UNIVERZITET CRNE GORE
PRIRODNO-MATEMATIČKI FAKULTET
DOKTORSKE STUDIJE

VIJEĆU PRIRODNO-MATEMATIČKOG FAKULTETA

Predmet: Prijava teme doktorske disertacije i predlog Komisije za ocjenu podobnosti teme

U skladu sa članom 33, stav 4, Pravila doktorskih studija, doktorand **mr Kosta Pavlović** je 15. 02. 2021. god. Vijeću Prirodno-matematičkog fakulteta podnio **Prijavu teme doktorske disertacije** (**PD Obrazac** sa pratećom dokumentacijom) pod naslovom **Umetanje vodenog žiga u digitalne audio signale korišćenjem dubokih neuronskih mreža**.

Komisija za doktorske studije PMF-a je na sjednici održanoj 15. 02. 2021. god. razmatrala formalne uslove dostavljene prijave sa stanovišta neophodnih podataka i ispunjavanja uslova za prijavu teme i podnosi Vijeću

P R E D L O G

sastava **Komisije za ocjenu podobnosti teme**:

1. **Dr Slobodan Đukanović**, redovni profesor Elektrotehničkog fakulteta Univerziteta Crne Gore (naučna oblast: mašinsko učenje)

2. **Dr Milenko Mosurović**, redovni profesor Prirodno-matematičkog fakulteta Univerziteta Crne Gore (naučna oblast: vještačka inteligencija)

3. **Dr Igor Đurović**, redovni profesor Elektrotehničkog fakulteta Univerziteta Crne Gore (naučna oblast: obrada signala)

Podgorica, 15. 02. 2021. god.

ZA KOMISIJU ZA DOKTORSKE STUDIJE

Doc. dr Goran Popivoda

PRIJAVA TEME DOKTORSKE DISERTACIJE

OPŠTI PODACI O DOKTORANDU	
Titula, ime i prezime	MSc Kosta Pavlović
Fakultet	Prirodno-matematički fakultet
Studijski program	Doktorske studije
Broj indeksa	1/2018
Ime i prezime roditelja	Jovan Pavlović
Datum i mjesto rođenja	08.05.1994, Berane, Crna Gora
Adresa prebivališta	Ul. Generala Anđelića br. 3/A Kolašin
Telefon	069-528-960
E-mail	kosta@ucg.ac.me
BIOGRAFIJA I BIBLIOGRAFIJA	
Obrazovanje	❑ 01.09.2016-19.10.2018 master studije Univerzitet Crne Gore, Prirodno-matematički fakultet, prosječna ocjena: 10.00 ❑ 05.09.2016-10.07.2016 specijalističke studije Univerzitet Crne Gore, Prirodno-matematički fakultet, prosječna ocjena: 10.00 ❑ 07.09.2012-30.06.2015 osnovne studije Univerