio signals lasting 2 seconds, as well as the spatial requirements of our models. Model B runs faster than Model A since it has a significantly smaller number of parameters, and there is no need for performing STFT and inverse STFT operations. The smaller number of parameters in Model B is also part of the reason for the difference in performance compared to Model A.

8. Conclusions

Copyright infringement, identity theft, and the spread of false information are increasingly common in the modern world. Creators of audio content, whether they are musicians, actors, politicians, or some people who are not public figures, are targets of

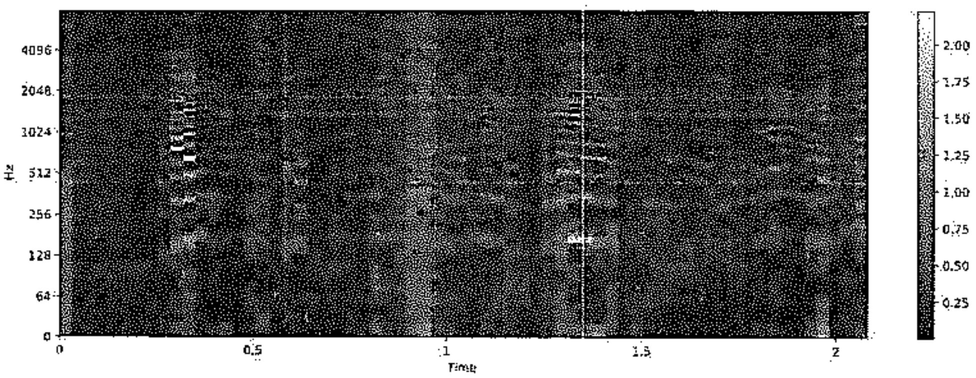


Fig. 9. STFT magnitude of the watermark signal from Fig. 7.

Table 7
Spatial complexities and runtimes of the embedder and the detector for Model A and Model B.

Watermarking scheme	Embedder param#	Detector param#	Embedder runtime (ms)	Detector runtime (ms)
Model A	7,375,778	9,186,560	34	18
Model B	2,948,425	1,397,600	21	9

these attacks, and it is of great importance that their rights are protected. Watermarking is an important procedure in this regard, as it helps to preserve copyrights and protect the information transmitted by the signal.

In this paper, a DNN-based speech watermarking system is proposed. The system consists of two neural networks, the embedder and the detector networks, that behave similarly to adversarial networks. While the embedder has the task of embedding the watermark so that it is completely hidden into the signal carrier, the detector should be able to always detect it, even when its input is a signal struck by several attacks. Two networks are jointly trained to converge in a carefully controlled procedure in which the priority between their tasks slowly changes. We trained two models with this type of architecture and training procedure. These models mainly differ in their inputs. One model (Model A) receives the STFT of the signal, and the other (Model B) works with raw speech signals. Model A exhibits better overall performance than Model B according to the main performance measures for audio watermarking systems.

A realistic dataset with over 800 hours of speech was processed and adapted for use. Our tests on this dataset show that the accuracy of the detector network is almost 100 percent. Additionally, the system shows superior robustness to comparative methods on a set of common watermarking attacks. As the number of designed attacks on audio watermarking systems is extremely large, the current set of attacks will be continuously expanded in the future with specific emphasis on desynchronization attacks. An ongoing development handles desynchronization attacks such as time scaling, sample cropping and resampling.

The obtained results for SNR and PESQ indicate that the watermarked signals are virtually indistinguishable from the original recordings. Model A-V2 achieved signal quality comparable to the best methods.

CRediT authorship contribution statement

Kosta Pavlović: Conceptualization, Data curation, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **Slavko Kovačević:** Data curation, Methodology, Software, Writing – original draft, Writing – review & editing. **Igor Djurović:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Adam Wojciechowski:** Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Ingemar J. Cox, Joe Kilian, Frank T. Leighton, Talal Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Process. 6 (12) (1997) 1673–1687.
[2] Fabien A.P. Petitcolas, Ross Anderson, Markus G. Kuhn, Information hiding-a survey, Proc. IEEE 87 (7) (1999) 1062–1078.
[3] Frank Hartung, Martin Kutter, Multimedia watermarking techniques, Proc. IEEE 87 (7) (1999) 1079–1107.
[4] Nikos Nikolaidis, Ioannis Pitas, Robust image watermarking in the spatial domain, Signal Process. 66 (3) (1998) 385–403.
[5] Wenjun Zeng, Bede Liu, A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images, IEEE Trans. Image Process. 8 (11) (1999) 1534–1548.
[6] Igor Djurović, Srdjan Stanković, Ioannis Pitas, Digital watermarking in the fractional Fourier transformation domain, J. Netw. Comput. Appl. 24 (2) (2001) 167–173.
[7] Srdjan Stanković, Igor Djurović, Ioannis Pitas, Watermarking in the space/spatial-frequency domain using two-dimensional Radon-Wigner distribution, IEEE Trans. Image Process. 10 (4) (2001) 650–658.
[8] Srdjan Stanković, Igor Djurović, Rainer Herpers, Ljubiša Stanković, An approach to optimal watermark detection, AEU, Int. J. Electron. Commun. 57 (5) (2003) 355–357.
[9] Lixin Luo, Zhenyong Chen, Ming Chen, Xiao Zeng, Zhang Xiong, Reversible image watermarking using interpolation technique, IEEE Trans. Inf. Forensics Secur. 5 (1) (2010) 187–193.
[10] Chih-Chin Lai, Cheng-Chih Tsai, Digital image watermarking using discrete wavelet transform and singular value decomposition, IEEE Trans. Instrum. Meas. 59 (11) (2010) 3060–3063.
[11] Frank Hartung, Bernd Girod, Watermarking of uncompressed and compressed video, Signal Process. 66 (3) (1998) 283–301.
[12] Guang Hua, Jiwu Huang, Yun Q. Shi, Jonathan Goh, Vrizlynn L.L. Thing, Twenty years of digital audio watermarking—a comprehensive review, Signal Process. 128 (2016) 222–242.
[13] Byeong-Seob Ko, Ryouichi Nishimura, Yoiti Suzuki, Time-spread echo method for digital audio watermarking, IEEE Trans. Multimed. 7 (2) (2005) 212–221.
[14] Yong Xiang, Iynkaran Natgunanathan, Dezhong Peng, Wanlei Zhou, Shui Yu, A dual-channel time-spread echo method for audio watermarking, IEEE Trans. Inf. Forensics Secur. 7 (2) (2012) 383–392.

[15] Guang Hua, Jonathan Goh, Vrizlynn L.L. Thing, Time-spread echo-based audio watermarking with optimized imperceptibility and robustness, IEEE/ACM Trans. Audio Speech Lang. Process. 23 (2) (2015) 227–239.

[16] Paraskevi Bassia, Ioannis Pitas, Nikos Nikolaidis, Robust audio watermarking in the time domain, IEEE Trans. Multimed. 3 (2) (2001) 232–241.

[17] Wen-Nung Lie, Li-Chun Chang, Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification, IEEE Trans. Multimed. 8 (1) (2006) 46–59.

[18] Shijun Xiang, Jiwu Huang, Histogram-based audio watermarking against time-scale modification and cropping attacks, IEEE Trans. Multimed. 9 (7) (2007) 1357–1372.

[19] Xingyuan Liang, Shijun Xiang, Robust reversible audio watermarking based on high-order difference statistics, Signal Process. 173 (2020) 107584.

[20] Darko Kirovski, Henrique S. Malvar, Spread-spectrum watermarking of audio signals, IEEE Trans. Signal Process. 51 (4) (2003) 1020–1033.

[21] Henrique S. Malvar, Dinei A. Florencio, Improved spread spectrum: a new modulation technique for robust watermarking, IEEE Trans. Signal Process. 51 (4) (2003) 898–905.

[22] Li Wei, Xiangyang Xue, Peizhong Lu, Localized audio watermarking technique robust against time-scale modification, IEEE Trans. Multimed. 8 (1) (2006) 60–69.

[23] David Megías, Jordi Serra-Ruiz, Mehdi Fallahpour, Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification, Signal Process. 90 (12) (2010) 3078–3092.

[24] Yong Xiang, Iynkaran Natgunanathan, Dezhong Peng, Guang Hua, Bo Liu, Spread spectrum audio watermarking using multiple orthogonal PN sequences and variable embedding strengths and polarities, IEEE/ACM Trans. Audio Speech Lang. Process. 26 (3) (2018) 529–539.

[25] Brian Chen, Gregory W. Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, IEEE Trans. Inf. Theory 47 (4) (2001) 1423–1443.

[26] Shaoquan Wu, Jiwu Huang, Daren Huang, Yunqing Shi, Efficiently self-synchronized audio watermarking for assured audio data transmission, IEEE Trans. Broadcast. 51 (1) (2005) 69–76.

[27] Xiang-Yang Wang, Hong Zhao, A novel synchronization invariant audio watermarking scheme based on DWT and DCT, IEEE Trans. Signal Process. 54 (12) (2006) 4835–4840.

[28] Baiying Lei, Ing Yann Soon, Ee-Leng Tan, Robust SVD-based audio watermarking scheme with differential evolution optimization, IEEE Trans. Audio Speech Lang. Process. 21 (11) (2013) 2368–2378.

[29] Hwai-Tsu Hu, Ling-Yuan Hsu, Robust, transparent and high-capacity audio watermarking in DCT domain, Signal Process. 109 (2015) 226–235.

[30] Min-Jae Hwang, JeeSok Lee, MiSuk Lee, Hong-Goo Kang, SVD-based adaptive QIM watermarking on stereo audio signals, IEEE Trans. Multimed. 20 (1) (2018) 45–54.

[31] Hwai-Tsu Hu, Tung-Tsun Lee, Frame-synchronized blind speech watermarking via improved adaptive mean modulation and perceptual-based additive modulation in DWT domain, Digit. Signal Process. 87 (2019) 75–85.

[32] Daniel Gruhl, Anthony Lu, Walter Bender, Echo hiding, in: Information Hiding, Springer, Berlin Heidelberg, 1996, pp. 295–315.

[33] Michael Arnold, Audio watermarking: features, applications and algorithms, in: 2000 IEEE International Conference on Multimedia and Expo, ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No. 00TH8532), vol. 2, February 2000, pp. 1013–1016.

[34] In-Kwon Yeo, Hyoung Joong Kim, Modified patchwork algorithm: a novel audio watermarking scheme, IEEE Trans. Speech Audio Process. 11 (4) (2003) 381–386.

[35] Eurchi Salah, Khaldi Amine, Kafi Redouane, Kahlessenane Fares, A Fourier transform based audio watermarking algorithm, Appl. Acoust. 172 (2021) 107652.

[36] Zhenghui Liu, Yuankun Huang, Jiwu Huang, Patchwork-based audio watermarking robust against de-synchronization and recapturing attacks, IEEE Trans. Inf. Forensics Secur. 14 (5) (2019) 1171–1180.

[37] Slami Saadi, Ahmed Merrad, Ali Benziane, Novel secured scheme for blind audio/speech norm-space watermarking by Arnold algorithm, Signal Process. 154 (2019) 74–86.

[38] Xiaojuan Xu, Hong Peng, Chengyuan He, DWT-based audio watermarking using support vector regression and subsampling, in: Applications of Fuzzy Sets Theory, Springer, Berlin Heidelberg, 2007, pp. 136–144.

[39] Lukas Tegendal, Watermarking in audio using deep learning, Master's thesis, Linköping University, 2019.

[40] Seung-Min Mun, Seung-Hun Nam, Haneol Jang, Dongkyu Kim, Heung-Kyu Lee, Finding robust domain from attacks: a learning framework for blind watermarking, Neurocomputing 337 (2019) 191–202.

[41] Haribabu Kandi, Deepak Mishra, Subrahmanyam R.K. Sai Gorthi, Exploring the learning capabilities of convolutional neural networks for robust image watermarking, Comput. Secur. 65 (2017) 247–268.

[42] Farah Deeba, She Kun, Fayaz Ali Dharejo, Hameer Langah, Hira Memon, Digital watermarking using deep neural network, Int. J. Comput. Math. Learn. 10 (2) (2020) 277–282.

[43] Jiren Zhu, Russell Kaplan, Justin Johnson, Fei-Fei Li Hidden, Hiding data with deep networks, in: 15th European Conference, Munich, Germany, September 2018, pp. 682–697.

[44] Kosta Pavlović, Slavko Kovačević, Igor Djurović, Speech watermarking using deep neural networks, in: 2020 28th Telecommunications Forum (TELFOR), 2020, pp. 1–4.

[45] Olaf Ronneberger, Philipp Fischer, Thomas Brox U-net, Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, vol. 9351, October 2015, pp. 234–241.

[46] Sergey Ioffe, Christian Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in: Proceedings of the 32nd International Conference on Machine Learning, Lille, France, in: Proceedings of Machine Learning Research, vol. 37, PMLR, July 2015, pp. 448–456.

[47] Bing Xu, Naiyan Wang, Tianqi Chen, Mu Li, Empirical evaluation of rectified activations in convolutional network, CoRR, arXiv:1505.00853 [abs], 2015.

[48] Vinod Nair, Geoffrey E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), Omnipress, Haifa, Israel, June 2010, pp. 807–814.

[49] Xiang Li, Shuo Chen, Xiaolin Hu, Jian Yang, Understanding the disharmony between dropout and batch normalization by variance shift, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019, pp. 2677–2685.

[50] Günter Klambauer, Thomas Unterthiner, Andreas Mayr, Sepp Hochreiter, Self-normalizing neural networks, CoRR, arXiv:1706.02515 [abs], 2017.

[51] Yann A. LeCun, Léon Bottou, Genevieve B. Orr, Klaus-Robert Müller, Efficient BackProp, in: Lecture Notes in Computer Science, Springer, Berlin Heidelberg, 2012, pp. 9–48.

[52] Xavier Glorot, Yoshua Bengio, Understanding the difficulty of training deep feedforward neural networks, J. Mach. Learn. Res. - Proceedings Track 9 (2010) 249–256.

[53] Martin Steinebach, Fabien A.P. Petitcolas, Frederic Raynal, Jana Dittmann, Caroline Fontaine, Christian Seibel, Nazim Fates, Lucilla Ferri, Stirmark benchmark: audio watermarking attacks, in: Proceedings International Conference on Information Technology: Coding and Computing, May 2001, pp. 49–54.

[54] Andreas Lang, Jana Dittmann, Ryan Spring, Claus Vielhauer, Audio watermark attacks: from single to profile attacks, in: Proceedings of the 7th Workshop on Multimedia and Security - MM&Sec '05, January 2005, pp. 39–50.

[55] Maryam Tanha, Seyed Dawood Sajjadi Torshizi, Mohd Taufik, Fazirulhisyam Hashim Abdullah, An overview of attacks against digital watermarking and their respective countermeasures, in: Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), June 2012, pp. 265–270.

[56] The European Broadcasting Union, Sound quality assessment material recordings for subjective tests, https://tech.ebu.ch/publications/sqamcd, 2008.

[57] John S. Garofolo, Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett, Nancy L. Dahlgren, Victor Zue, TIMIT acoustic-phonetic continuous speech corpus LDC93S1, https://catalog.ldc.upenn.edu/LDC93S1, 1993.

[58] Thomas Baer, Brian Moore, Karolina Kluk, Effects of low pass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies, J. Acoust. Soc. Am. 112 (2002) 1133–1144.

[59] Benjamin Hornsby, Todd Ricketts, The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding, J. Acoust. Soc. Am. 113 (2003) 1706.

[60] Patrick Bas, Teddy Furon, A new measure of watermarking security: the effective key length, IEEE Trans. Inf. Forensics Secur. 8 (8) (2013) 1306–1317.

[61] Yi Hu, Philipos C. Loizou, Evaluation of objective quality measures for speech enhancement, IEEE Trans. Audio Speech Lang. Process. 16 (1) (2008) 229–238.

[62] Antony W. Rix, John G. Beerends, Michael P. Hollier, Andries P. Hekstra, Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs, in: 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221), vol. 2, February 2001, pp. 749–752.

Kosta Pavlović was born in Montenegro in 1994. He received the B.Sc and M.Sc. degrees in computer science from the Faculty of Natural Sciences and Mathematics, University of Montenegro in 2016 and 2018, respectively. His research interests include digital watermarking, machine learning, audio signal processing, document processing, computer vision and optimization.

Slavko Kovačević received the B.Sc and M.Sc. degrees in computer science from the Faculty of Eletrical Engineering, University of Montenegro in 2019 and 2021. His research interests include digital signal processing, machine learning and computer vision.

Igor Djurović was born in Montenegro in 1971. He received the B.S., M.S., and Ph.D. degrees, all in electrical engineering, from the University of Montenegro, in 1994, 1996, and 2000, respectively. He is currently

a Professor with the University of Montenegro and Visiting Scholar at the Lodz University of Technology. During 2002, he was on leave with the Department of Mechanical and System Engineering, Kyoto Institute of Technology, Japan, as the Japan Society for the Promotion of Science fellow. He also was a visiting professor and joint researcher with ENSIETA, Brest, France, during 2003, and visiting researcher at the Gipsa Lab, INP Grenoble, France 2009, supported by the CNRS. He was on short stays in the AIIA Laboratory, Aristotle University, Thessaloniki, the Signal Theory Group, Ruhr University, Bochum, Germany, International Center for Signal Processing, Tampere University of Technology, Tampere, Finland, National Aerospace University, Kharkov, Ukraine, etc. He published more than 250 papers in international scientific journals and conferences. He has published eight book chapters including contribution to Time-Frequency Signal Analysis and Processing, B. Boashash, ed. He is coauthor of Time-frequency based feature extraction and classification: Considering energy concentration as a feature using the Stockwell transform and related approaches, VDM Verlag. He was member of editorial board of the Research Letters in Signal Processing, the Elsevier Signal Processing and Journal of Electrical and Computer Engineering and leading guest editor of the special issue of Eurasip Journal on Advances in Signal Processing, Special issue on Robust Signal Processing of Nonstationary Signals. He is member of the Montenegrin Academy of Sciences and Arts. He has run numerous national, regional, bilateral and international projects. He was director of the first Montenegrin Centre of Excellence in Bio-informatics (BIO-ICT). His current research interests include application of virtual instruments, time-frequency analysis-based methods for signal estimation and filtering, fractional Fourier transform applications, image processing, robust estimation, parametric and nonparametric estimation, spectral analysis, motion estimation, and digital watermarking.

Adam Wojciechowski received the M.Sc. degree in computer science from the Faculty of Technical Physics, Information Technology and Applied Mathematics, Lodz University of Technology, in 1999, the Ph.D. degree in computer science, (specialization in computer graphics and vision) from the Faculty of Automatics, Electronics and Computer Science, Silesian University of Technology, in 2005, and the D.Sc. degree in computer science from the Faculty of Mechanical Engineering and Computer Science, Czestochowa University of Technology, in 2014. His research interests include human–computer interaction, machine learning, computer graphics, computer vision, computer games as well as virtual environments.

International Journal of Wavelets, Multiresolution
and Information Processing
(2023) 2350009 (20 pages)

DOI: 10.1142/S0219691323500091

DNN-based speech watermarking resistant to desynchronization attacks

Kosta Pavlović
Faculty of Natural Sciences and Mathematics
University of Montenegro
Podgorica, Montenegro
kosta@ucg.ac.me

Slavko Kovačević
Faculty of Electrical Engineering
University of Montenegro
Podgorica, Montenegro
skovacevic@ucg.ac.me

Igor Djurović
Faculty of Electrical Engineering
University of Montenegro
Podgorica, Montenegro
Montenegrin Academy of Sciences and Arts
Podgorica, Montenegro
igordj@ucg.ac.me

Adam Wojciechowski
Institute of Information Technology
Lodz University of Technology
Łódź, Poland
adam.wojciechowski@p.lodz.pl

Received 31 December 2022
Revised 30 January 2023
Accepted 31 January 2023
Published

Desynchronization attacks proved to be the greatest challenge to audio watermarking systems as they introduce misalignment between the signal carrier and the watermark. This paper proposes a DNN-based speech watermarking system with two adversarial networks jointly trained on a set of desynchronization attacks to embed a randomly generated watermark. The detector neural network is expanded with spatial pyramid pooling layers to be able to handle signals affected by these attacks. A detailed training procedure of the aforementioned DNN system with gradual attack introduction is proposed in order to achieve robustness. Experiments performed on a speech dataset show that the system achieves satisfactory results according to all the benchmarks it was tested against. The system preserves signal quality after watermark embedding. Most importantly, the system achieved resistance to all considered desynchronization attacks. The majority of the attacks cause less than 1.70% of incorrectly detected watermarked bits on average, which outperforms comparative techniques in this regard.

Keywords: Deep neural networks; desynchronization attacks; spatial pyramid pooling; speech watermarking.

AMS Subject Classification: 68T07

1. Introduction

Intellectual property is considered to be any creativity in science or art, whether in written, spoken or another form. The intangible nature of the intellectual property often makes it challenging to prevent copyright infringement. The protection of intellectual property in the modern digital world is becoming increasingly important as the expansion of the Internet made intellectual theft more frequent. Recent breakthroughs in the field of deep neural networks (DNNs) made generating artificial data (so-called "deep fakes") possible. These data could be aimed at creating forgeries, spreading misinformation, defaming public figures, identity theft, etc. Watermarking is a process during which digital data get marked by a watermark in order to preserve its copyright and authenticity. Watermark detection is a process in which a watermark gets extracted from the signal carrier.

In the context of audio data, embedding a watermark in audio signals is necessary not only for the purpose of protecting intellectual property but also for the purpose of protecting the information which that signal carries. The watermark can be a message of any kind, but it must be of a certain length. Unlike other signal types, it is often unacceptable for audio watermarks to be perceptible to humans. Watermarking systems should be resistant to noise, compression and malicious attacks in order to maintain their effectiveness. Therefore, watermarking should represent a trade-off between robustness, data quality and data rate.

Digital watermarking has been an active field of research since the 1990s.[2] Since then, numerous techniques have been introduced for watermarking every type of digital multimedia, with the main ideas originating from image watermarking.[6,17,24,30,42,46,51]

Ultrasound watermarking is the most obvious way to embed the watermark into audio signals, but these kinds of straightforward approaches are susceptible to any sort of interference. Therefore, a watermark should be present in the entirety of the frequency spectrum in order to ensure its robustness. Over the years, a number of advanced watermarking techniques have been developed. These techniques can perform watermark embedding in the time domain with time-aligned[26] and echo-based[16,44] methods. However, the embedding is more often performed in one of the well-known transform domains, such as the discrete Fourier transform

(DFT), the discrete cosine transform (DCT) or some discrete wavelet transform (DWT).[11,12,32,48] Various techniques are used, such as spread spectrum,[45] patchwork algorithm,[27,34] quantization index modulation (QIM),[15,18,20] as well as singular value decomposition (SVD).[18,47]

Kernel techniques were the first to introduce the concept of learning into the area of multimedia watermarking.[5,38] The detector is trained to memorize the differences between the original and the watermarked signal, while in the embedding phase, traditional watermarking techniques are proposed. The revitalization of deep learning has led to its penetration into various areas of signal processing,[3,19,28,29] as well as the area of digital signal watermarking. Several papers exploit DNNs for both watermark embedding and watermark detection.[21,33,37,50] According to multiple universal approximation theorems, neural networks can approximate any continuous function to arbitrary precision. Neural networks, with one[4,14] or two[23] hidden layers, with suitable activation functions, and an abundant but finite number of neurons, have universal approximation capabilities. A recent paper[49] proved the universal approximation theorem for deep convolutional neural networks with a sufficient number of layers. This makes DNNs appropriate for overcoming problems such as the nonlinearity of desynchronization attacks. The DNN architecture proposed in this paper balances the tasks of watermark embedding and detection as they oppose each other. This system could be divided into two networks: the embedder, which generates its own transform domain and embeds the watermark in it, and the detector, which extracts the watermark from the signal carrier. The embedder should essentially perform a certain form of fusion between the signal carrier and the watermark so that the quality of the output is maximized. Fusion techniques for images[25] manage to produce excellent results which are more informative and of higher quality for both human and machine perception. The detector has to be able to extract watermark bits even when the signal is corrupted to some extent.

Audio watermarking systems are evaluated according to several criteria: their ability to preserve signal quality, the amount of data they can transmit per unit of time, and their resistance to various attacks. For each of these criteria, there are different measures that numerically characterize how successfully that criterion has been met. Robustness is measured in the terms of bit error rate (BER), i.e. the percentage of incorrectly detected watermark bits. System capacity is measured in bits per second (bps). A number of measures have been proposed for signal quality. The most widely used is the signal-to-noise ratio (SNR). In this paper, apart from SNR, we calculated perceptual evaluation of speech quality (PESQ).[40] This is a measure that incorporates the properties of the human auditory system when estimating signal quality, and its usage is recommended by the ITU Telecommunication Standardization Sector (ITU-T). PESQ more severely penalizes changes in frequency regions to which the human ear is more sensitive while placing less emphasis on less disturbing changes. SNR measure, conversely, treats all changes equally.

Robustness is the main requirement for audio watermarking systems due to the multitude and continuous development of possible attacks. Attacks are most often designed to disable watermark detection, but sometimes also to cause unauthorized or erroneous detection. The proposed network is trained to be robust against the most common desynchronization attacks. These attacks constitute a third part of our system, which is placed between the embedder and the detector networks and is an integral part of the entire network architecture. One of the main features of these attacks is that they change the length of the signal. Conventional DNNs require a fixed size input which means that this feature of desynchronization attacks could potentially cause performance degradation or completely disable the network. We propose solving this problem by introducing spatial pyramid pooling (SPP) layers[13] inside our network architecture. The system is trained on batches with a specific, randomly generated watermark message, random false watermarks and no watermark at all, to prevent the network to produce false positives. Due to the opposing objectives of the two networks and the abundance of attacks, bringing this system to the point of convergence is a demanding task. Hence, significant attention is paid to the design of the training procedure. It is thoroughly monitored with careful selection of the values for all the hyperparameters.

The rest of this paper is organized as follows. Section 2 describes the signal distortions, i.e. attacks, against which the network is trained. A detailed overview of the architecture and training procedure is given in Secs. 3 and 4, respectively. The obtained results and comparisons with other prominent techniques are presented in Sec. 5, with conclusions in Sec. 6.

2. Watermarking Attacks

Desynchronization attacks have proven to be the most effective type of attack on audio watermarking systems. Current watermarking techniques are resistant to common watermarking attacks, such as noise addition, filtering, etc. However, they are unable to achieve robustness against desynchronization attacks or could achieve it but at a very low embedding rate or without retaining high signal quality. Watermarking systems usually expect an ideal alignment of the audio signal and the watermark during detection. Desynchronization attacks disturb this alignment, which is the main reason for their effectiveness. Furthermore, these attacks are nonlinear operations. It is difficult to adequately reconstruct a signal altered with nonlinear operations with linear procedures, such as common watermark detection algorithms. Therefore, desynchronization attacks remain the focus of many researchers. Techniques from Refs. 20, 27, 34 and 47 are designed with a special intent to tackle this class of the attacks. A patchwork-based audio watermarking technique is proposed in Ref. 34. This method embeds watermark bits by altering means of absolute-valued DCT fragments. Repeating the same procedure several times results in a multi-layer watermark embedding system, which increases embedding capacity but reduces signal quality. To prevent significant quality degradation, an ordering process of the DCT coefficients is devised.

Authors in Ref. 47 suggest using the frequency singular value coefficient (FSVC) feature for watermark embedding. FSVC is extracted by performing SVD on mid-frequency band DCT coefficients of two consecutive fragments of a signal. This feature is then modified to embed watermark bits. Authors claim that the FSVC feature is less sensitive to desynchronization attacks, making it more robust than its counterparts. Both of the techniques mentioned above introduce an error buffer to increase robustness.

Some techniques,[15,27] use synchronization codes to deal with desynchronization attacks. They embed synchronization codes into the frames of the signal carrier along with watermark bits. During watermark extraction, synchronization codes are extracted first. If these codes are extracted successfully, the frame is considered to be synchronized, and watermark bits can be extracted from it. Otherwise, the frame is discarded. Desynchronization attacks affect the detection of these codes and cause frames to be discarded and watermark extraction to not be performed. Hence, synchronization codes can serve to indicate if a desynchronization attack has occurred and prevent wrong watermark detection. However, they do not provide the means for achieving robustness in the full sense. This imminent loss of information in the case of the attacks prevents us from determining whether the information is copyrighted or not. Just acknowledging that the attacks occurred is not sufficient. The failure to detect the watermark in such situations would make copyright protection infeasible. The changes introduced by the attacks could be subtle and undetectable to the human auditory system which would enable unhindered distribution and consumption of such audio content.

In determining the exact set of attacks against which we will test our approach, we were guided by the renowned StirMark benchmark[43] for audio watermarking. We consider four types of desynchronization attacks from this benchmark: jitter, sample permutations, resampling and time scaling. Every attack is associated with a set of parameters that control its severity. Exact values of certain attack parameters are randomly sampled from a predefined ranges before passing each training batch through the network. The parameter ranges were chosen in such a way that the attacks would disrupt the signal as much as possible, but not cause the loss of the information carried by the signal. This is thought to be the most common attacking scenario, as the attacker does not want to completely destroy the signal, because that would render it useless to him too.

Although low-pass filtering and random noise do not represent desynchronization attacks, they are included in our set. These are two essential attacks that every audio watermarking system should be resistant to. Achieving robustness against desynchronization attacks while skipping these basic attacks would not be credible. We also include a time delay attack from the StirMark benchmark. This attack adds a delayed copy to the signal. It could seriously disturb traditional watermarking techniques because it can disrupt the division of the signal into frames, which is the very first step of the vast majority of these techniques.

Attacks are incorporated into watermarking system as neural network layers. These layers are used to present potential watermark detection difficulties to the system during training and allow it to produce a more robust watermarking scheme. All of the attacks against which the system is trained, along with their selected severity levels, are listed here

(1) Jitter — up to 5% samples,
(2) Samples permutation — up to 5% samples,
(3) Downsampling — to 8 kHz,
(4) Upsampling — to 32 kHz,
(5) Time fold — up to 10% reduction,
(6) Time stretch — up to 10% increase,
(7) Low-pass filtering — 4 kHz cutoff frequency,
(8) Random noise — up to 30 dB,
(9) Time delay — up to 10% delay.

More details on the attacks, their parameters and realizations are given in Appendix A.

3. DNN Architecture

Significant efforts have been invested in researching domains in which it is most suitable to analyze and process signals. The domain of the input data is also an important factor in the architectural design of a watermarking system. Domain knowledge can significantly contribute to the overall results. The time domain is the default for audio watermarking. Audio watermarking in the time domain can be expressed via the generic formula:

$$y(n) = x(n) + \alpha w(n), \tag{3.1}$$

where y stands for a watermarked signal, while x and w represent the original signal and the watermark, respectively. Parameter α scales the watermark strength and impacts its audibility and detectability.

However, as stated in Sec. 1, time-domain watermarking is commonly considered inadequate, which led to the usage of various transform domains such as DFT, DCT, FrFT, etc. Even more general techniques could be used. Wavelet decomposition[32] is a state-of-the-art technique for effective and exact analysis of different types of signals. It is highly valued in the field of image processing as it provides a complete image representation and performs decomposition according to both scale and orientation. Some recent breakthroughs that further improved the representation efficiency of the wavelet decomposition[18] could also be applied in the domain of audio processing. Different types of wavelets are considered,[11,12] and their properties, as well as potential applications in fractal analysis,[9,10] are evaluated. However, these techniques are yet to be fully explored in the area of digital watermarking.

Audio watermarking in the generic transform domain can be expressed via the following formula:

$$Y(k) = X(k) + \alpha W(k), \tag{3.2}$$

where Y stands for a watermarked signal in a transform domain, while X and W represent the original signal and the watermark in the same domain, respectively.

In this paper, short-time Fourier transform (STFT) is used as the input domain. It is a generalization of the DFT technique. The STFT is superior to frequency-domain representations as it enables the embedder to insert the watermark in suitable time intervals rather than only on certain frequencies. The advantage of STFT is also that it can be viewed as an initial feature extractor, which facilitates system training. The STFT is defined as follows:

$$\mathrm{STFT}(n,k) = \sum_{m=-N_w/2}^{N_w/2-1} g(m)x(n+m)e^{-j2\pi mk/N_w}, \tag{3.3}$$

where g represents a window function that localizes the signal in the time–frequency plane. N_w is the length of the window function. It is obvious that the quality of the STFT representation of the signal depends on the type of the window function and its length. One of the features of the STFT is that it is reversible, so the signal carrier can be reconstructed after watermark embedding without reducing its quality. STFT clearly retains information about both signal phase and amplitude. Hiding the watermark only within the signal amplitude is not recommended, as even the slightest changes could be detected by the human auditory system. If the change is not sufficient enough, the watermark itself would be hardly detectable. The network should learn to hide the watermark mostly within the signal phase, as the human auditory system is relatively insensitive to its changes.

The proposed architecture is mostly based on our previous work[37] and it is shown in Fig. 1. The architecture is divided into two networks, the embedder and the

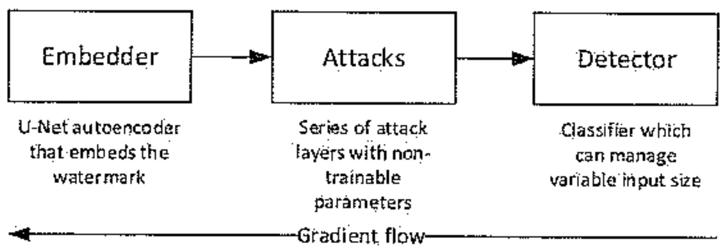


Fig. 1. Architecture.

Table 1. Embedder architectural parameters.

Type	Filters	Size/stride	Output
Convolutional	16	$5 \times 5/2$	256×32
Convolutional	32	$5 \times 5/2$	128×16
Convolutional	64	$5 \times 5/2$	64×8
Convolutional	128	$5 \times 5/2$	32×4
Convolutional	256	$5 \times 5/2$	16×2
Watermark embedding			
Convolutional	256	$5 \times 5/2$	16×2
Transposed convolutional	128	$5 \times 5/2$	32×4
Transposed convolutional	64	$5 \times 5/2$	64×8
Transposed convolutional	32	$5 \times 5/2$	128×16
Transposed convolutional	16	$5 \times 5/2$	256×32
Transposed convolutional	8	$5 \times 5/2$	512×64
Convolutional	2	$5 \times 5/1$	512×64

detector. However, there is a pivotal advancement that tackles desynchronization attacks which are insurmountable for the aforementioned architecture.

The embedder network remained the same as in Ref. 37. It is based on the U-net design[41] with five downsampling and five upsampling blocks and it uses batch normalization. The input signal is downsampled until it has a size of $16 \times 2 \times 256$. In this latent space representation, the watermark, which is a one-dimensional array, gets embedded by repeating it 16×2 times. The upsampling part of the embedder expands the signal back to its original size. The embedder architectural parameters are presented in Table 1.

To achieve robustness, the embedder needs to insert watermark bits in such a way that the attacks affect the detection as little as possible. This is why the detection error caused by watermarking attacks must be propagated to the embedder. Therefore, neural network layers that simulate watermarking attacks are included in the training procedure. These layers provide information to the embedder about attacks that could occur so that it can learn to perform robust watermark embedding. Attack layers do not have trainable parameters but are viewed as an integral part of the neural network since the detection error flows through them to the embedder.

The detector has the task of extracting the watermark from the signal carrier. Its architecture is based on an image classifier that ends with a fully connected layer. However, desynchronization attacks, which are the focal point of this paper, impose certain architectural features on the detector. While convolutional layers can work with inputs that vary in size, fully connected layers enforce the requirement of fixed-length inputs. Neural networks are usually trained with fixed-length inputs but desynchronization attacks break this norm as they change the length of the signal. Although the input dimensions of the convolutional layer can vary in size, variable input size would result in variable output size. The detector output,

Table 2. Detector architectural parameters. Number L_w represents the length of the watermark (number of watermark bits). The second dimension of layers' outputs is marked with ? due to unknown signal length after applying desynchronization attacks.

Type	Filters	Size/Stride	Output
Convolutional	32	$5 \times 5/2$	$256 \times ?$
Convolutional	32	$5 \times 5/2$	$128 \times ?$
Convolutional	64	$5 \times 5/1 \times 2$	$128 \times ?$
Convolutional	64	$5 \times 5/1 \times 2$	$128 \times ?$
Convolutional	128	$5 \times 5/1 \times 2$	$128 \times ?$
Convolutional	128	$5 \times 5/1 \times 2$	$128 \times ?$
SPP		$1, 8, 32, 128$	128×169
Fully connected		$21632 \times L_w$	L_w

however, must be of a fixed size as the watermark is of a fixed size as well. Therefore, in order to allow the network to properly handle batches whose signal lengths have been changed through desynchronization attacks, a two-dimensional SPP layer[13] is introduced. These layers are placed before the fully connected layers of the detector. SPP consists of four one-dimensional adaptive max pooling layers. Bin sizes for these layers are given in Table 2. This ensures that the input dimensions of the fully connected layer are constant. The SPP layers not only made our neural networks agnostic of input size but reduced the chance of overfitting the model. That is why the coarsest level has a single bin that performs a global pooling operation. Binary-cross entropy loss is calculated between the output of the detector and the original watermark. Other detector architectural parameters are also presented in Table 2.

When training the network with variable input sizes, it is advised to form batches of data that match in size. In this paper, the size of the data used to input into the network is constant. However, the dimensions change after desynchronization attacks are applied between the two networks. Hence, to increase the likelihood of convergence, attacks are performed on the whole batch, one attack at a time.

4. Training Procedure

Conducting a successful training procedure, in this case, means fulfilling the tasks of the two networks, the embedder and the detector. These two networks have different and somewhat opposing goals. The embedder strives to maintain signal quality, while the detector strives to extract all watermark bits. Apart from preserving signal quality and errorless watermark detection, robustness is the third requirement for the watermarking system. It is also the most demanding. The training procedure is designed by taking robustness into special consideration. We train our system to be resistant to all of the desynchronization attacks from Sec. 2.

Transfer learning is one of the most important methods in machine learning today. It is used as one of the main means to solve various tasks across all application areas of machine learning. The idea is to reuse an already trained model as a starting point for solving a new task. That model could be trained using different data, but on a task that is similar to the task at hand. This was also the guiding idea when designing our training procedure. We split our problem into several subtasks and created checkpoints after solving each task to be used as a starting setup before initiating the next task. Our idea differs in certain points from transfer learning in the traditional sense. We use the same dataset throughout all the subtasks and do not perform any architectural changes during training. Furthermore, training for previous subtasks is not terminated when a new subtask is initiated. It continues, but at a reduced scale.

Introducing all of the attacks at once showed to be overwhelming for the optimization procedure. Hence, they are introduced one by one as subtasks. We divided the training procedure into N_c cycles of 200 batches. Training lasts 60 cycles, which equals to one pass through the whole training dataset, i.e. an epoch. This is sufficient due to the amount of available data. Changes to the training procedure are made on a per cycle basis. At the start of the training, the network is unable to embed or extract the watermark. Hence, the first two cycles are conducted without attacks to train the network to perform its initial tasks, thus reaching a good starting point for combating attacks. After that, we introduce a new attack every two cycles. The first cycle upon establishing a new attack is used for the system to learn the characteristics of that attack. The attack is performed on the selected batches without any supplementary operations. Starting from the second cycle upon presenting the new attack to the system, and further on, the attack is coupled with low-pass filtering. This is necessary in order to prevent the network to embed watermark bits in high frequencies. Without it, networks would almost certainly use high frequencies to fight attacks, because there isn't much audio content in those frequencies, especially human speech, and that is where it is much easier to add new information and extract it later. The embedder could make more subtle changes in the high-frequency range, making it easy for the detector to locate them. Conversely, these changes in high frequencies, even subtle in their intensity, would be fairly disturbing to human hearing. By coupling all the attacks with low-pass filtering, we force the networks to use lower frequencies for embedding the watermark.

Each attack is associated with the probability of its occurrence in a given cycle i. This probability can be defined as

$$p(i,j) = \begin{cases} p_a \cdot p_{\text{new}}, & \lfloor i/2 \rfloor = j, \\ p_a \cdot (1 - p_{\text{new}}) \cdot \max(0, \operatorname{sgn}(i-j))/\lfloor i/2 \rfloor, & \lfloor i/2 \rfloor \neq j \ \& \ \lfloor i/2 \rfloor \leq N_a, \\ p_a \cdot N_a, & \lfloor i/2 \rfloor > N_a, \end{cases} \tag{4.1}$$

where j represents the ordinal number of the attack and i is the cycle index. N_a is the total number of the attacks. In Ref. 37, we concluded that it is necessary to retain a certain portion of batches that are not attacked in every cycle. The hyperparameter p_a is intended to control the size of this portion. It represents the probability of a batch being attacked. Previously, we decided to have 50% of non-attacked batches in each cycle. Here, we reduce this portion to 25%, since we deal with a significantly higher number of attacks $N_a = 9$. It follows that $p_a = 0.25$. The probabilities of the attacks in each cycle are differently distributed. The newly introduced attack is given priority, by assigning it a unique probability of occurrence p_{new}. We use value of 0.5 for p_{new} during our experiments. The rest of the attacks introduced up until that point have evenly distributed occurring probabilities. After the last attack is introduced in cycles 19 and 20, the system is trained for additional 40 cycles, with each attack having the same probability of occurrence.

Different goals lead to different loss functions for the two networks. Cross-entropy is the straightforward choice for the detection loss function. To determine the loss function for quality preservation, we considered two options, namely, mean absolute error (MAE) and mean squared error (MSE). As expected, using MSE leads to superior performance. There are several reasons for this. MSE loss function is much smoother than MAE, so there is less risk that the optimization process will reach a ridge. Introducing a new attack after every two cycles disrupts signal quality preservation as significant drops of signal quality measures can be seen in Fig. 2. Using a smoother loss function makes this process easier to steer. MSE loss emphasizes large errors, which could be considered as a drawback in certain applications because it makes the system susceptible to outliers. However, this drawback does not hold here since our data are outlier free. Furthermore, MSE contains a term similar to the SNR measure which is one of the main measures for the estimation of signal quality. It would be natural to conclude that this measure would be maximized by using a loss function as similar to it as possible. Following equations show SNR and MSE, respectively

$$\mathrm{SNR} = 10 \cdot \log_{10} \frac{\sum_{n=1}^{N} x(n)}{\sum_{n=1}^{N} (x(n) - y(n))^2}, \tag{4.2}$$

$$\mathrm{MSE} = \frac{1}{N} \sum_{n=1}^{N} (x(n) - y(n))^2. \tag{4.3}$$

The similarity between these two metrics that is highlighted here is the squared difference of the original signal x and the watermarked signal y.

Figure 2 shows trajectories of SNR and PESQ measures during training. These values are calculated on a small test set selected just for this purpose. MSE loss is used in one training setup, while MAE loss is used in the other. Both models reach similar performance in terms of robustness, but it can be seen that a higher signal quality is achieved when using MSE than when using MAE throughout the whole training process. Also, more stable growth of SNR is observed, compared to the PESQ, caused by the already mentioned similarity of SNR and MSE.

Quality preservation is the sole responsibility of the embedder, but both networks must participate in the minimization of the detection loss if robustness is

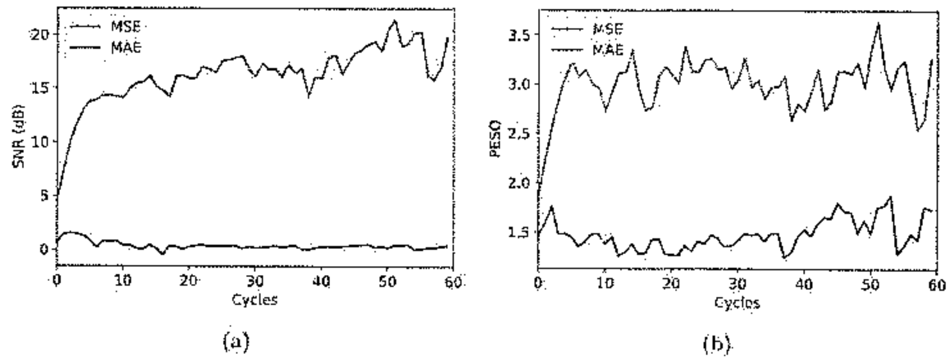


Fig. 2. Trajectories of SNR (a) and PESQ (b) measures during training.

ought to be achieved. The embedder must be trained to embed watermark bits in such a way that they are not lost after an attack. Both losses are unified for a joint training procedure and weight balancing is introduced to prevent the embedder from prevailing. If the embedder learns to reconstruct the original signal perfectly, the detector will not be able to extract any watermark bits, much less overcome all of the attacks. The quality preservation loss must be held at a certain level below optimal for the system to be able to achieve robustness. Weights w_q and w_d are assigned to the quality and the detection loss, respectively. The joint loss function is then calculated as

$$\text{Loss} = w_q \cdot Q_{\text{loss}} + w_d \cdot D_{\text{loss}}, \tag{4.4}$$

where Q_{loss} and D_{loss} are loss functions for quality preservation and watermark detection, respectively. We have experimentally concluded that the best results are obtained when the weights are in a 3 : 1 ratio, on the side of the detection.

The system is trained on the dataset from Ref. 37 that consists of 868 h of preprocessed speech data. These signals are divided into segments of 32 768 samples, sampled at 16 kHz, which is approximately 2 s. There is a stark difference in the number of messages embedded during the training procedure in contrast to our previous work.[37] While in Ref. 37, we used a randomly sampled pool of six messages, our new architecture is trainable on a virtually infinite number of messages. To achieve this, a significant change in the detector architecture is required as well. This change is caused by several drawbacks of the previous solution. First of all, any message that is not part of the original finite sample would be incorrectly classified as one of the messages from the pool. Second, not having a zero vector message as a valid option during training would result in an error while classifying audio signals that are not watermarked at all — the system would assume that they carry a certain message. Therefore, 40% of training batches are watermarked with "the watermark", a unique message reserved for the entirety of the dataset, and 40% of batches are watermarked with a random watermark so that the system could learn to differentiate between the right and the wrong message, while the remaining batches are not watermarked at all. When processing a non-watermarked batch, only the detector weights are updated.

In short, optimization is the process of finding the set of weights of the neural network layers that minimize the loss function. Finding an optimum solution in the N-dimensional problem is not an easy task — one that could be practically impossible without a proper optimizer. The fragility of the gradient descent is even more expressed in the case of the two networks that are jointly trained. Therefore, the appropriate optimization algorithm is chosen by trial and error method. We selected Nadam,[7] an improved Adam[22] algorithm with the Nesterov momentum.[35] The learning rate parameter is set to 0.0002.

5. Results

Since our primary goal is to overcome desynchronization attacks, the most emphasis is placed on robustness. However, this goal is achieved without seriously compromising other performance evaluation measures. We tested our system against the attacks from Sec. 2 and the obtained results are given in Table 3. For testing, we use a separate part of the dataset that is not used in training. Our system generally shows good performance across all experiments, with BER below 2% for almost all the attacks. No significant differences are observed between watermarked and non-watermarked signals in terms of robustness, so the results from Table 3 represent an average of the values for all three aforementioned types of batches. Results for each technique are given in two columns. The first column contains BER values when an attack is not coupled with low-pass filtering and the second column is for BER values when attacks are paired with a filter. As expected, the most problematic attacks are those in which signal samples are lost, i.e. jitter and time-fold. The reason for this is that these attacks, along with the loss of signal samples, cause the loss of information about watermark embedding. Downsampling attack also leads to sample loss, but more rigidly, and it is therefore easier to overcome.

We compare our approach with a recently published technique,[47] designed specifically to target desynchronization attacks in audio watermarking. In addition, a prominent patchwork-based multilayer watermark embedding technique from Ref. 34 is tested in the same setting. The experimental results that are obtained using these watermarking schemes are also given in Table 3. The proposed method shows significantly better overall performance in the terms of robustness. More pronounced differences in the BER values can be observed especially in the case of desynchronization attacks. This is completely expected as our system is allowed to adapt to these signal processing operations during training. For as fair a comparison as possible, all methods have been tested at a 20 bps embedding rate, which is generally accepted as the lower limit for the capacity of watermarking systems in the literature. Source code for the proposed technique is available at https://github.com/kosta-pmf/dnn-audio-watermarking. We also share

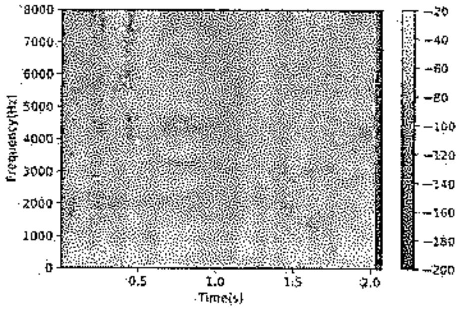
Table 3. Results in terms of robustness at 20 bps embedding rate.

Attack	Proposed		Ref. 34		Ref. 47	
	No filter	With filter	No filter	With filter	No filter	With filter
No attack	0.00	—	0.00	—	0.04	—
Jitter	1.34	1.70	50.02	50.29	16.37	17.39
Samples permutation	0.00	0.03	16.37	16.62	3.46	3.50
Downsampling	0.02	0.05	50.00	50.17	49.80	50.02
Upsampling	0.01	0.01	50.31	50.50	49.06	50.24
Time fold	3.02	7.85	39.44	39.86	38.33	38.62
Time stretch	0.38	0.57	44.06	44.57	36.33	36.70
Low-pass filtering	0.00	—	0.00	—	1.30	—
Random noise	0.00	0.00	4.20	4.23	1.21	1.30
Time delay	0.02	1.58	18.22	18.39	48.44	49.06

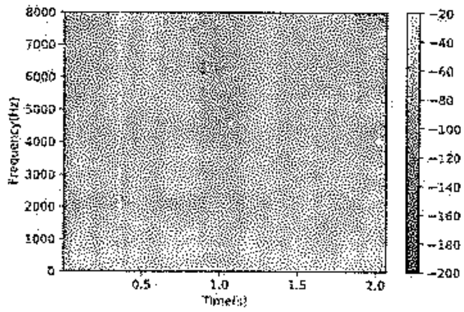
our implementations of the referenced techniques at https://github.com/kosta-pmf/audio-watermarking.

The proposed approach also provides a complete framework for creating audio watermarking systems. The system designer only needs to identify a set of watermarking attacks relevant to their application and implement appropriate neural network layers that approximate the attacks. It is required that these approximations are differentiable in order for network to be able to adapt to them and devise an embedding procedure resistant to the attacks. Conversely, traditional techniques would require substantial methodological changes, if some new class of attacks emerged.

Unlike music signals, sound quality is not crucial with speech signals. However, the information they carry is preeminent and must not be compromised. For this reason, a certain drop in quality can be tolerated in speech watermarking, as long as the information is preserved. Figure 3 contains spectrograms of an input speech signal, from the test set, and a watermarked signal, the output of the embedder network. This figure shows that the content of the speech is well preserved as there



(a)



(b)

Fig. 3. Spectrograms of an input signal (a) and a watermarked signal (b) on a dB scale.

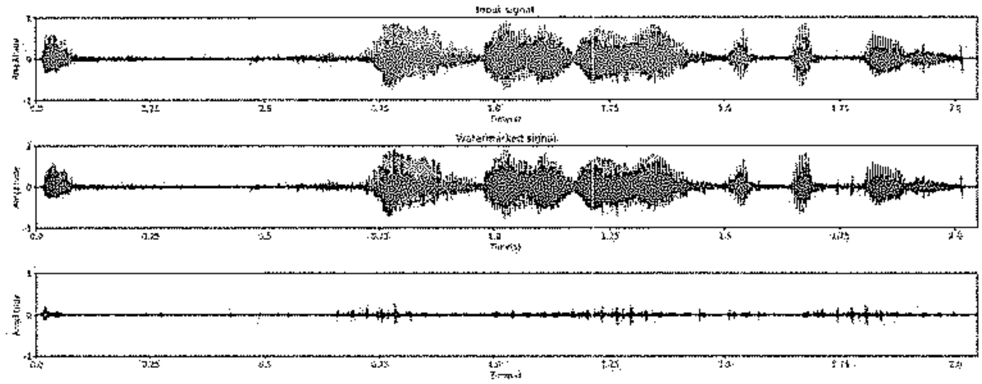


Fig. 4. Input signal, watermarked signal and their difference.

are no clearly noticeable differences between the two spectrograms in areas with speech content. The rest of the spectral content is somewhat worse reconstructed, but as this content is subtle, it will not significantly affect the values of the signal quality metrics. There are no highly noticeable differences on high frequencies, so the network learned not to use that range of frequencies for embedding watermark bits. Figure 4 shows these two signals in the time domain, as well as their difference. This figure also confirms a negligible difference between the two signals, suggesting that the interference introduced by the watermarking system is not too extreme.

Quality of the watermarked signals from our experiments can be seen in Fig. 2. SNR is around 20 dB, and PESQ is around 3, depending on which model checkpoint is selected. This is lower than our counterparts, but they are high enough to preserve speech content. The reason for the lower signal quality is the presence of the watermark in the entire spectrum necessary to achieve robustness. Both PESQ and SNR are lower for the signals with the real watermark, contrary to the ones watermarked with fake or nonexistent watermarks. This may be an indication that the system has learned not to watermark signals in these situations. In these cases, SNR reaches values around 23 dB, and PESQ reaches 4.

6. Conclusion

A robust system design is imperative in almost all areas of digital multimedia watermarking. In the case of audio watermarking, a class of desynchronization attacks presents the biggest obstacle. The nonlinearity and randomness of these attacks make them difficult for traditional watermarking techniques. In this paper, we proposed a set of desynchronization attacks that can be used as a benchmark for existing and future efforts. We also presented a modified architecture of a DNN-based watermarking system, and conducted a strictly controlled training procedure to bring this complex system to convergence. The system demonstrated resistance to set attacks while maintaining fair signal quality and respectable capacity. The proposed system should be improved primarily in terms of signal quality, so that it could be applied for other types of audio signals where quality is of greater importance, such as music signals. A number of watermarks that the system is able to embed, as an essential practical feature, remain for further research.

Appendix A. Desynchronization Attacks

This section contains a more detailed overview of the set of desynchronization attacks used within this study.

Jitter attack deletes a random set of samples from the watermarked signal. It can be represented as: $y(m) = x(n_m), n_m \in S$, where x and y are the input and output signal, respectively, $x(n_m)$ is a subsequence of $x(n)$ and S ($S \subset N$) is a set of sample indices that are to be retained. The severity of this attack can be regulated by changing the cardinality of the set S.

Samples permutation attack randomly shuffles a selected group of samples. Similarly to the jitter attack, this attack has only one hyperparameter which is the number of samples to be selected for the permutation.

Resampling changes the sampling rate of the signal, which qualifies this type of audio effect as a desynchronization attack. There are two types of resampling, namely downsampling and upsampling. A downsampling operation is often used in various audio systems as a completely legitimate operation for bandwidth reduction and signal compression purposes. Since this operation is so common, it is crucial that audio watermarking systems be resistant to it. It can be represented with: $y(m) = x(m \cdot F_d)$, where F_d is the downsampling factor. This type of downsampling could introduce aliasing when $\omega_N \geq \pi/F_d$, where ω_N is the normalized Nyquist frequency of the signal. Here, we are dealing with speech signals that have a low maximum frequency around 4 kHz, which means that we can perform downsampling to as low as 50% of our original sampling rate of 16 kHz without creating aliasing artifacts.

We also consider the upsampling attack that increases sampling rate by a factor of F_u. This attack consists of two steps. First, an expanded signal y_e is created by adding $F_u - 1$ zeros between every two samples. This can be represented with

$$y_e(m) = \sum_{n=0}^{N} x(n) \cdot \delta(m - n \cdot F_u), \tag{A.1}$$

where δ is the unit impulse signal. This operation has a very simple form in the frequency domain,[36] which is why it is chosen for the realization of this attack. The DFT of the expanded signal y_e can be expressed as: $Y_e(k) = X(k \cdot F_u)$. This means that the DFT of the expanded signal is the frequency-scaled DFT of the input signal x. Finally, interpolation needs to be performed to fill in the missing samples in the expanded signal. In our case, linear interpolation is adequate. It is performed by applying the linear interpolation filter h_{lin}, whose impulse response is given with: $h_{\text{lin}}(m) = 1 - |m|/F_u$ for $|m| \leq F_u$ and 0 elsewhere. The upsampled signal is obtained by convolving y_e with h_{lin}.

Time scaling is an archetypal desynchronization attack. It changes the playback time of the audio recording. We differentiate two types of time scaling attacks, namely time-fold and time-stretch. Time-fold represents a time scaling attack in which the duration of the signal is reduced. Time-stretch represents time scaling which increases the duration of the signal. This attack could be performed by interpolation, i.e. resampling of the audio signal. However, resampling also affects the pitch of the signal. In order to study time scaling as a separate watermarking attack, different techniques must be implemented. There are two general approaches to time scaling while preserving the pitch of the signal. The first group of approaches is based on the phase vocoder algorithm.[8] The second group consists of time-domain methods for audio and speech time scaling.[1,31,39] Because time-domain techniques are more prone to producing reverberating and clicking artifacts, this makes them more disruptive attacks. As the quality of the output signal is not crucial because these effects are treated as attacks, we have chosen time-domain methods for their efficiency. All these time-domain approaches essentially share a similar idea. The signal is first divided into a set of segments, which are then combined using the overlap-add technique. We use the synchronous-overlap-add algorithm (SOLA).[39]

References

1. F. Charpentier and M. Stella, Diphone synthesis using an overlap-add technique for speech waveforms concatenation, in *ICASSP '86: IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Vol. 11, Tokyo, Japan, 1986, pp. 2015–2018.
2. I. J. Cox, J. Kilian, F. T. Leighton and T. Shamoon, Secure spread spectrum watermarking for multimedia, *IEEE Trans. Image Process.* **6**(12) (1997) 1673–1687.
3. Z. Cui and C. Bao, Power exponent based weighting criterion for DNN-based mask approximation in speech enhancement, *IEEE Signal Process. Lett.* **28** (2021) 618–622.
4. G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signals Systems* **2** (1989) 303–314.
5. N. Dabas and R. P. Singh, ELM-kernel and reduced kernel ELM based watermarking scheme, *J. Inf. Secur. Appl.* **46** (2019) 173–192.
6. I. Djurović, S. Stanković and I. Pitas, Digital watermarking in the fractional Fourier transformation domain, *J. Netw. Comput. Appl.* **24**(2) (2001) 167–173.
7. T. Dozat, Incorporating Nesterov momentum into Adam, in *Int. Conf. Learning Representations 2016*, San Juan, Puerto Rico, 2016.
8. J. L. Flanagan and R. M. Golden, Phase vocoder, *Bell Syst. Tech. J.* **45**(9) (1966) 1493–1509.
9. E. Guariglia, Harmonic Sierpinski gasket and applications, *Entropy* **20**(9) (2018) 714.
10. E. Guariglia, Primality, fractality, and image analysis, *Entropy* **21**(3) (2019) 304.
11. E. Guariglia and R. Guido, Chebyshev wavelet analysis, *J. Funct. Spaces* **2022** (2022) 1–17.
12. E. Guariglia and S. Silvestrov, Fractional-wavelet analysis of positive definite distributions and wavelets on $\mathscr{D}'(\mathbb{C})$, in *Engineering Mathematics II* (Springer International Publishing, Cham, 2016), pp. 337–353.

13. K. He, X. Zhang, S. Ren and J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition, in *Computer Vision — ECCV 2014* (Springer International Publishing, 2014), pp. 346–361.
14. K. Hornik, M. Stinchcombe and H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* **2**(5) (1989) 359–366.
15. H.-T. Hu and T.-T. Lee, Frame-synchronized blind speech watermarking via improved adaptive mean modulation and perceptual-based additive modulation in DWT domain, *Digit. Signal Process.* **87** (2019) 75–85.
16. G. Hua, J. Goh and V. L. L. Thing, Time-spread echo-based audio watermarking with optimized imperceptibility and robustness, *IEEE/ACM Trans. Audio Speech Lang. Process.* **23**(2) (2015) 227–239.
17. Y. Huang, B. Niu, H. Guan and S. Zhang, Enhancing image watermarking with adaptive embedding parameter and PSNR guarantee, *IEEE Trans. Multimed.* **21**(10) (2019) 2447–2460.
18. M.-J. Hwang, J.-S. Lee, M.-S. Lee and H.-G. Kang, SVD-based adaptive QIM watermarking on stereo audio signals, *IEEE Trans. Multimed.* **20**(1) (2018) 45–54.
19. A. Jadda and I. S. Prabha, Speech enhancement via adaptive wiener filtering and optimized deep learning framework, *Int. J. Wavelets Multiresolut. Inf. Process.* **21** (2023) 2250032.
20. W. Jiang, X. Huang and Y. Quan, Audio watermarking algorithm against synchronization attacks using global characteristics and adaptive frame division, *Signal Process.* **162** (2019) 153–160.
21. H. Kandi, D. Mishra and S. R. S. Gorthi, Exploring the learning capabilities of convolutional neural networks for robust image watermarking, *Comput. Secur.* **65** (2017) 247–268.
22. D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, *CoRR*, preprint (2014), arXiv:abs/1412.6980.
23. V. Kurková, Kolmogorov's theorem and multilayer neural networks, *Neural Netw.* **5**(3) (1992) 501–506.
24. H. Y. Leung, L. M. Cheng and L. L. Cheng, Robust watermarking schemes using selective curvelet coefficients based on a HVS model, *Int. J. Wavelets Multiresolut. Inf. Process.* **08** (2010) 941–959.
25. S. Li, X. Kang, L. Fang, J. Hu and H. Yin, Pixel-level image fusion: A survey of the state of the art, *Inf. Fusion* **33** (2017) 100–112.
26. X. Liang and S. Xiang, Robust reversible audio watermarking based on high-order difference statistics, *Signal Process.* **173** (2020) 107584.
27. Z. Liu, Y. Huang and J. Huang, Patchwork-based audio watermarking robust against de-synchronization and recapturing attacks, *IEEE Trans. Inf. Forensics Sec.* **14**(5) (2019) 1171–1180.
28. Y. Liu, C. Xia, X. Zhu and S. Xu, Two-stage copy-move forgery detection with self deep matching and proposal superglue, *IEEE Trans. Image Process.* **31** (2021) 541–555.
29. N. Lopac, F. Hržić, I. P. Vuksanovic and J. Lerga, Detection of non-stationary GW signals in high noise from Cohen's class of time–frequency representations using deep learning, *IEEE Access* **10** (2022) 2408–2428.
30. S. P. Maity and M. K. Kundu, Performance improvement in spread spectrum image watermarking using wavelets, *Int. J. Wavelets Multiresolut. Inf. Process.* **09** (2011) 1–33.

31. D. Malah, Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals, *IEEE Trans. Acoust. Speech Signal Process.* **27**(2) (1979) 121–133.
32. S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7) (1989) 674–693.
33. S.-M. Mun, S.-H. Nam, H. Jang, D. Kim and H.-K. Lee, Finding robust domain from attacks: A learning framework for blind watermarking, *Neurocomputing* **337** (2019) 191–202.
34. I. Natgunanathan, Y. Xiang, G. Hua, G. Beliakov and J. Yearwood, Patchwork-based multilayer audio watermarking, *IEEE/ACM Trans. Audio Speech Lang. Process.* **25**(11) (2017) 2176–2187.
35. Y. Nesterov, A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$, *Dokl. Akad. Nauk SSSR* **269** (1983) 543–547.
36. A. Oppenheim, *Discrete-time Signal Processing* (Prentice Hall, Upper Saddle River, NJ, 1999).
37. K. Pavlović, S. Kovačević, I. Djurović and A. Wojciechowski, Robust speech watermarking by a jointly trained embedder and detector using a DNN, *Digit. Signal Process.* **122** (2021) 103381.
38. H. Peng, B. Li, X. Luo, J. Wang and Z. Zhang, A learning-based audio watermarking scheme using kernel fisher discriminant analysis, *Digit. Signal Process.* **23**(1) (2013) 382–389.
39. L. Rabiner and R. Schafer, *Digital Processing of Speech Signals* (Prentice Hall, Englewood Cliffs, NJ, 1978).
40. A. W. Rix, J. G. Beerends, M. P. Hollier and A. P. Hekstra, Perceptual evaluation of speech quality (PESQ) — a new method for speech quality assessment of telephone networks and codecs, in *2001 IEEE Int. Conf. Acoustics, Speech, and Signal Processing Proc.* (*Cat. No. 01CH37221*), Vol. 2, Salt Lake City, Utah, USA, 2001, pp. 749–752.
41. O. Ronneberger, P. Fischer and T. Brox, U-net: Convolutional networks for biomedical image segmentation, in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, Lecture Notes in Computer Science, Vol. 9351 (Singapore, 2015), pp. 234–241.
42. S. Stankovic, I. Djurovic and I. Pitas, Watermarking in the space/spatial-frequency domain using two-dimensional Radon–Wigner distribution, *IEEE Trans. Image Process.* **10**(4) (2001) 650–658.
43. M. Steinebach, F. A. P. Petitcolas, F. Raynal, J. Dittmann, C. Fontaine, C. Seibel, N. Fates and L. Ferri, StirMark benchmark: Audio watermarking attacks, in *Proc. Int. Conf. Information Technology: Coding and Computing*, Las Vegas, Nevada, 2001, pp. 49–54.
44. S. Wang, W. Yuan and M. Unoki, Multi-subspace echo hiding based on time-frequency similarities of audio signals, *IEEE/ACM Trans. Audio Speech Lang. Process.* **28** (2020) 2349–2363.
45. Y. Xiang, I. Natgunanathan, D. Peng, G. Hua and B. Liu, Spread spectrum audio watermarking using multiple orthogonal PN sequences and variable embedding strengths and polarities, *IEEE/ACM Trans. Audio Speech Lang. Process.* **26**(3) (2018) 529–539.
46. D. Xiao, A. Zhao and F. Li, Robust watermarking scheme for encrypted images based on scrambling and Kronecker compressed sensing, *IEEE Signal Process. Lett.* **29** (2022) 484–488.

47. J. Zhao, T. Zong, Y. Xiang, L. Gao, W. Zhou and G. Beliakov, Desynchronization attacks resilient watermarking method based on frequency singular value coefficient modification, *IEEE/ACM Trans. Audio Speech Lang. Process.* **29** (2021) 2282–2295.
48. X. Zheng, Y. Tang and J. Zhou, A framework of adaptive multiscale wavelet decomposition for signals on undirected graphs, *IEEE Trans. Signal Process.* **67**(7) (2019) 1696–1711.
49. D.-X. Zhou, Universality of deep convolutional neural networks, *Appl. Comput. Harmon. Anal.* **48**(2) (2020) 787–794.
50. J. Zhu, R. Kaplan, J. Johnson and L. Fei-Fei, Hidden: Hiding data with deep networks, in *15th Eur. Conf.*, Munich, Germany, 2018, pp. 682–697.
51. T. Zong, Y. Xiang, I. Natgunanathan, S. Guo, W. Zhou and G. Beliakov, Robust histogram shape-based method for image watermarking, *IEEE Trans. Circuits Syst. Video Technol.* **25**(5) (2015) 717–729.

УНИВЕРЗИТЕТ ЦРНЕ ГОРЕ

Ул. Цетињска бр. 2
П. фах 99
81000 ПОДГОРИЦА
ЦРНА ГОРА
Телефон: (020) 414-255
Факс: (020) 414-230
E-mail: rektor@ac.me

UNIVERSITY OF MONTENEGRO

Ul. Cetinjska br. 2
P.O. BOX 99
81 000 PODGORICA
MONTENEGRO
Phone: (+382) 20 414-255
Fax: (+382) 20 414-230
E-mail: rektor@ac.me

Број: 08- 619
Датум, 10. 04. 2014 г.

Ref:
Date,

УНИВЕРЗИТЕТ ЦРНЕ ГОРЕ
Природно-математички факултет
Број 311
Подгорица, 11. 04. 2014 год

Na osnovu člana 75 stav 2 Zakona o visokom obrazovanju (Sl.list RCG, br. 60/03 i Sl.list CG, br. 45/10 i 47/11) i člana 18 stav 1 tačka 3 Statuta Univerziteta Crne Gore, Senat Univerziteta Crne Gore, na sjednici održanoj 10.aprila.2014. godine, donio je

ODLUKU
O IZBORU U ZVANJE

Dr MILENKO MOSUROVIĆ bira se u akademsko zvanje **redovni profesor** Univerziteta Crne Gore za predmete: Strukture podataka, Teorija složenosti algoritama i Paralelni algoritmi, **na Prirodno-matematičkom fakultetu**.

REKTOR

Prof.dr Predrag Miranović

Milenko Mosurović (selektovana bibliografija)

1. **M. Mosurovic**, N. Krdzavac, H. Graves and M. Zakharyaschev. A Decidable Extension of SROIQ with Complex Role Chains and Unions, Journal of Artificial Intelligence Research (JAIR) 47 (2013), 809-851.

2. **Milenko Mosurovic**, Michael Zakharzaschev. On the Complexity of (Restricted) ALCIr, Publications de l'Institut Mathematique, Vol. 95[109], pp. 133–147 (2014).

3. **Milenko Mosurović**. Disjunction in modal description logics, Publications de l'Institut Mathématique, Vol. 82(96), pp. 25–35 (2007)

4. **M. Mosurovic**, F. Wolter and M. Zakharyaschev, Temporal description logics: a point-based approach, AI*IA Notizie, Anno XIV, no. 1, Marzo 2001, pp. 16-20.

5. **Milenko Mosurovic**, Tatjana Stojanovic, Ana Kaplarevic-Mališic. Reasoning in basic description logics and description logics with modal operators. Logic in computer science, Beograd: Matematički Institut SANU. Zbornik Radova (Beograd) 12(20), 113-158 (2009).

6. Artale, E. Franconi, **M. Mosurovic**, F. Wolter and M. Zakharyaschev, The DRL(US) temporal description logic, DL2001. Published on CEUR-WS, Vol. 49, pp. 96-105.

7. Artale, E. Franconi, **M. Mosurovic**, F. Wolter and M. Zakharyaschev, Reasoning over Conceptual Schemas and Queries in Temporal Databases, In Proceedings of the 9th Italian Symposium on Advanced Database Systems, Venice, Italy, June 27-29, 2001.

8. **M. Mosurovic** and M. Zakharyaschev, On the complexity of description logics with modal operators, In P. Kolaitos and G. Koletos, editors, Proceedings of the 2nd Panhellenic Logic Symposium, pp. 166-171, Delphi, 1999.

9. **M. Mosurovic**, F. Wolter, M. Zakharyaschev, Complexity of temporalized ALC, In Proceedings of the 4th International Conference on Discrete Mathematics in Control System Theory, Krasnovidovo (June 19-25, 2000), Moscow, MAX Press, 2000, pp. 141-144.

10. **M. Mosurovic**, Henson Graves, Nenad Krdzavac: A Decidable Extension of SRIQ with Disjunctions in Complex Role Inclusion Axioms. DL2012, Published on CEUR-WS, Vol-846, p. 497-507

BIOGRAFIJA

IME I PREZIME: Milenko Mosurović

Rođen sam 1968. godine u Prijepolju, Republika Srbija, gdje sam završio osnovnu i srednju školu. Za postignute rezultate iz matematike dobitnik sam diplome "Mihailo Petrović Alas". Diplomirao sam 1992. godine, na PMF-u u Podgorici sa prosječnom ocjenom 9,71. Kao najbolji student završne godine studija PMF-a, dobio sam nagradu Univerziteta Crne Gore.

Postdiplomske studije matematike na Matematičkom fakultetu (Katedra za računarstvo) u Beogradu završio sam 1996. godine odbranom magistarskog rada "Paralelizacija algoritama za množenje velikih brojeva i njihova primena na ispitivanje primalnosti Fermaovih brojeva". Doktorsku disertaciju "Složenost opisnih logika sa modalnim operatorima" odbranio sam 2001. godine na Matematičkom fakultetu Univerziteta u Beogradu.

Koautor sam dvije univerzitetske zbirke zadataka.

Tokom ljetnjeg semestra školske 1994/95. godine boravio sam na naučnom usavršavanju na Matematičkom institutu SANU u Beogradu pod rukovodstvom prof. dr Žarka Mijajlovića.

U periodu od 1997. do 2000. godine, boravio sam na naučnom usavršavanju u Rusiji, na Moskovskom državnom univerzitetu "M. V. Lomonosov", na Fakultetu za primijenjenu matematiku i kibernetiku. Moje naučno usavršavanje se odvija na Katedri za matematičku kibernetiku pod rukovodstvom prof. dr Mihaila V. Zaharjaševa (Michael Zakharyaschev).

Od diplomiranja 1992. godine do danas radim na PMF-u u Podgorici. U zvanje redovnog profesora izabran sam u aprilu 2014. godine.

УНИВЕРЗИТЕТ ЦРНЕ ГОРЕ
Ул. Цетињска бр. 2
П. фах 99
81000 ПОДГОРИЦА
ЦРНА ГОРА
Телефон: (020) 414-255
Факс: (020) 414-230
E-mail: rektor@ac.me

UNIVERSITY OF MONTENEGRO
Ul. Cetinjska br. 2
P.O. BOX 99
81 000 PODGORICA
MONTENEGRO
Phone: (+382) 20 414-255
Fax: (+382) 20 414-230
E-mail: rektor@ac.me

Број: 08-824
Датум, 02.06.2011 г.

Ref:
Date,

Na osnovu člana 75 stav 2 Zakona o visokom obrazovanju (Sl.list RCG, br. 60/03 i Sl.list CG, br. 45/10) i člana 18 stav 1 tačka 3 Statuta Univerziteta Crne Gore, Senat Univerziteta Crne Gore, na sjednici održanoj 02.06.2011. godine, donio je

ODLUKU
O IZBORU U ZVANJE

Dr IGOR ĐUROVIĆ bira se u akademsko zvanje **redovni profesor** Univerziteta Crne Gore za predmete: Programiranje I (osnovne studije, ETF), Programiranje II (osnovne studije, ETR) i Teorija informacija i kodova (osnovne studije, ETR) na **Elektrotehničkom fakultetu**.

УНИВЕРЗИТЕТ ЦРНЕ ГОРЕ
ЕЛЕКТРОТЕХНИЧКИ ФАКУЛТЕТ
Број 02/2-764
Подгорица, 09.06. 2011 год.

REKTOR
Miranović Predrag
Prof.dr Predrag Miranović

Igor Đurović je rođen 29. 08. 1971. u Cetinju. Osnovnu i srednju školu prirodno-matematičkog smjera završio je u Herceg Novom. Dobitnik je više priznanja na republičkim takmičenjima učenika srednjih škola iz matematike. Diplomirao je na smjeru Elektronika na Elektrotehničkom fakultetu u Podgorici, 1994. godine. Na istom fakultetu je magistrirao („Funkcija jezgra u vremensko-frekvencijskoj analizi i softverski paket za realizaciju distribucija") i doktorirao („Vremensko-frekvencijske reprezentacije u estimaciji parametara signala sa primjenom u digitalnom watermarking-u"), 1996. i 2000. respektivno. U zvanja docenta, vanrednog profesora i redovnog profesora biran je 2001, 2006. i 2011. godine na Elektrotehničkom fakultetu. Bio je šef Katedre za računare, rukovodilac postdiplomskih studija, rukovodilac doktorskih studija na Fakultetu, član Senata Univerziteta Crne Gore (2011–2014) i član Strukovnog vijeća za prirodne i tehničke nauke Univerziteta. Predavao je na više drugih jedinica Univerziteta Crne Gore a bio je i gostujući nastavnik na Fakultetu za proizvodnju i menadžment, Trebinje, Univerziteta u Istočnom Sarajevu, BiH.

Autor i koautor više univerzitetskih udžbenika kao i više skripata, praktikuma itd. Bio je mentor na više doktorskih disertacija i magistarskih radova.

Autor je oko 200 radova, od toga preko 100 u vodećim međunarodnim časopisima iz obrade signala i srodnih oblasti. Editor je jedne monografije publikovane u našoj zemlji i autor radova u dvije monografije. Autor je 6 poglavlja u naučnim monografijama izdatim od renomiranih međunarodnih izdavača i koautor jedne knjige publikovane u Njemačkoj. Radovi su vezani za više tema iz obrade signala sa primjenama: estimacija parametara signala; vremensko-frekvencijska analiza sa primjenama; multimedijalni signali i sistemi; obrada signala u telekomunikacijama; primjena u električnim kolima i električnim mjerenjima itd.

Do sada su ovi radovi citirani više od 2200 puta u okviru ISI Web of Knowledge, od kojih samo radovi iz oblasti digitalnog watermarking-a oko 300 puta. Rad dr Igora Đurovića je ostao veoma zapažen pa je recenzent sa više stotina recenzija u više od 40 vodećih međunarodnih časopisa. Pored ovoga, bio je recenzent i više domaćih i regionalnih časopisa. Član editorijalnog odbora ili pridruženi editor bio je u više uglednih međunarodnih časopisa, od kojih se izdvaja jedan od najuglednijih časopisa iz oblasti obrade signala Elsevier „Signal Processing". Bio je član editorijalnog odbora Journal of Electrical and Computer Engineering, Hindawi kao i časopisa Research Letters in Signal Processing, Hindawi. Bio je vodeći gostujući editor za Eurasip (Evropsko udruženje za obradu signala) Journal on Advances in Signal Proceesing za specijalni broj „Robust processing of non-stationary signals". Član je editorijalnih bordova, recenzent i član programskih komiteta nekoliko međunarodnih i regionalnih konferencija. Pored toga, veći broj radova je proglašavan najboljim u sekcijama na domaćim konferencijama. Senior Member IEEE (vodeće svjetsko udruženje inženjera elektrotehnike i elektronike) je od 2006. godine.

Bio je rukovodilac lokalnih timova, partner i učesnik na većem broju nacionalnih, bilateralnih i međunarodnih projekata finansiranih od strane Volkswagen stuftung, FP 7, Tempus, CNRS, JSPS, DoD Canada, WUS Austria, Ministarstva nauke Crne Gore itd. Osmislio je i bio prvi direktor prvog domaćeg Centra izvrsnosti BIO-ICT u periodu 2014–2015.

U periodu od novembra 2001. do novembra 2002. boravio je kao stipendista Japanskog društva za unapređenje nauke (JSPS) na Kyoto Institute of Technology. Bio je na kraćim boravcima na inostranim univerzitetima i to: Univerzitet Aristotel Solun, Aritifical Inteligence and Image Analysis Laboratory (Laboratorija za vještačku inteligenciju i analizu slike), Grčka, veći broj univerziteta u SAD u okviru International Visitors Program, Ruhr University Bochum, Signal Theory Group, Njemačka u okviru Volkswagen stiftung projekta, ENSIETA, Brest, Francuska, u okviru PAI Pelikan projekta, GIPSA lab, INP Grenoble, Francuska u okviru CNRS projekta, Nacionalni aerokosmički univerzitet, Kharkov, Ukrajina, Tampere univerzitet za tehnologiju, Tampere, Finska, itd. Bio je član Komisije za odbranu doktorske disertacije na Department for Mathematical Statistics, Lund University, Lund, Švedska.

Dobitnik je Nagrade Crnogorske akademije nauka i umjetnosti iz Fonda Petra Vukčevića 2002. godine i Trinaestojulske nagrade za 2016. godinu. Član je i prvi predsjednik Centra mladih naučnika CANU. Organizovao je više skupova u okviru CANU: „Mobilne i bežične komunikacije: stanje i perspektive" 2009. godine, „Visoko obrazovanje u Crnoj Gori: stanje i perspektive", u organizaciji Centra mladih naučnika CANU, sa izlaganjem „Nauka u Crnoj Gori" i skup „Važnost GEO inicijativa i crnogorski kapaciteti u ovim oblastima" 2011. godine. Trenutno rukovodi Komisijom CANU za naučne i umjetničke djelatnosti i Odborom CANU za informaciono-komunikacione tehnologije. Organizator je skupa o trendovima u savremenim mobilnim komunikacionim sistemima, koji je održan u CANU 2014. godine.

Za vanrednog člana CANU izabran je 29. novembra 2011. godine, a za redovnog 18. decembra 2018. godine.

1. Quasi-maximum-Likelihood Estimator of PPS on the Uniform Linear Array - I Djurović, M Simeunović, P Raković, Circuits, Systems, and Signal Processing 38 (10), 4874-4889, 2019.
2. A modified Viterbi algorithm-based if estimation algorithm for adaptive directional time–frequency distributions - NA Khan, M Mohammadi, I Djurović, Circuits, Systems, and Signal Processing 38 (5), 2227-2244, 2019.
3. Parametric estimation of 2D cubic phase signals using high-order Wigner distribution with genetic algorithm - M Simeunović, I Djurović, A Pelinković, Multidimensional Systems and Signal Processing 30 (1), 451-464, 2019.
4. A method for efficient maximization of PPS estimation functions - M Simeunović, I Djurović, Digital Signal Processing 84, 38-45, 2019.
5. Combined Centre-Weighted Median Filter and BM3D to Filter Digital Images in Mixed Gaussian and Impulsive Environments - I Djurović, IETE Journal of Research 64 (6), 796-806, 2018.
6. Hardware implementation of the quasi-maximum likelihood estimator core for polynomial phase signals - NR Brnović, I Djurović, VN Ivanović, M Simeunović, IET Circuits, Devices & Systems 13 (2), 131-138, 2018.
7. Parameter estimation of coupled polynomial phase and sinusoidal FM signals - I Djurović, P Wang, M Simeunović, PV Orlik, Signal Processing 149, 1-13, 2018.
8. Parameter estimation of 2D polynomial phase signals using NU sampling and 2D CPF - I Djurović, M Simeunović, IET Signal Processing 12 (9), 1140-1145, 2018.
9. Combination of the Viterbi algorithm and cross-Wigner distribution for the instantaneous frequency estimation phase signals in high noise environments - I Djurović, Journal of Electrical Engineering 69 (3), 255-258, 2018.
10. QML-RANSAC instantaneous frequency estimator for overlapping multicomponent signals in the time-frequency plane - I Djurović, IEEE Signal Processing Letters 25 (3), 447-451, 2018.
11. Review of the quasi-maximum likelihood estimator for polynomial phase signals - I Djurović, M Simeunović, Digital Signal Processing 72, 59-74, 2018.
12. Estimation of sinusoidal frequency-modulated signal parameters in high-noise environment - I Djurović, Signal, Image and Video Processing 11 (8), 1537-1541, 2017.
13. Post-processing of time-frequency representations in instantaneous frequency estimation based on ant colony optimization - M Brajović, V Popović-Bugarin, I Djurović, S Djukanović, Signal Processing 138, 195-210, 2017.
14. Cubic phase function: A simple solution to polynomial phase signal analysis - I Djurović, M Simeunović, P Wang, Signal Processing 135, 48-66, 2017.
15. On improvement of joint estimation of DOA and PPS coefficients impinging on ULA - P Raković, M Simeunović, I Djurović, Signal Processing 134, 209-213
16. Combination of the adaptive Kuwahara and BM3D filters for filtering mixed Gaussian and impulsive noise - I Djurović, Signal, Image and Video Processing 11 (4), 753-760, 2017.
17. The STFT-based estimator of micro-Doppler parameters - I Djurović, V Popović-Bugarin, M Simeunović, IEEE Transactions on Aerospace and Electronic Systems 53 (3), 1273-1283, 2017.
18. QML-RANSAC: PPS and FM signals estimation in heavy noise environments - I Djurović, Signal Processing 130, 142-151, 2017.

19. BM3D filter in salt-and-pepper noise removal - I Djurović, EURASIP Journal on Image and Video Processing 2016 (1), 13, 2016
20. Resolving aliasing effect in the QML estimation of PPSs - I Djurovic, M Simeunovic, IEEE Transactions on Aerospace and Electronic Systems 52 (3), 1494-1499, 2016.
21. Time–frequency feature representation using energy concentration: An overview of recent advances - E Sejdić, I Djurović, J Jiang, Digital signal processing 19 (1), 153-183, 2009.
22. Fractional Fourier transform as a signal processing tool: An overview of recent developments - E Sejdić, I Djurović, LJ Stanković, Signal Processing 91 (6), 1351-1369, 2011.
23. Digital watermarking in the fractional Fourier transformation domain - I Djurovic, S Stankovic, I Pitas, Journal of Network and Computer Applications 24 (2), 167-173, 2001.
24. Watermarking in the space/spatial-frequency domain using two-dimensional Radon-Wigner distribution - S Stankovic, I Djurovic, I Pitas, IEEE transactions on image processing 10 (4), 650-658, 2001.
25. Separation of target rigid body and micro-Doppler effects in ISAR imaging - L Stankovic, I Djurovic, T Thayaparan, IEEE Transactions on Aerospace and Electronic Systems 42 (4), 1496-1506, 2006.
26. Micro-Doppler-based target detection and feature extraction in indoor and outdoor environments - T Thayaparan, LJ Stanković, I Djurović, Journal of the Franklin Institute 345 (6), 700-722, 2008.
27. An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment - I Djurović, LJ Stanković, Signal Processing 84 (3), 631-643, 2004.
28. Integrated cubic phase function for linear FM signal analysis - P Wang, H Li, I Djurovic, B Himed, IEEE Transactions on Aerospace and Electronic Systems 46 (3), 963-977, 2010.
29. Frequency-based window width optimization for S-transform - I Djurović, E Sejdić, J Jiang, AEU-International Journal of Electronics and Communications 62 (4), 245-250, 2008.
30. Robust L-estimation based forms of signal transforms and time-frequency representations - I Djurovic, LJ Stankovic, JF Bohme, IEEE Transactions on Signal Processing 51 (7), 1753-1761, 2003.

УНИВЕРЗИТЕТ У БЕОГРАДУ
ЕЛЕКТРОТЕХНИЧКИ
ФАКУЛТЕТ

Булевар краља Александра 73
П.Ф. 35-54, 11120 Београд, Србија
Тел: +381 11 3248464
Факс: +381 11 3248681

ТР: 840-0000001438666-48
ПИБ: 100206130
МБ: 07032498
ЈБКЈС: 02239

Број: 203

Датум: 07.02.2024. године

На лични захтев именованог, а увидом у службену евиденцију, Универзитет у Београду – Електротехнички факултет издаје следећу

ПОТВРДУ

Др **Горан (Стеван) Квашчев**, рођен 12.07.1975. године у Кикинди, Република Србија, ЈМБГ: 1207975840015, налази се у радном односу са пуним радним временом на Универзитету у Београду – Електротехничком факултету на радном месту „ванредни професор".

Именовани је први пут изабран у наведено звање ванредног професора 04.02.2018. године, у коме се и даље налази.

Ова потврда се издаје на лични захтев именованог.

У Београду,

07.02.2024. године

Секретар ЕТФ-а

Милош Дивјак

Горан Квашчев – Биографија и библиографија

А. Биографски подаци

Др Горан С. Квашчев је рођен у Кикинди, 12 јула 1975. године, где је завршио основну школу и гимназију. Електротехнички факултет у Београду уписао је 1994. године. Дипломирао је 05.09.2000. године са темом "Примена ПЛЦ контролера у реализацији аутоматског подешавања ПИ регулатора из одскочног одзива" оценом 10. У току студирања остварио је просечну оцену 8.86.

По дипломирању је запослен на Електротехничком факултету, Универзитета у Београду, у Београду, Катедра за аутоматику, где активно учествује у настави, као и раду на пројектима. Постдипломске студије, смер Управљање системима на Електротехничком факултету у Београду, уписује 2000. године. Магистрирао је 2005. године одбраном тезе "Даљи развој и упоредна анализа процедура за експериментално пројектовање и подешавање индустријских регулатора". Докторску дисертацију по насловом „Робусна идентификација индустријских процеса" одбранио је 9.3.2012. године на Електротехничком факултету у Београд, а 17.9.2012. је промовисан у доктора електротехничких наука од стране Универзитета у Београду.

Горан Квашчев је објавио (15) петнаест радова у водећим међународним часописима са impact factor-ом, 42 радова на међународним конференцијама, 6 у домаћим часописима и 30 на домаћим конференцијама. Учествовао је на 26 пројеката: *TEMPUS*, *FP7*, *EUREKA* и WUS пројеката, као и на више пројеката финансираних од Министарства. Кандидат је ангажован на следећим предметима основних, мастер и докторских студија: Системи аутоматског управљања, Моделирање и идентификација, Управљање у реалном времену, Управљање индустријским процесима, Неуралне мреже, Практикум из софтверских алата, Управљање сложеним системима, Класификација и естимација сигнала... Учествовао је у реализацији већег броја међународних и националних иновационих, истраживачких, развојних и мултидисциплинарних пројеката. Горан Квашчев је члан националног друштва ЕТРАН и међународне организације IEEE, као и Инжењерске коморе Србије.

Област истраживања Горана Квашчева обухвата пројектовање, пуштање у рад и оптимизација система управљања и регулације за велике термоенергетске и индустријске објекте и постројења, моделирање и идентификацију процеса, детекција и изолација отказа.

Б. Дисертације

Б.1. **Г. С. Квашчев**, Даљи развој и упоредна анализа процедура за експериментално пројектовање и подешавање индустријских регулатора, Магистарска теза, Универзитет у Београду - Електротехнички факултет, Београд, Србија, 2005.

Б.2. **Г. С. Квашчев**, Робусна идентификација индустријских процеса, Докторска дисертација, Универзитет у Београду - Електротехнички факултет, Београд, Србија, 2012.

В. Наставна активност

Горан Квашчев је, као предметни наставник, тренутно ангажован на следећим предметима дипломских, мастер и докторских студија Електротехничког факултета у Београду:

- Управљање у реалном времену, обавезни за студенте ОС
- Неуралне мреже, изборни за студенте ОС
- Управљање индустријским процесима, изборни са студенте СИ
- Неуралне мреже и системи за обраду сигнала, изборни са мастер студенте ОС
- Системи аутоматског управљања, обавезни са студенте ОГ
- Управљање сложеним индустријским процесима, изборни са мастер студенте ОС
- Управљање сложеним системима, изборни за докторске студенте УСОС
- Неуралне мреже, изборни за докторске студенте УСОС
- Класификација и естимација сигнала, изборни за докторске студенте УСОС

У оцењивањима од стране студената, добијао је високе оцене Просечна оцена: 4,36 и Просечна оцена на предметима са 10 и више анкетираних студената: 4,37.

Од избора у наставничко звање, Горан Квашчев је руководио израдом: 104 завршна рада (студије 4 и 5 год), 52 завршних - мастер радова, 1 докторском дисертацијом. Учествовао је комисијама за одбрану радова и то: 219 завршних радова (студије 4 и 5 год), 103 завршних - мастер радова, као и комисијама за оцену и за усмену одбрану 10 докторских дисертација на Електротехничком факултету у Београду.

Кандидат је био члан 3 комисије за избор у звање за сарадника у настави на Универзитету у Београду – ЕТФ-у, 2015-2022.

Горан Квашчев је коаутор универзитетског уџбеника:
Б. Д. Ковачевић, **Г. С. Квашчев**, *Идентификација процеса*, Univerzitet u Beogradu – Elektrotehnički fakultet, Akademska misao, ISBN 978-86-7466-732-3, 2017.

Горан Квашчев је аутор помоћног наставног уџбеника: „ПИД регулација - анализа и синтеза", Универзитет у Београду – Електротехнички факултет, ISBN 978-86-7225-096-1, 2023.

Г. Библиографија научних и стручних радова

Горан Квашчев је аутор или коаутор 15 (петнаест) радова у међународним научним часописима са impact factor-ом, 42 (четрдесет два) рада на међународним конференцијама, 30 (тридесет) радова на домаћим конференцијама, као 11 (једанаест) техничких решења. Списак радова, категорисан према Правилнику о поступку и начину вредновања, и квантитативном исказивању научноистраживачких резултата истраживача, дат је у наставку.

Категорија М20 - Радови објављени у научним часописима међународног значаја

M20-1. M. R. Mataušek, **G. S. Kvaščev**, "A unified step response procedure for autotuning of PI controller and Smith predictor for stable processes", Journal of Process Control, ISSN:0959-1524, Volume 13, Pages 787-800, December 2003. (M21), IF:1.248

M20-2. **Kvascev, G.S.**, Djurovic, Z.M., Kovacevic, B.D., "Adaptive recursive M-robust system parameter identification using the QQ-plot approach", Control Theory & Applications, IET, ISSN:1350-2379, Vol. 5 Issue 4, pp. 579 – 593, DOI: 10.1049/iet-cta.2009.0647, 2011. (M22), IF:0.990

Категорија М30 - Зборници међународних научних скупова
(сви радови су поткатегорије М33 - Саопштење са међународног скупа штампано у целини)

M30-1. Mladen Majstorović, Ivan Nikolić, Jelena Radović, Goran Kvaščev, *"Neural Network Control Approach for a Two-Tank System"*, NEUREL 2008. 9th Symposium on, Belgrade, DOI:10.1109/NEUREL.2008.4685619

M30-2. Jorgovanovic, M. , Pajic, M., Kvascev, G., Popovic, J., "FPGA design of arbitrary down-sampler", 26th international conference on microelectronics proceedings, Vol.1, pp.391-394, 2008

M30-3. Goran S. Kvascev, Zeljko M. Djurovic, Veljko D. Papic: "One approach to fault detection in steam temperature control system", Conference DECOM-IFAC-09, Ohrid

M30-4. Goran Kvascev, Predrag Tadic, Ruben Puche Panadero, Predrag Todorov, "Thermal Power Plant Fan Drive Load Distribution Control", IFAC Conference on Control Methodologies and Technology for Energy Efficiency, CMTEE 2010, Vilamoura, Portugal, March 29-31, 2010

M30-5. P. Tadic, Z. Djurovic, G. Kvascev, V. Papic, "Coal-shortage detection in power plants by means of fixed size sample strategy", IFAC Conf. on Control Methodologies & Tech. for Enegry-Efficency, Vilamoura, Portugal, March 2010

M30-6. Papic, V. Djurovic, Z. Kvascev, G. Tadic, P., "On signal-to-noise ratio estimation", MELECON 2010, 15th IEEE Mediterranean Electrotechnical Conference, Valletta, Malta, 26-28 April 2010 DOI:10.1109/MELCON.2010.5476314

M30-7. Goran Kvascev, Predrag Tadic, Zeljko Djurovic, "An Application of Model Based Fault Detection in Power Plants", Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis, pp. 130-134, 18-19 November, 2010, Ferrara, Italy

M30-8. Nebojša Malešević, Lana Popović, Goran Bijelić and Goran Kvaščev, "Muscle twitch responses for shaping the multi-pad electrode for functional electrical stimulation", NEUREL 2010. 10th Symposium on, Belgrade

M30-9. Aleksandra Lj. Marjanović, Željko M. Đurović, Goran S. Kvaščev, Predrag R. Tadić, "Fault Detection and Isolation in Steam Separator System Using Hidden Markov Models", 9th European Workshop on Advanced Control and Diagnosis (ACD 2011), Hungary

M30-10. Goran S. Kvascev, Aleksandra Lj. Marjanovic, Predrag R. Tadic, Zeljko M. Djurovic, *"To Robust Identification of Water Steam-separator Process in Thermal Power Plants"*, IEEE International Conference on Industrial Technology (ICIT2012), Athens, 2012

M30-11. Predrag Milosavljević, Nenad Baščarević, Kosta Jovanović, Goran Kvaščev, *„Neural Networks in Feedforward Control of a Robot Arm Driven by Antagonistically Coupled Drives"*, 11th Symposium on Neural Network Applications in Electrical Engineering (NEUREL 2012), 2012, Belgrade, Serbia

M30-12. Goran Kvaščev, Maja Gajić-Kvaščev, Željko Đurović, *„Radial Basis Function Network Based Feature Extraction for Improvement the Procedure of Sourcing Neolithic Ceramics"*, 11th Symposium on Neural Network Applications in Electrical Engineering (NEUREL 2012), 2012, Belgrade, Serbia

M30-13. Marjanovic Aleksandra, Kvascev Goran, Djurovic Zeljko, *„Comparison of Identification Procedures in the Frame of Fault Detection and Isolation",2012 IEEE Multi-conference on Systems and Control, October 3-5, 2012, Dubrovnik, Croatia*

„Non-destructive characterisation and classification of ceramic artefacts using pEDXRF and statistical pattern recognition", Chemistry Central Journal, ISSN:1752-153X, 6:102, doi:10.1186/1752-153X-6-102, 2012 (M21), IF:3.281

M20-4. Nasar Aldian Ambark Shashoa, **Goran Kvascev**, Aleksandra Marjanovic, Zeljko Djurovic, Sensor Fault Detection and Isolation in a Thermal Power Plant Steam Separator, Control Engineering Practice vol. 21, issue 7, (2013), ISSN:0967-0661, pp. 908–916, DOI:10.1016/j.conengprac.2013.02.012 (M21), IF: 1.912

M20-5. Y. Abuadlla, S. Gajin, **G. Kvaščev**, Z. Jovanović, Flow-Based Anomaly Intrusion Detection System Using Two Stages Neural Network, Computer Science and Information Systems, COMPUTER SCIENCE AND INFORMATION SYSTEMS, ISSN:1820-0214, pp. 1-10, 2014, (M23), IF:0.477

M20-6. Antić S., Djurovic Z., and **Kvascev G.**, Application of Structured and Directional Residuals for Fault Detection and Isolation on Permanent-Magnet DC Motor with Amplifier, Qual. Reliab. Eng. Int., 2016, (DOI: 10.1002/qre.1962), ISSN:0748-8017, (M22), IF: 1.366

M20-7. S. Vujnović, Ž. Đurović, **G. Kvaščev**, Fan mill state estimation based on acoustic signature analysis, CONTROL ENGINEERING PRACTICE, ISSN:0967-0661, Vol. 57, pp. 29-38, 2016. (M21), IF: 2.602

У последњем петогодишњем периоду

M20-8. V. Bobić, M. Djurić-Jovičić, N. Dragašević, M. B. Popović, V. S. Kostić, **G. Kvaščev**, An Expert System for Quantification of Bradykinesia Based on Wearable Inertial Sensors, SENSORS, ISSN:1424-8220, Vol. 19, No. 11, pp. 1-17, Jun, 2019. DOI:10.3390/s19112644 (M21), IF:3.275

M20-9. D. Pavlovic, M. Czerkawski, C. Davison, O. Marko, C. Michie, R. Atkinson, V. Crnojevic, I. Andonovic, B. Рајовић, **G. Kvaščev,** C. Tachtatzis, Behavioural Classification of Cattle Using Neck-Mounted Accelerometer-Equipped Collars, SENSORS, ISSN:1424-8220, Vol. 22, No. 6, pp. 1-18, Mar, 2022. DOI:10.3390/s22062323 (M21), IF:3.576

M20-10. **G. Kvaščev**, Ž. Đurović, Water Level Control in the Thermal Power Plant Steam Separator Based on New PID Tuning Method for Integrating Processes, ENERGIES, ISSN:1996-1073, Vol. 15, No. 17, pp. 6310-6326, Aug, 2022. DOI:10.3390/en15176310 (M23) IF:3.252

M20-11. S. Ičagić, **G. Kvaščev**, A Smart Alcoholmeter Sensor Based on Deep Learning Visual Perception, SENSORS, ISSN:1424-8220, Vol. 22, No. 19, pp. 1-17, Sep, 2022, DOI:10.3390/s22197394, (M21) IF: 3.576

M20-12. M. Zivkovic, N. Bacanin, M. Antonijevic, B. Nikolić, **G. Kvaščev**, M. Marjanovic, N. Savanovic, Hybrid CNN and XGBoost Model Tuned by Modified Arithmetic Optimization Algorithm for COVID-19 Early Diagnostics from X-ray Images, Electronics, ISSN:2079-9292, Vol. 11, No. 3798, pp. 1-30, Nov, 2022. DOI:10.3390/electronics11223798 (M22) IF: 2.690

M20-13. I. Vajs, **G. Kvaščev**, T. Papić, M. Janković, Eye-tracking Image Encoding: Autoencoders for the Crossing of Language Boundaries in Developmental Dyslexia Detection, IEEE ACCESS, ISSN:2169-3536, pp. 3024-3033, Jan, 2023. DOI: 10.1109/ACCESS.2023.3234438 (M22) IF: 3.476

M20-14. M. Bjekic, A. Lazovic, K. Venkatachalam, N. Bacanin, M. Zivkovic, **G. Kvaščev**, B. Nikolić, Wall segmentation in 2D images using convolutional neural networks, PEERJ Computer Science, ISSN:2376-5992, pp. 1-18, Sep, 2023. doi:10.7717/peerj-cs.1565, (M21), IF:3.8

M20-15. J. Kljajić, **G. Kvaščev**, Ž. Đurović, Reconstructing Nerve Structures from Unorganized Points., Applied Sciences., ISSN:2076-3417, 2023; 13(20):11421. DOI: 0.3390/app132011421, (M22), IF: 2.7

M30-14. Goran Kvaščev, Zeljko Djurovic, and Aleksandra Marjanovic, *A Practical Solution for Control of Fan Drives Load in Thermal Power Plants,* 2012 International Conference on Advanced Electrical Engineering, Hong Kong, September 4-5, 2012)

M30-15. *Kvaščev, G., Đurović, Z., Kovačević, B., Kovačević, I.K., „Adaptive estimation of time-varying parameters in AR models with variable forgetting factor"*, 17th IEEE Mediterranean Electrotechnical Conference (MELECON 2014), 13-16 April 2014, Beirut, pp: 68 – 73, DOI: 10.1109/MELCON.2014.6820509

M30-16. *Kvascev Goran, Djurovic Zeljko, Kovacevic Branko, Vujnovic Sanja, eLab as a tool for remote laboratory in process control, CD Proceedings on 11th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, pp 393-394, Polytechnic of Porto (ISEP) in Porto, Portugal from 26-28 February 2014, DOI: 10.1109/REV.2014.6784198

M30-17. *Aleksandra Marjanovic, Miroslav Krstic, Zeljko Djurovic, Goran Kvascev and Veljko Papic: Combustion distribution control using the extremum seeking algorithm*, The 11th European Workshop on Advanced Control and Diagnosis (ACD 2014), Berlin, Germany, 13 - 14 November 2014

M30-18. Ivan R. *Nikolić, Vesna N. Petkovski, Goran S. Kvaščev, „Neural Network-Based Modeling of a Thermal Power Plant Feedwater Pump"*, 12th Symposium on Neural Network Applications in Electrical Engineering (NEUREL 2014), 25-27. November 2014, Belgrade, Serbia, ISBN: 978-1-4673-1569-2

M30-19. *Maja Gajić-Kvaščev, V. Andrić, D. Korolija Crkvenjakov, G. Kvaščev, S. Gadžurić, S. Ridolfi, „Feature extraction from paint layer EDXRF spectra using pattern recognition techniques"*, CMA4CH 2014, Mediterraneum Meeting, Employ the Multivariate Analysis and Chemometrics in Cultural Heritage and Environment Fields, 5th ed., 14-17 December 2014, Rome, Italy, 2014

M30-20. *G. Kvaščev, Ž. Đurović, V. Vlatković, A Wide Pulse Response Procedure for Tuning of PD/PID Controller for Integrating Processes, 2015 International Conference on Computer and Computational Sciences (ICCCS), pp. 56-60, IEEE, Paris, Jan, 2015.*

M30-21. N. Cimbaljevic, A. Marjanović, G. Kvascev, License Plate Character Recognition Based on Hopfield Networks, IcETRAN 2015, Srebrno jezero, Srbija, Jun, 2015

M30-22. S. Draskovic, G. Kvascev, Z. Djurovic, V. Petrovic, B. Kovacevic, Speech Signal Parameter Identification Using Adaptive Forgetting Factor, IcETRAN 2015, Srebrno jezero, Srbija, Jun, 2015.

M30-23. B. Brković, G. Kvaščev, M. Bebić, Model Predictive Controller for Flying Shear Application, 18th International Symposium POWER ELECTRONICS Ee2015, pp. 1-5, Novi Sad, Serbia, Oct, 2015.

M30-24. S. Vujnović, A. Marjanović, L. Cokić, G. Kvaščev, Ž. Đurović, Modelovanje laboratorijskog modela sušare, 6th International conference on technics and informatics in education TIO 2016, Čačak, Serbia, 2016.

M30-25. A. Marjanović, S. Vujnović, L. Cokić, G. Kvaščev, Ž. Đurović, Upravljanje vazdušnom levitacijom, 6th International conference on technics and informatics in education TIO 2016, Čačak, Serbia, 2016.

M30-26. S. Vujnović, *A. Al*-Hasaeri, P. Tadić, G. Kvaščev, Acoustic noise detection for state estimation, IcETRAN 2016, Zlatibor, Serbia, 2016

M30-27. П. Тадић, A. Marjanović, С. Вујновић, Г. Квашчев, Fault Diagnosis in Redundant-Sensors Schemes: A Combined Parity Relations and Dedicated Observer Approach, XIII International

Conference on Systems, Automatic Control and Measurements (SAUM2016), Nis, Srbija, Nov, 2016

M30-28. B. Lukić, K. Jovanović, G. Kvaščev, Feedforward neural network for controlling qbmove maker pro variable stiffness actuator, Neural Networks and Applications (NEUREL), 2016 13th Symposium on, Belgrade, Serbia, Nov, 2016

M30-29. V. Bobić, P. Tadić, G. Kvaščev, Hand gesture recognition using neural network based techniques, Neural Networks and Applications (NEUREL), 2016 13th Symposium on, pp. 35 - 38, IEEE, Belgrade, Serbia, Nov, 2016

M30-30. N. Vlahović, G. Kvaščev, Noise reduction by using Autoassociative Neural Networks, Neural Networks and Applications (NEUREL), 2016 13th Symposium on, Belgrade, Serbia, Nov, 2016

M30-31. M. Novičić, M. Janković, G. Kvaščev, M. Popović, Classification of forearm movements based on kinematic parameters using artificial neural networks, Proc of the 25rd Telecommunications forum TELFOR 2017, pp. 1-4, Belgrade, Nov, 2017.

У последњем петогодишњем периоду

M30-32. A. Marjanović, G. Kvaščev, V. Stevanović, B. Kovačević, Ž. Đurović, Steam temperature regulation at the output of rapid water heater evaporator by means of the extremum seeking method, International Conference Energy and Ecology Industry, Oct, 2018.

M30-33. T. Miljković, M. Bjelić, D. Šumarac Pavlović, G. Kvaščev, Prepoznavanje duvačkih instrumenata pomoću hromaprofila i neuralne mreže, 18th International Symposium INFOTEH-JAHORINA, pp. 362-367, Jahorina, 2019.

M30-34. A. Al-Hasaeri, A. Marjanović, S. Vujnović, G. Kvaščev, Ž. Đurović, Probability of Detection and False alarm Density Estimation in Target Tracking Systems with Unknown Measurement Noise Statistics, 6th International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2019, Srebrno jezero, Serbia, 2019.

M30-35. G. Kvaščev, V. Panić, M. Gajić--Kvaščev, B. Terzić, Automatization of distillation process for spirit beverages production, Yucorr, pp. 167-172, Serbian Society of Corrosion and Materials Protection, Tara, Srbija, Sep, 2019.

M30-36. Goran Kvascev, Zeljko Djurovic, Avram Avramovic, Application of cascade control in the process of flue-gas desulfurization of thermal power plant, IcETRAN 2021, Stanišići, Republika Srpska, 2021

M30-37. Natalija Đorđević, Nenad Džamić, Aleksa Stojić and Goran Kvaščev, Denoising the open-loop step response using an encoder-decoder convolutional neural network, IcETRAN 2022, Novi Pazar, Serbia, 2022

M30-38. Tanja Boljanić, Jovana Malešević and Goran Kvaščev, Deep neural network approach for artifact detection in raw ECG, IcETRAN 2022, Novi Pazar, Serbia, 2022

M30-39. G. Kvaščev, V. Andrić, M. Gajić--Kvaščev, Ž. Đurović, The use of neural networks for EDXRF spectra processing, European Conference on X-ray Spectrometry 2022, Bruges. Belgium, Jun, 2022.

M30-40. S. Vukojičić, L. Ristić, G. Kvaščev, Comparation between PI and Model Predictive Control of Two Mass Resonant Mechanical System, 7th International Symposium on Environment-Friendly Energies and Applications (EFEA 2022), pp. 1-6, IEEE Xplore, Mauritius, Dec, 2022.

M30-41. S. Vukojičić, L. Ristić, G.Kvaščev, Model Predictive Control of Two Mass Resonant Mechanical System Optimized by Neural Network, 7th International Symposium on

Environment-Friendly Energies and Applications (EFEA 2022), pp. 1-6, IEEE Xplore, Mauritius, Dec, 2022.

M30-42. S. Ičagić, D. Drašković, B. Nikolić, G. Kvaščev, Surface Crack-specific Knowledge Localization in Pretrained Convolutional Neural Networks, PROCEEDINGS of X International Conference on Electrical, Electronic and Computing Engineering IcETRAN, pp. 1-5, ETRAN Society, Istočno Sarajevo, Bosnia and Herzegovina, Jun, 2023.

Категорија М50 – Часописи националног значаја

M50.1. Miroslav R. Mataušek, Goran S. Kvaščev, "Step Response Tuning of the PI Controller – Revisited", J. Aut. Control, Vol. 11(1): 59-71, 2001.

M50.2. Nebojša Radmilović, Slaviša Stojaković, Goran Kvaščev "One Solution of Master Controller in Thermal Power Plants", Journal of Automatic Control, Vol 18(1), Pages 5-8, October 2008.

M50.3. Nebojša Malešević, Lana Popović, Goran Bijelić and Goran Kvaščev, "Muscle twitch responses for shaping the multi-pad electrode for functional electrical stimulation", Journal of automatic control, University of Belgrade, Vol. 20:53-57, 2010.

M50.4. A. Marjanović, G. Kvaščev, P. Tadić, Ž. Đurović, „Application of predictive maintenance techniques in industrial systems", Serbian journal of electrical engineering, Vol. 8, No. 3, pp. 263-279, 2011

M50.5. Goran S. Kvaščev, Aleksandra Lj. Marjanovic, and Željko M. Đurovic, "Robust adaptive system identification of steam separator process in thermal power plants", ELECTRONICS, VOL. 15, NO. 2, pp. 49-53, 2011

У последњем петогодишњем периоду

M50.6. M. Zakić, G. Kvaščev, Procena mesta nastanka kvara na električnom vodu primenom veštačkih neuralnih mreža, ENERGIJA, EKONOMIJA, EKOLOGIJA, No. 4, pp. 68-74, 2022.

Категорија М60 - Зборници скупова националног значаја
(сви радови су поткатегорије М63 - Саопштење са скупа нац. значаја штампано у целини)

M60-1. M.R. Mataušek, A.D. Micić, G.S. Kvaščev, "New experimental methods for designing and tuning linear controllers", XLV Konferencija ETRAN 2001, Bukovička Banja.

M60-2. Kvaščev G.S., Jovanović D.S., Mataušek M.R., "Eksperimentalna analiza novog postupka podešavanja PI regulatora", XLV Konferencija ETRAN 2001, Bukovička Banja.

M60-3. Goran S. Kvaščev, Miroslav R. Mataušek, "Uporedna analiza PI/PID regulartora i kompenzatora transportnog kašnjenja (DTC)", XLVI Konferencija ETRAN 2002, Banja Vrućica - Teslić.

M60-4. Milosavljević M.M., Kvaščev G.S., Rakić, A.Ž., "Jedna klasa sistema za monitorisanje distribucije audiovizuelnih sadržaja", Konferencija Informacione tehnologije, Žabljak 2002.

M60-5. Miroslav R. Mataušek, Goran S. Kvaščev, Aleksandar Ž. Rakić, "Analiza postupaka za eksperimentalno podešavanje PI/PID" regulatora, XLVII Konferencija ETRAN 2003, Herceg Novi.

M60-6. Goran Kvaščev, Đačić Boban, Srbijanka Turjalić: "Eksperimentalna verifikacija SRT metode podešavanja PI regulatora", XLIX Konferencija ETRAN 2005, Budva

M60-7. Igor Petričković, Goran Kvaščev, Branko Kovačević: "Prikaz uređaja Microbox dSpace i mogućnosti njegove primene", XLIX Konferencija ETRAN 2005, Budva

M60-8. Goran Kvaščev, Srbijanka Turajlić: "Izbor radnih tačaka u rešavanju 'tracking' problema nelinearnih sistema", L Konferencija ETRAN 2006, Beograd

M60-9. Rakić, A. Ž., Kvaščev, G. S. and Kovačević, B. D. "Maneuver target tracking using robust min-max kalman filtering", Konferencija TELFOR 2001, Beograd.

M60-10. Rakić A. Ž., Petrović T.B., Kvaščev G.S. "Easing EMI Problems in Power-supply Switching Devices by Robust Random Control", Konferencija TELFOR 2003, Beograd

M60-11. Goran Kvaščev, Milena Petrović, Nebojša Radmilović, Slaviša Stojaković "Raspodela uglja po mlinskim krugovima termoenergetskih blokova", LII Konferencija ETRAN 2008, Palić

M60-12. Nebojša Radmilović, Slaviša Stojaković, Goran Kvaščev, "Jedno rešenje glavnog regulatora u termoenergetskim postrojenjima", LII Konferencija ETRAN 2008, Palić

M60-13. Goran Kvaščev, Branko Kovačević: "Detekcija otkaza u sistemu ubrizgavanja", LIII Konferencija ETRAN 2009, Vrnjačka Banja

M60-14. Nebojša Radmilović, Milena Milojević, Nikola Krajnović, Goran Kvaščev: "Modelovanje i upravljanje sistemom loženja mazutom u termoelektranama ", LIII Konferencija ETRAN 2009, Vrnjačka Banja

M60-15. Vesna Petkovski, Goran Kvaščev, "Primena Kaskadne Strukture Regulacije za Kontrolu Temperature Sveže Pare u Termoenergetskim Postrojenjima", LIV Konferencija ETRAN 2010, Donji Milanovac

M60-16. Goran Kvaščev, Kovačevič Branko "Rekurzivna Robusna Identifikacija Sistema Zasnovana na QQ-krivama", LIV Konferencija ETRAN 2010, Donji Milanovac

M60-17. Veljko Papić, Željko Đurović, Goran Kvaščev, Predrag Tadić: A New Approach to Doppler Filter Adaptation in Radar Systems (Invited paper), 19th Telecommunications Forum TELFOR 2011, Belgrade

M60-18. Goran Kvaščev, Aleksandra Marjanović, Nasar Aldian Ambark Shashoa, „Robusna Adaptivna Identifikacija Parametara Separatora Pare Kotlovskih Termoenergetskih Postrojenja", LV Konferencija ETRAN 2011, Banja Vrućica (Teslić)

M60-19. Aleksandra Marjanović, Goran Kvaščev, Predrag Todorov, "Jedan Pristup Detekciji Otkaza na Bazi Identifikacije Procesa i Sekvencijalnog Testiranja Hipoteza", LV Konferencija ETRAN 2011, Banja Vrućica (Teslić)

M60-20. Predrag R. Tadić, Goran S. Kvaščev, Željko M. Đurović, „Dijagnoza kvarova senzora u separatoru termoelektrane testom generalizovanog količnika verodostojnosti", LVI Konferencija ETRAN 2012, Zlatibor, Srbija

M60-21. Aleksandra Marjanović, Goran Kvaščev, Predrag Tadić, „Skriveni markovljevi modeli u detekciji otkaza na industrijskim postrojenjima", LVI Konferencija ETRAN 2012, Zlatibor, Srbija

M60-22. A. Marjanović, G. Kvaščev, P. Todorov, Ž. Đurović, Управљање просторном расподелом температуре у котлу на основу визуелизације процеса сагоревања, 57. конференција ЕТРАН 2013, Друштво за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику, Златибор, Србија, Jun, 2013.

M60-23. N. Ambark Shashoa, G. Kvaščev, G. Kvaščev, B. Kovačević, Applicarion of two stage identification procedure in the frame of FDI, 57. конференција ЕТРАН 2013, Друштво за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику, Златибор, Србија, Jun, 2013.

M60-24. S. Vujnović, G. Kvaščev, P. Todorov, Udaljeno upravljanje vazdušnom levitacijom: laboratorijska postavka, ETRAN 2015, Srebrno Jezero, Srbija, 2015.

M60-25. L. Cokić, A. Marjanović, P. Tadić, G. Kvaščev, Analiza uticaja reda prediktora i pola govornika na odnos signal/šum za diferencijalni kvantizator, ETRAN 2016, Zlatibor, Srbija, 2016.

У последњем петогодишњем периоду

M60-26. Anita Lupšić, Marija Tomić, Goran Kvaščev, Algoritmi veštačke inteligencije i njihova primena za analizu ljudskih pokreta, Društvo za ETRAN, Palić, 2018.

M60-27. Т. Миљковић, M. Bjelić, Д. Шумарац Павловић, Г. Квашчев, Upotreba različitih obeležja za prepoznavanje drvenih duvačkih instrumenata korišćenjem neuralnih mreža, ЕТРАН, pp. 69-74, Društvo za ETRAN, Srebrno jezero, Jun, 2019.

M60-28. Milan Marinković, Uroš Romić, Goran Kvaščev, Boško Nikolić, Predikcija prosečne ocene i dužine trajanja studija upotrebom neuralnih mreža, Zbornik radova 28. IKT konferencije YU INFO 2022, pp. 41-44, Informaciono društvo Srbije, Kopaonik, Mar, 2022.

M60-29. Aleksa Stojić and Goran Kvaščev, regulacija procesa sa astatizmom primenom PID kontrolera sa modifikovanim integralnim dejstvom, Društvo za ETRAN, Jun 2023

M60-30. Vladislava Bobić, Milica Đurić-Jovičić, Nataša Dragašević-Mišković, Vladimir Kostić and Goran Kvaščev, Comparison of two deep learning models for the recognition of parkinson's disease gait patterns, Društvo za ETRAN, Jun 2023

Категорија M80 - Техничка и развојна решења

M80-1. Поступак расподеле оптерећења по млинским круговима термоенергетског постројења, 2012, Ивана Бачвански-Јањатовић; Милан Бједов; Милисав Богдановић; Жељко Ђуровић; Горан Квашчев; Бранко Ковачевић; Бојан Папић; Вељко Папић; Весна Петковски; Небојша Радмиловић; Драган Радојевић; Милена Милојевић; Никола Крајновић; Иван Николић;

M80-2. Нова метода и реализација управљања расподелом оптерећења дуалних вентилатора у термоенергетском постројењу, 2012, Биљана Антић; Жељко Ђуровић; Љубиша Јовановић; Горан Квашчев; Владимир Неранџић; Вељко Папић; Весна Петковски; Небојша Радмиловић; Александар Сушић; Ђорђе Човић; Вања Чукалевски; Александра Марјановић; Милена Милојевић; Никола Крајновић; Иван Николић;

M80-3. Решење индустријског ПИД регулатора за примену у аутоматском управљању разноврсним процесима у термоелектрани, 2012, Милан Бједов; Младен Вучинић; Жељко Ђуровић; Горан Квашчев; Бојан Папић; Вељко Папић; Весна Петковски; Богдан Поповић; Небојша Радмиловић; Драган Радојевић; Срђан Сударевић; Александра Марјановић; Милена Милојевић; Никола Крајновић; Иван Николић; Милош Станковић;

M80-4. Библиотека функција за одређивање параметара воде у различитим фазним стањима оптимизованих за рад у реалном времену, 2013, -, Милан Бједов; Милисав Богдановић; Драган Бојанић; Жељко Ђуровић; Василије Јовановић; Горан Квашчев; Никола Крајиновић; Милена Милојевић; Миленко Николић; Небојша Пањевац; Бојан Папић; Весна Петковски; Небојша Радмиловић; Драган Радојевић; Иван Николић;

M80-5. Један начин реализације координисане контроле система више парних котлова и турбине за потребе напређених система оптимизације рада термоелектране, 2013, -

Биљана Антић; Мирсад Бахтијаревић; Драган Бојанић; Младен Вучинић; Жељко Ђуровић; Горан Квашчев; Никола Крајиновић; Бојан Папић; Весна Петковски; Небојша Радмиловић; Александар Супић; Иван Николић;

M80-6. Реализација граничника пада градијента притиска свеже паре испред турбине у систему турбинске регулације парне турбине, 2014, Ана Вучуревић; Жељко Ђуровић; Горан Квашчев; Бранко Ковачевић; Никола Крајиновић; Милена Милојевић; Дарко Новаковић; Весна Петковски; Небојша Радмиловић; Срђан Сударевић; Ђорђе Човић; Иван Николић; Тамара Јовановић;

M80-7. Алгоритам аутоматског тестирања функционисања стоп вентила парне турбине са одвојеним управљачким сервопогонима регулационих и стоп вентила - пример турбине 18-К-350, 2014, Биљана Антић; Милисав Богдановић; Жељко Ђуровић; Горан Квашчев; Никола Крајиновић; Милена Милојевић; Владимир Неранџић; Весна Петковски; Небојша Радмиловић; Предраг Тадић; Иван Николић; Тамара Јовановић; Александар Латиновић;

M80-8. Реализације главног регулатора количине ваздуха за сагоревање угља у котловском постројењу термоелектране, 2014, Мирсад Бахтијаревић; Жељко Ђуровић; Горан Квашчев; Никола Крајиновић; Милена Милојевић; Вељко Папић; Весна Петковски; Небојша Радмиловић; Вања Чукалевски; Иван Николић; Тамара Јовановић; Зоран Стојковић; Радиша Рајић;

M80-9. Један приступ моделовању воденог тракта котла за потребе симулатора-тренажера термоенергетског блока, 2015, Жељко Ђуровић; Горан Квашчев; Никола Крајиновић; Милена Милојевић; Весна Петковски; Небојша Радмиловић; Тамара Јовановић;

M80-10. Симулатор типских извршних органа термоенергетског блока као додатна компонента VIEW® T-POWER DCS система, 2015, Жељко Ђуровић; Горан Квашчев; Никола Крајиновић; Милена Милојевић; Бојан Папић; Весна Петковски; Небојша Радмиловић; Тамара Јовановић;

M80-11. S. Ičagić, G. Kvaščev, System and method for digitalization of breathalyzer measurements based on artificial intelligence, Aug, 2022.

Цитираност

У бази података SCOPUS кандидат има 42 радова у часописима и на конференцијама су цитирани у укупно 209 пута, без аутоцитата и то M20-1:(14 пута), M20-2:(5 пута), M20-3:(17 пута), M20-4:(11 пута), M20-5:(35 пута), M20-6:(3 пута), M20-7:(2 пут), M20-8:(22 пут), M20-9:(9 пут), M20-10:(3 пут), M20-12:(21 пут), M20-13:(1 пут), M30-29:(20 пута), M30-1:(7 пута), M30-28:(6 пута), M30-30:(5 пута), M30-16:(4 пута), M30-6(4 пута), M30-2:(4 пута)...

Хиршов индекс кандидата у анализи без аутоцитата је H=8.

Д. Пројекти

Горан Квашчев је учествовао у реализацији 26 међународног, националног, иновационог, истраживачког, развојног и мултидисциплинарног пројекта, и то:

1. Развој и реализација дигиталног регулатора са аутоматским подешавањем за управљање индустријским процесима, Министарство за науку и технолошки развој Р. Србије IT.1.05.0177.B, 2002-2004.
2. Информационе и комуникационе технологије у здравственој заштити (оригинални назив: Information and Communication Technologies in Health Care *INCO-Health*), EU TEMPUS CD-

3. Projektovanje i implementacija sistema regulacije i upravljanja kotlovskog postrojenja, Benson tip, Blok A1 (210MW), TENT "Nikola Tesla A", Obrenovac, Srbija 2005, Institut Mihajlo Pupin
4. Optimizacija i projektovanje sistema regulacije sagorevanja radi minimizacije emisije NOx gasova, TE "Kostolac B", Kostolac, Serbia 2014-2015, Siemens Srbija
5. Pilot e-Lab Experiment, UNESCO & Hewlett-Packard: Piloting Solutions for Alleviating Brain Drain in South East Europe, 2005-2006.
6. Пројекат WUS-Austria, Course Development Program Plus, "Support to Higher Education in Serbia and Montenegro 2005-2007", 2005-2007.
7. Пројекат Министарства науке и заштите животне средине републике Србије, технолошки развој, "Развој интегрисаног навигационог система за примену у аутоматском лоцирању возила", 2007-2010.
8. Пројекат Министарства науке и заштите животне средине републике Србије, технолошки развој, "Развој нових метода за моделирање телекомуникационих система", 2007-2010.
9. Даљинско управљање роботизованим системима путем гласа, Министарство за науку и заштиту животне средине Р. Србије ТР-6147, 2005-2007.
10. Аутоматизовани систем противградне заштите, Министарство за науку и заштиту животне средине Р. Србије ТР-6124, 2005-2007.
11. Projektovanje i implementacija sistema regulacije i upravljanja kotlovskog postrojenja, Benson tip, Blok A4 (300MW) TENT "Nikola Tesla A", Obrenovac, Srbija 2007, Institut Mihajlo Pupin
12. Projektovanje i implementacija sistema regulacije i upravljanja kotlovskog postrojenja, Sulzer tip, Blok B1 (1000t/h, 348MW), TE "Kostolac B", Kostolac, Srbija 2008, Institut Mihajlo Pupin
13. Projektovanje i implementacija sistema regulacije i upravljanja kotlovskog postrojenja, Sulzer tip, Blok A6 (300MW), TE "Nikola Tesla A", Obrenovac, Srbija 2009, Institut Mihajlo Pupin
14. Projektovanje i podešavanje sistema regulacije, TGME-464/S tip, Kotao 3 (500t/h), TETO "Novi Sad", Novi Sad, Srbija 2009 , Institut Mihajlo Pupin
15. Projektovanje i implementacija sistema regulacije i upravljanja kotlovskog postrojenja, Sulzer tip, Blok B2 (1000t/h, 350MW), TE "Kostolac B", Kostolac, Srbija 2010, Institut Mihajlo Pupin
16. Power Plants Robustification Based on Fault Detection and Isolation Algorithms (*PRODI*), EU FP7-ICT INFSO-ICT-224233, 2008-2011.
17. Building Network of Remote Labs for Strengthening University - Secondary Vocational Schools Collaboration (*NeReLa*), EU TEMPUS 543667-2013, 2013-2016.
18. Robust Decentralised Estimation for Large-Scale Systems (*RODEO*), Executive Program for Scientific and Technical Cooperation between Italy and Serbia MAE-PGR00152, 2013-2015.
19. Пројекат Министарства за науку и технолошки развој, "Повећање енергетске ефикасности и расположивости у системима за производњу и пренос електричне енергије развојем нових метода за дијагностику и рану детекцију отказа", 2011-2014.
20. Пројекат Министарства за науку и технолошки развој, "Систем за оптимизацију рада термоблока са турбоагрегатором снаге веће од 300 MW", 2011-2014.
21. Optimizacija i projektovanje DCS sistema regulacije kotlovskog postrojenja, Blok B1 (650MW), TE "Nikola Tesla B" Obrenovac, Srbija, 2014-2016, EPS
22. FAult and STate detection of Rotary machineries based on acoustic signals (FASTER) EUREKA projekt, Beograd, 2018-2020

23. Automatic SVR Ball Inspection, HENKEL SRBIJA d.o.o. Kruševac, 2019, design, testing, commissioning, guarantee testing. Serbia
24. Projektovanje električnih instalacija i upravljanja borbenih vozila Milos 4x4, BORBENI SLOŽENI SISTEMI D.O.O., 2019
25. Izrada tehnicke dokumentacije elektricnih instalacija i racunarskog upravljanja visenamenskog oklopnog borbenog vozila Lazar, BORBENI SLOŽENI SISTEMI D.O.O., 2019.
26. Izrada tehnicke dokumentacije elektricnih instalacija i racunarskog upravljanja visenamenskog oklopnog borbenog vozila Dusan, ЈУГОИМПОРТ - СДПР Ј.П., 2020

Ђ. Остали резултати

Горан Квашчев је рецензент међународних часописа: IEEE Transaction on Education, IET Science, Measurement & Technology, Computer Science and Information Systems. Такође, вишегодишњи је рецензент конференција НЕУРЕЛ, ТЕЛФОР, (Иц)ЕТРАН. Члан је међународног удружења IEEE, као и националног друштва ЕТРАН. Од јануара 2001. је технички едитор часописа Journal of automatic control.

У факултетским оквирима, ангажовање Горана Квашчева огледало се кроз учешће у раду комисија и руководећим позицијама:

- 2023-данас: члан Савета факултета
- 2022-данас: председник Статутарне комисије
- 2021-данас: шеф Катедре за сигнале и системе
- у два мандата је био члан дисциплинске комисије, од тога у једном мандату је био председник комисије
- 2012-2015: члан финансијске комисије
- 2015-2018. године врши је функцију заменика Шефа Катедре за сигнале и системе
- 2015-2019. године је продекан за финансије Електротехничког факултета.

Београд, 2.2.2024. године

Др Горан Квашчев, ванредни професор
Универзитет у Београду – Електротехнички факултет

Univerzitet Crne Gore
adresa / address_Cetinjska br. 2
81000 Podgorica, Crna Gora
telefon / phone_00382 20 414 255
fax_00382 20 414 230
mail_rektorat@ucg.ac.me
web_www.ucg.ac.me
University of Montenegro

Broj / Ref 03 - 2401

Datum / Date 04.06.2020

05.06.2020

02/1 609

Na osnovu člana 72 stav 2 Zakona o visokom obrazovanju („Službeni list Crne Gore" br 44/14, 47/15, 40/16, 42/17, 71/17, 55/18, 3/19, 17/19, 47/19) i člana 32 stav 1 tačka 9 Statuta Univerziteta Crne Gore, Senat Univerziteta Crne Gore na sjednici održanoj 04.06.2020. godine, donio je

ODLUKU
O IZBORU U ZVANJE

Dr Vesna Popović Bugarin bira se u akademsko zvanje redovni profesor Univerziteta Crne Gore za **oblasti Računarstvo i Digitalna obrada signala,** na Elektrotehničkom fakultetu Univerziteta Crne Gore, na neodređeno vrijeme.

SENAT UNIVERZITETA CRNE GORE
PREDSJEDNIK

Prof. dr Danilo Nikolić, rektor

Prof. dr Vesna Popović-Bugarin

BIOGRAFIJA

Vesna Popović-Bugarin je rođena 03. 05. 1978. godine u Podgorici. Osnovnu i srednju školu (Gimnazija "Slobodan Škerović", prirodno-matematički smjer) završila je u Podgorici. U toku školovanja učestvovala je i osvajala nagrade na opštinskim i republičkim takmičenjima u znanju iz fizike. Diplomirala je, magistrirala i doktorirala 2001, 2005. i 2009. godine, respektivno, na Elektrotehničkom fakultetu (ETF) u Podgorici.

Tokom postdiplomskih studija boravila je u Ženevi, Švajcarska, na institutu za nuklearna istraživanja – CERN, u periodu od 08. 06. 2004. do 18. 07. 2004. godine, dok je tokom doktorskih studija boravila po mjesec dana u: Brestu, Francuska, na ENSIETA-i (École Nationale Supérieure d'Ingénieurs), kao i u Bonu, Njemačka, na Univerzitetu primijenjenih nauka, Bonn-Rhein-Sieg University of Applied Sciences.

Vesna Popović-Bugarin je zaposlena na ETF-u od 2002. godine, 27.05.2010. godine je izabrana u zvanju docenta, 24.06.2015. u zvanje vanrednog profesora, a 04. 07. 2020. u zvanje redovnog profesora.

Oblasti njenog interesovanja uključuju vještačku inteligenciju, vremensko-frekvencijsku analizu signala, obradu radarskih signala i analizu mikro-Doppler efekta u radarskim signalima.

Vesna Popović-Bugarin je bila angažovana na velikom broju domaćih i međunarodnih naučnih projekata, kao i na dva FP7 projekta. Objavila je preko 40 naučnih radova, od čega 15 u međunarodnim časopisima sa SCI liste. Koautor je jednog domaćeg udžbenika i po jednog poglavlja u dvijema monografijama izdatim od strane inostranih izdavača.

Član je profesionalnih udruženja: Institute of Electrical and Electronics Engineers (IEEE) i IEEE Signal Processing Society. Bila je član odbora za informaciono-komunikacione tehnologije pri CANU i Centra za mlade naučnike pri CANU.

Recenzent je u više vodećih međunarodnih časopisa sa SCI liste, kao i na nekoliko naučnih konferencija.

Bila je član radne grupe za izradu obrazovnih programa (uključujući standarde zanimanja i standarde kvalifikacija) Elektrotehničar elektronike, Elektrotehničar elektronskih komunikacija, Elektrotehničar računarskih sistema i mreža i Elektrotehničar za razvoj veb i mobilnih aplikacija (septembar 2017-januar 2018), pri Centru za stručno obrazovanje.

Bila je član ekspertskog tima za evaluaciju obrazovnih programa srednjih stručnih škola elektrotehnike u Crnoj Gori: Srednja elektrotehnička škola „Vaso Aligrudić" Podgorica (maj 2014) i Srednja stručna škola Nikšić (maj 2014). Evaluaciju je organizovao Centar za stručno obrazovanje.

Jedan je od autora Elaborata za transformaciju Elektrotehničkog fakulteta u Fakultet za elektrotehniku, računarski inženjering i informacione tehnologije. Takođe je jedan od članova tima koji je pripremao projektnu dokumentaciju i učestvovao u pregovorima koji su rezultovali u dobijanju granta za BIO-ICT projekat (Centar uspješnosti u bioinformatici). Projekat je planirano da traje tri godine, a u cjelosti ga je finansiralo Ministarstvo nauke Crne Gore iz kredita Svjetske banke u okviru HERIC (Higher Education and Research for Innovation and Competitiveness) projekta. Učestvovala je u pisanju i realizaciji TEMPUS projekta: „Razvoj kurikuluma postdiplomskih studija primenjenog računarskog inženjerstva i internacionalizacije postdiplomskih studija ETF-a". Trenutno učestvuje u realizaciji ERASMUS projekta DUALMON: „Strengthening capacities for the implementation of dual education in Montenegro higher education / DUALMON".

Bila je član sam tima koji je osvojio prvu nagradu na takmičenju Otvorene Ideje za Crnu Goru sa aplikacijom BUDI ODGOVORAN. Aplikacija (Web portal www.budiodgovoran.me i Android aplikacija) se u saradnji sa Vladom Crne Gore, a uz značajnu podršku Programa za razvoj Ujedinjenih nacija (UNDP), Ambasade Ujedinjenog Kraljevstva u Crnoj Gori, trenutno koristi u borbi protiv sive ekonomije. Projekat BUDI ODGOVORAN je osvojio drugu nagradu međunarodne inicijative *Partnerstva otvorenih vlada* u konkurenciji najboljih projekata iz 33 zemlje svijeta koji afirmišu uključivanje građana u javne politike.

Koordinator je Startup akademije koju drugu godinu zaredom relizuje Elektrotehnički fakultet, Univerziteta Crne Gore, u saradnji sa Kompanijom Amplitudo.

Obavljala je funkciju zamjenika naučnog direktora BIO-ICT Centra izvrsnosti u periodu od 2013. do 2018. godine. Bila je prodekan za razvoj i istraživanje na Elektrotehničkom fakultetu, Univerziteta Crne Gore, u periodu od 2019-2022. godine.

Prof. dr Vesna Popović-Bugarin

BIBLIOGRAFIJA

Međunarodni časopisi i poglavlja u monografijama

1. **V. Popović-Bugarin**, and S. Djukanović, "A Low Complexity Model Order and Frequency Estimation of Multiple 2-D Complex Sinusoids," Digital Signal Processing, Vol. 104, September 2020, https://doi.org/10.1016/j.dsp.2020.102794
2. **V. Popović-Bugarin**, and S. Djukanović, "Efficient instantaneous frequency estimation in high noise based on the Wigner distribution," *Signal Processing*, Vol. 157, pp. 25-29, April 2019.
3. S. Djukanović, and **V. Popović-Bugarin**, "Efficient and accurate detection and frequency estimation of multiple sinusoids," *IEEE Access*, Vol. 7, pp. 1118 – 1125, December 2018., DOI: 10.1109/ACCESS.2018.2886397.
4. E. Salković, I. Djurović, M. Knežević, **V. Popović-Bugarin**, A. Topalović, "Digitization and mapping of national legacy soil data of Montenegro," *Soil and Water Research*, 81/2017-SWR, Vol. 13, No.2, pp. 83–89, 2018.
5. M. Brajović, **V. Popović-Bugarin**, I. Djurović, and S. Djukanović, "Post-processing of Time-Frequency Representations in Instantaneous Frequency Estimation Based on Ant Colony Optimization," *Signal Processing*, vol. 138, pp. 195–210, September 2017.
6. I. Djurović, **V. Popović-Bugarin**, and M. Simeunović, "The STFT-based estimator of micro-Doppler parameters," *IEEE Transactions on Aerospace and Electronic Systems*, DOI: 10.1109/TAES.2017.2669741. Vol. 53, No. 3, June 2017.
7. E. Hot, **V. Popović-Bugarin**, "Soil data clustering by using K-means and fuzzy K-means algorithm," *Telfor Journal*, Vol. 8, No. 1, pp. 56-61, 2016.
8. LJ. Stanković, M. Daković, T. Thayaparan, and **V. Popović-Bugarin**, "Inverse Radon Transform Based Micro-Doppler Analysis from a Reduced Set of Observations," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 51, No. 2, pp.1155-1169, April 2015.
9. LJ. Stanković, M. Daković, T. Thayaparan, and **V. Popović-Bugarin**, "Signal Decomposition of Micro-Doppler Signatures," *Radar Micro-Doppler Signatures: Processing and Applications*, Ed. V. C. Chen, D. Tahmoush, and W. J. Miceli, IET, pp.273-327, 2014.
10. LJ. Stanković, **V. Popović-Bugarin**, and F. Radenović, "Genetic algorithm for rigid body reconstruction after micro-doppler removal in the radar imaging analysis," *Signal Processing*, Volume 93, Issue 2013, Jan 2013.
11. LJ. Stanković, M. Daković, T. Thayaparan, and **V. Popović-Bugarin**, "Micro-Doppler Removal in the Radar Imaging Analysis," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 49, No. 2, April 2013, pp.1234-1250
12. S. Djukanović, and **V. Popović-Bugarin**, "A parametric method for multicomponent interference suppression in noise radars," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2730–2738, July 2012.

13. S. Djukanović, **V. Popović-Bugarin**, M. Daković, and LJ. Stanković, "A parametric method for non-stationary interference suppression in direct sequence spread-spectrum systems," *Signal Processing*, Vol. 91, No. 6, pp. 1425-1431, June 2011

14. I. Djurović, C. Ioana, T. Thayaparan, LJ. Stanković, P Wang, **V. Popović-Bugarin**, and M. Simeunović, "Cubic-phase function evaluation for multicomponent signals with application to SAR imaging," *IET Signal Processing*, vol. 4, no. 4, August 2010, pp. 371-381

15. T. Thayaparan, LJ. Stanković, M. Daković, and **V. Popović-Bugarin**, "Micro-Doppler parameter estimation from a fraction of the period," *IET Signal Processing*, vol. 4, no. 3, pp. 201-212, June 2010. doi: 10.1049/iet-spr.2009.0093, ISSN: 1751-9675

16. **V. Popović-Bugarin**, I. Djurović, LJ. Stanković, T. Thayaparan, and M. Daković, "Autofocusing of SAR images based on parameters estimated from the PHAF," *Signal Processing*, Vol. 90, No. 5, May 2010, Pages 1382-1391

17. I. Djurović, LJ. Stanković, **V. Popović-Bugarin**, M. Daković, and T. Thayaparan, "Time-frequency analysis for SAR and ISAR imaging," *in GeoSpatial Visual Analytics*, Springer 2009, ISBN 978-90-481-2898-3, pp. 113-127

18. LJ. Stanković, T. Thayaparan, **V. Popović-Bugarin**, I. Djurović, and M. Daković, "Adaptive S-Method for SAR/ISAR Imaging," *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, No. Jan. 2008, Page(s) 10 pages

Nacionalni časopisi:

1. S. Vujović, M. Brajović, **V. Popović-Bugarin**, N. Latinović, J. Latinović, and M. Bajčeta, "A web service for grapevine monitoring and forecasting a disease," *ETF Journal of Electrical Engineering*, Vol. 22, No. 1, 2016.

Međunarodne konferencije

1. I. Knežević, I. Čavor, T. Dlabač, and **V. Popović-Bugarin**: "Using Machine Learning Techniques for Predicting Electrical Data of PV Panels from RGB Images", in Proc of 10th International Maritime Science Conference (IMSC 2023), May 2023, Croatia.

2. J. Vujović, **V. Popović-Bugarin** and Š. Ivošević, "Corrosion prediction of fuel tanks on bulk carrier," 2nd Kotor International Maritime Conference, November 2022, Kotor, Montenegro.

3. M. Đukanović, S. Rogić, L. Novićević, **V. Popović-Bugarin** and M. Jovanović: "Application of Apriori Algorithm for CRM Improvement - Case Study from Montenegro", ICCTA 22: Proceedings of the 2022 8th International Conference on Computer Technology Applications, pp. 48–56, May 2022.

4. M. Vukčević, **V. Popović-Bugarin**, E. Dervić, "DBSCAN and CLARA clustering algorithms and their usage for the soil data clustering", 8th Mediterranean Conference on Embedded Computing (MECO), 10-14 Jun 2019, ISBN: 2637-9511

5. M. Bajčeta, P. Sekulić, S. Djukanović, T. Popović, and **V. Popović-Bugarin**, "Retinal blood vessels segmentation using ant colony optimization," *2016 13th Symposium on Neural Networks and Applications (NEUREL)*, Belgrade, November 22-24, 2016.

6. E. Hot, and **V. Popović-Bugarin**, "Soil data clustering by using K-means and fuzzy K-means algorithm," *23rd Telecommunications Forum TELFOR 2015*, Serbia, Belgrade, November 24-26, 2015.

7. E. Hot, **V. Popović-Bugarin**, A. Topalović, M. Knežević, "Generating thematic pedologic maps by using data mining and interpolations," *3rd International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2016*, Zlatibor, Serbia, June 2016.

8. M. Brajović, and **V. Popović-Bugarin**, "Instantaneous Frequency Estimation Using Ant Colony Optimization and Wigner Distribution," *4th Mediterranean Conference on Embedded Computing, MECO 2015*, Budva, Montenegro, June 2015.

9. **V. Popović-Bugarin**, F. Radenović, and LJ. Stanković, "Analytical Rigid Body Reconstruction after Micro-Doppler Removal," *in Proc. of ISPA 2013*,

10. S. Djukanović, and **V. Popović-Bugarin**, "An efficient peak frequency estimator for product high-order ambiguity function," *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011.

11. S. Djukanović, and **V. Popović-Bugarin**, "A Robust Estimation of Polynomial-Phase Coefficients," *18th Telecommunications Forum TELFOR 2010*, Serbia, Belgrade, November 23-25, 2010.

12. **V. Popović-Bugarin**, I. Djurović, LJ. Stanković, T. Thayaparan, and M. Daković, "Autofocusing of SAR Images Based on the LPFT and the PHAF," *ANTEM/AMEREM 2010, 14TH International Symposium on Antennas and Electromagnetics and The American Electromagnetics Conference*, July 2010, Ottawa, ON, CANADA

13. T. Thayaparan, LJ. Stanković, M. Daković, I. Djurović, and **V. Popović-Bugarin**, "Image enhancement and motion compensation of moving targets in ISAR using S-method," *International Radar Conference Surveilance for a Safer World 2009, RADAR*, 2009, pp. 1-5.

14. I. Djurović, M. Daković, and **V. Popović-Bugarin**, "Numerical modeling in radar data analyzing," *6-th International Seminar Mathematical Models & Modeling in Laser – Plasma Processes*, Budva, Crna Gora, pp. 43, jun 2009.

15. LJ. Stanković, T. Thayaparan, M. Daković, and **V. Popović-Bugarin**, "S-Method in radar imaging," *EUSIPCO 2006*, Florence, Italy, 4-8 September 2006.

16. **V. Popović-Bugarin**, M. Daković, T. Thayaparan, and LJ. Stanković, "Sar Images Improvements by Using The S-Method," *in Proc. of ICASSP 2006*, Vol.3, 21-24 May 2006, pp.III-177-III-180

17. **V. Popović-Bugarin**, T. Thayaparan, and LJ. Stanković, "Noise analysis of the high resolution methods in ISAR," *ISPA 2005*, 15-17 Sept. 2005 pp.489 - 493

18. LJ. Stanković, **V. Popović-Bugarin**, and M. Daković, "On the Capon's method application in time-frequency signal analysis," *in Proc. of ISSPIT 2003*, Darmstadt, Germany, 14-17 December 2003, pp. 721-724.

Nacionalne i regionalne konferencije

1. E. Hot, and **V. Popović-Bugarin**, "Analiza fuzzy K-means klasterizacije koristeći pedološku bazu podataka Crne Gore," *Informacione Tehnologije - IT 2016*, March 2016.

2. S. Vujović, M. Brajović, **V. Popović-Bugarin**, N. Latinović, J. Latinović, and M. Bajčeta, "WEB servis za monitoring vinograda i predviđanje oboljenja vinove loze," *Informacione Tehnologije - IT 2016*, March 2016.

3. M. Brajović, S. Vujović, **V. Popović-Bugarin**, S. Djukanović, M. Knežević, and A. Topalović, "Soil analysis database and the expert system for recommendations of fertilization in agriculture," *Informacione Tehnologije - IT 2016, March 2016*,
4. P. Sekulić, M. Bajčeta, **V. Popović-Bugarin**, S. Djukanović, and A. Mihailović, "One approach to the optimization of public transport," *Informacione Tehnologije - IT* 2016, *March 2016*,
5. **V. Popović-Bugarin**, F. Radenović, and LJ. Stanković, "Genetički algoritam za rekonstrukciju nedostajućih odbiraka kratkotrajne Fourier-ove transformacije," *Informacione Tehnologije - IT 2013*,
6. I. Djurović, M. Simeunović, and **V. Popović-Bugarin**, "Primjena modifikovane kubične fazne funkcije u estimaciji parametara SAR signala," *Informacione tehnologije IT '09*, Žabljak, 2009.

Udžbenici

1. I. Djurović, **V. Popović-Bugarin**, M. Simeunović, and P. Raković, *Programiranje II – praktikum za laboratorijske vježbe*. Elektrotehnički fakultet, Podgorica 2010
2. I. Djurović, S. Djukanović, and **V. Popović-Bugarin**, *Programiranje I – praktikum za laboratorijske vježbe*. Elektrotehnički fakultet, Podgorica 2010
3. I. Djurović, S. Djukanović, and **V. Popović-Bugarin**, *Programski jezik C sa zbirkom urađenih zadataka*. textbook, Elektrotehnički fakultet, Podgorica, 2006.
4. I. Djurović, S. Djukanović, and **V. Popović-Bugarin**, *Programski jezik C: Predavanja sa zbirkom riješenih zadataka*. ETF Edicija Udžbenici, Podgorica 2004
5. I. Djurović, and **V. Popović-Bugarin**, *Objektno orijenisano programiranje na programkom jeziku C++ sa zbirkom riješenih zadataka*. Skripta, ETF, Podgorica 2004.

Univerzitet Crne Gore

University of Montenegro

Na osnovu člana 72 stav 2 Zakona o visokom obrazovanju („Službeni list Crne Gore" br. 44/14, 47/15, 40/16, 42/17, 71/17, 55/18, 3/19, 17/19, 47/19, 72/19 i 74/20 i 104/21) i člana 32 stav 1 tačka 9 Statuta Univerziteta Crne Gore, Senat Univerziteta Crne Gore na sjednici održanoj 09.03.2022. godine, donio je

O D L U K U
O IZBORU U ZVANJE

Dr IGOR JOVANČEVIĆ bira se u akademsko zvanje **docent Univerziteta Crne Gore iz oblasti Računarske nauke na Prirodno-matematičkom fakultetu Univerziteta Crne Gore,** na period od pet godina.

SENAT UNIVERZITETA CRNE GORE
PREDSJEDNIK
Božović
Prof. dr Vladimir Božović, rektor

BIOGRAFIJA - dr Igor Jovančević, docent

OBRAZOVANJE

Igor Jovančević je diplomirao 2008.g na Prirodno-matematičkom fakultetu, Univerziteta Crne Gore sa **prosječnom ocjenom 9.2**, odbranom diplomskog rada: *Segmentacija objekta u prvom planu u video snimku tehnikom sekvencijalnog klasterisanja karakterističnih tačaka baziranog na pokretu.* Rad je izrađen na Institutu za sisteme za učenje u realnom vremenu (Institut für Echtzeit Lernsysteme), **Univerzitet Siegen-Njemačka** u okviru projekta AMOR (Autonomous Mobile Outdoor Robot).

Diplomirao je 2013.g na 2-godišnjem evropskom **Erasmus Mundus Joint Master programu VIBOT** iz oblasti kompjuterske vizije i robotike, školovanjem na 3 univerziteta (zajednička diploma): Université de Dijon, Dižon (Francuska); Universitat de Girona, Đirona (Španija) i Heriot Watt University, Edinburg (Velika Britanija). Tema magistarskog rada: *Praćenje više podmornica koristeći samo azimut podatke sa sonara.*

Na **Ecole des Mines d'Albi - Univerzitet u Tuluzu, Francuska** je doktorirao 2016.g radom na inovativnom naučno-industrijskom projektu „*Air-Cobot*" (https://en.wikipedia.org/wiki/Air-Cobot) za izradu robota za inspekciju aviona u okviru kampanje **Airbus**-a *Hangar of The Future*. Tema doktorske disertacije: *Vizuelna inspekcija spoljašnjosti aviona koristeći Pan-Tilt-Zoom kameru i 3D skener instalirane na pokretnom robotu: obrada 2D slika i 3D oblaka tačaka.*

PODACI O RADNIM MJESTIMA I IZBORIMA U ZVANJE

Od 2009.g do 2011.g je radio kao **IT inženjer** i sistem analitičar u Procesing centru Crnogorske Komercijalne Banke u Podgorici.

Od 2013. do 2016. godine je radio kao doktorand i saradnik u nastavi na **Ecole des Mines d'Albi - Univerzitet u Tuluzu, Francuska**. Uporedo, bio je ko-mentor na magistarskim studijama i držao praktičnu nastavu na sljedećim predmetima: operaciona istraživanja, nelinearna optimizacija, numerički alati, teroija vjerovatnoće, statistika.

2017.g dobio je **nagradu za najbolju doktorsku disertaciju u 2016.** godini u oblasti avio- i svemirske industrije od francuskog klastera ***Aerospace Valley***, prvog svjetskog naučno-industrijskog klastera za avio- i svemirsku industriju.

Od 2016. do 2021.g radio je kao **istraživač-inženjer i naučni mentor** u Tuluzu u francuskoj kompaniji Diota, gdje je naučni rad fokusirao na primjene kompjuterske vizije na probleme automatske vizuelne inspekcije u industriji, oslanjajući se na savremena dostignuća robotike i lokalizacije senzora u realnom vremenu. Bio je odgovoran za kolaboraciju kompanije sa naučnim institucijama kao i ko-mentorstvo pri izradi dvije doktorske disertacije i četiri magistarska rada. Paralelno, kratko je radio kao **saradnik u nastavi** na Univerzitetu u Tuluzu na predmetu Matematička analiza.

2020.g dobio je status **„Qualification"** od francuskog nacionalnosg savjeta za univerzitete **(Conseil national des universités - CNU)** koji mu omogućava konkurisanje za naučno-istraživačke pozicije ranga **Maître de Conferences** u Francuskoj.

Od aprila 2021. radi na **Prirodno-matematičkom fakultetu Univerziteta Crne Gore**, prvo kao **saradnik sa doktoratom**, a od 9.3.2022. kao **docent**. U zvanju docent drži nastavu na predmetima:

1. Na *Prirodno-matematičkom fakultetu UCG*: Paralelni algoritmi, Paralelno programiranje, Programski jezici, Matematički softverski paketi, Multimedija, Mašinsko učenje.
2. Na *Mašinskom fakultetu UCG*, studijski program Mehatronika: Programiranje.

U istraživačkom radu bavi se primjenama kompjuterske vizije, većinom na probleme vizuelne inspekcije u industriji ali i na druge, kao što je inteligentno nadgledanje i bezbjednost fabričkih ćelija opremljenih robotima. Svakodnevno saraduje sa naučnim timom profesora Jean-José Orteu sa **Instituta Clément Ader (Tuluz, Francuska)** i **inženjerske škole IMT Mines Albi-Univerzitet u Tuluzu**. Zajedno sa timom predlaže jedinstvenu metodologiju za automatsku vizuelnu inspekciju kompleksnih mehaničkih sklopova koristeći kompjutersku viziju. Integrišu se 2D/3D pristupi (obrada 2D slika i CAD modela) i 3D/3D pristupi (obrada 3D oblaka tačaka i CAD modela). Cilj je verifikacija mehaničkih struktura u odnosu na referentni CAD model koristeći brzinu 2D analize i kompletnost 3D informacija. Posebna pažnja se poklanja novim doprinosima u domenu, posebno u oblasti Vještačke inteligencije i dubokog učenja (Deep Learning).

Od početka 2020.g učestvuje u **naučno-industrijskom projektu DECADOM** finansiranom od **francuskog regiona Occitanie**. Projekat rješava problem detekcije, klasifikacije i lokalizacije oštećenja koja mogu uticati na otpornost i estetske aspekte mehaničkih struktura, koristeći 2D/3D kompjutersku viziju.

Bio je član tehničke komisije za evaluaciju naučnih radova za međunarodnu konferenciju *Quality Control by Artificial Vision (QCAV2021)* koja se održala u maju 2021.g.: http://www.tc-iaip.org/qcav/2021/about.html.

U svojstvu Predsjedavajućeg za publikacije - **Publication chair** učestvuje u organizaciji međunarodne konferencije *Quality Control by Artificial Vision (QCAV2023)* koja se održava u junu 2023.g.: https://qcav2023.sciencesconf.org/

Ko-mentor je na doktorskim studijama na Univerzitetu u Tuluzu. Bio je mentor pri izradi 9 master teza i 2 doktorske teze.

Rukovodilac je projekta 2022-1-PL01-KA220-HED-000088359 *FAAI - Erasmus+: The Future is in Applied Artificial Intelligence*, koji je odobren u okviru programa Erasmus+ (call 2022 Round 1) i sprovodi se na Prirodno-matematičkom fakultetu. Projekat je počeo 1/9/2022. i traje 24 mjeseca.

Rukovodilac je projekta sa industrijskim partnerom Roadguard AS, Norveška, u okviru koga se izrađuje jedan master rad u formi prakse u punom radnom vremenu na Prirodno-matematičkom fakultetu.

Spisak publikacija je na linku francuskog nacionalnog arhiva: https://cv.archives-ouvertes.fr/igor-jovancevic a dostavljen je i u nastavku.

Bibliografija

Q1 Radovi u međunarodnim naučnim časopisima

Igor Jovančević, Huy-Hieu Pham, Jean-José Orteu, Rémi Gilblas, Jacques Harvent, Xavier Maurice, Ludovic Brèthes. 3D Point Cloud Analysis for Detection and Characterization of Defects on Airplane Exterior Surface. Journal of Nondestructive Evaluation, Springer Verlag, 2017, 36(4), pp.74. DOI 10.1007/s10921-017-0453-1

https://link.springer.com/article/10.1007/s10921-017-0453-1

https://www.scimagojr.com/journalsearch.php?q=21179&tip=sid&clean=0

Q2 Radovi u međunarodnim naučnim časopisima

Ghimire Pamir, **Igor Jovančević**, and Jean-José Orteu. 2021. "Learning Local Descriptor for Comparing Renders with Real Images" Applied Sciences 11, no. 8: 3301. DOI 10.3390/app11083301

https://doi.org/10.3390/app11083301

https://www.scimagojr.com/journalsearch.php?q=21100829268&tip=sid&clean=0

Q3 Radovi u međunarodnim naučnim časopisima

Abdelrahman G. Abubakr, **Igor Jovančević**, Nour Islam Mokhtari, Hamdi Ben Abdallah, Jean-José Orteu, "Learning deep domain-agnostic features from synthetic renders for industrial visual inspection," J. Electron. Imag. 31(5) 051604 (20 June 2022).
https://doi.org/10.1117/1.JEI.31.5.051604
https://www.scimagojr.com/journalsearch.php?q=25978&tip=sid&clean=0

Igor Jovančević, Stanislas Larnier, Jean-José Orteu, Thierry Sentenac, "Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot," Journal of Electronic Imaging 24(6), 061110 (30 November 2015). DOI 10.1117/1.JEI.24.6.061110
https://doi.org/10.1117/1.JEI.24.6.061110

https://www.scimagojr.com/journalsearch.php?q=25978&tip=sid&clean=0

Hamdi Ben Abdallah, Jean-José Orteu, **Igor Jovančević**, Benoit Dolives, "Three-dimensional point cloud analysis for automatic inspection of complex aeronautical mechanical assemblies," Journal of Electronic Imaging 29(4), 041012 (25 April 2020). DOI 10.1117/1.JEI.29.4.041012

https://doi.org/10.1117/1.JEI.29.4.041012

https://www.scimagojr.com/journalsearch.php?q=25978&tip=sid&clean=0

Ivan Mikhailov, **Igor Jovancevic**, Nour Islam Mokhtari, and Jean-Jose Orteu "Classification using a three-dimensional sensor in a structured industrial environment," Journal of Electronic Imaging 29(4), 041008 (17 February 2020). DOI 10.1117/1.JEI.29.4.041008

https://doi.org/10.1117/1.JEI.29.4.041008

https://www.scimagojr.com/journalsearch.php?q=25978&tip=sid&clean=0

Q5 Radovi u međunarodnim naučnim časopisima koji nisu indeksirani na SCI/SCIE/SSCI/A&HCI listama

Ben Abdallah Hamdi, **Igor Jovančević**, Jean-José Orteu, and Ludovic Brèthes. 2019. "Automatic Inspection of Aeronautical Mechanical Assemblies by Matching the 3D CAD Model and Real 2D Images" Journal of Imaging 5, no. 10: 81. DOI 10.3390/jimaging5100081

https://doi.org/10.3390/jimaging5100081

https://www.scimagojr.com/journalsearch.php?q=21100900151&tip=sid&clean=0

(Časopis je Q2 - INDEKSIRAN NA WoS ESCI listi:

https://www.mdpi.com/journal/jimaging/indexing)

Igor Jovančević, Huy-Hieu Pham, Jean-José Orteu, Rémi Gilblas, Jacques Harvent, Xavier Maurice, Ludovic Brèthes. Détection et caractérisation de défauts de surface par analyse des nuages de points 3D fournis par un scanner. Instrumentation, Mesure, Métrologie, Lavoisier, 2017, 16 (1-4), p.261-282.

https://hal.archives-ouvertes.fr/hal-01660998

http://www.lavoisier.eu/books/physics/instrumentation-mesure-metrologie-volume-16-n-1-4-janvier-decembre-2017/voure-li/description-9782746248281

https://www.scimagojr.com/journalsearch.php?q=21100201515&tip=sid&clean=0

K2 Naučni rad na međunarodnom naučnom skupu (štampano u cjelini)

Mokhtari, N.I., **Jovančević, I.**, Ben Abdallah, H., Orteu, J.J. (2023). Image Classification Applied to the Problem of Conformity Check in Industry. In: , et al. 17th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2022). SOCO 2022. Lecture Notes in Networks and Systems, vol 531. Springer, Cham. https://doi.org/10.1007/978-3-031-18050-7_33
https://link.springer.com/chapter/10.1007/978-3-031-18050-7_33
Knjiga: https://www.springer.com/series/15179

Igor Jovančević, Jean-José Orteu, Thierry Sentenac, and Rémi Gilblas "Automated visual inspection of an airplane exterior", Proc. SPIE 9534, Twelfth International Conference on Quality Control by Artificial Vision 2015, 95340Y (30 April 2015)

https://doi.org/10.1117/12.2182811

Jovančević, I.; Viana, I.; Orteu, J.; Sentenac, T. and Larnier, S. (2016). Matching CAD Model and Image Features for Robot Navigation and Inspection of an Aircraft. In Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods - ICPRAM, ISBN 978-989-758-173-1; ISSN 2184-4313, pages 359-366. DOI: 10.5220/0005756303590366

https://www.scitepress.org/Link.aspx?doi=10.5220/0005756303590366

I. Jovančević, A. Arafat, J. Orteu and T. Sentenac, "Airplane tire inspection by image processing techniques," *2016 5th Mediterranean Conference on Embedded Computing (MECO)*, 2016, pp. 176-179, DOI: 10.1109/MECO.2016.7525733.

https://ieeexplore.ieee.org/document/7525733

Hamdi Ben Adallah, Jean-José Orteu, Benoit Dolives, and **Igor Jovančević** "3D point cloud analysis for automatic inspection of aeronautical mechanical assemblies", Proc. SPIE 11172, Fourteenth International Conference on Quality Control by Artificial Vision, 111720U (16 July 2019); DOI 10.1117/12.2521715

https://doi.org/10.1117/12.2521715

Abdelrahman G. Abubakr, **Igor Jovancevic**, Nour Islam Mokhtari, Hamdi Ben Abdallah, and Jean-José Orteu "On learning deep domain-invariant features from 2D synthetic images for industrial visual inspection", Proc. SPIE 11794, Fifteenth International Conference on Quality Control by Artificial Vision, 1179418 (16 July 2021); DOI 10.1117/12.2589040

(nagrada za najbolji rad na konferenciji)

https://doi.org/10.1117/12.2589040

Assya Boughrara, **Igor Jovancevic**, Hamdi Ben Abdallah, Benoit Dolives, Mathieu Belloc, and Jean-José Orteu "Inspection of mechanical assemblies based on 3D deep learning approaches", Proc. SPIE 11794, Fifteenth International Conference on Quality Control by Artificial Vision, 1179407 (16 July 2021); DOI 10.1117/12.2588986

https://doi.org/10.1117/12.2588986

Hamdi Ben Abdallah, **Igor Jovančević**, Jean-José Orteu, Benoit Dolives, Ludovic Brèthes. Automatic inspection of aeronautical mechanical assemblies using 2D and 3D computer vision. NDT AEROSPACE 2019 - 11th Symposium on NDT in Aerospace, Nov 2019, Paris-Saclay, France.

https://www.ndt.net/search/docs.php3?id=25051

K4 Saopštenje na međunarodnom naučnom skupu

Assya Boughrara, **Igor Jovančević**, Hamdi Ben Abdallah, Benoit Dolives, Mathieu Belloc, Jean-José Orteu. Inspection d'assemblages mécaniques par une approche Deep Learning 3D : résultats préliminaires. ORASIS 2021 - 18ème journées francophones des jeunes chercheurs en vision par ordinateur, Centre National de la Recherche Scientifique [CNRS], Sep 2021, Saint Ferréol, France. 9 p.

https://orasis2021.sciencesconf.org/

Hamdi Ben Abdallah, **Igor Jovančević**, Jean-José Orteu, Ludovic Brèthes, Benoît Dolives. Inspection automatisée d'assemblages mécaniques aéronautiques par vision 2D/3D en exploitant la maquette numérique. OPTIQUE TOULOUSE 2018, Jul 2018, Toulouse, France.

https://www.sfoptique.org/pages/congres-optique/optique-toulouse-2018/

Igor Jovančević, H.-H. Pham, Jean-José Orteu, R. Gilblas, J. Harvent, X. Maurice, L. Brèthes. Détection et caractérisation de défauts par analyse des nuages de points 3D fournis par un scanner. 15ème Colloque Méthodes et Techniques Optiques pour l'Industrie, Le Mans (France), 20-24 mars 2017., Le Mans, France

https://www.club-cmoi.com/colloque-2017

Igor Jovančević, Jean-José Orteu, Thierry Sentenac, Rémi Gilblas. Automated visual inspection of an airplane exterior. Congrès de l'Ecole Doctorale Systèmes (EDSYS), May 2015, Toulouse, France.

https://edsys15.sciencesconf.org/

Igor Jovančević, Jean-José Orteu, Thierry Sentenac, Rémi Gilblas. Inspection d'un aéronef à partir d'un système multi-capteurs porté par un robot mobile. 14ème Colloque Méthodes et Techniques Optiques pour l'Industrie, Nov 2015, Lannion, France.

https://www.club-cmoi.com/2015

A. Fardana, S. Jain, **Igor Jovančević**, Y. Suri, C. Morand, N.M. Robertson. Controlling a Mobile Robot with Natural Commands based on Voice and Gesture. IEEE International Conference on Robotics and Automation (ICRA) - Workshop on Human Robot Interaction for Assistance and Industrial Robots, May 2013, Karlsruhe, Germany.

https://ieeexplore.ieee.org/xpl/conhome/6615630/proceeding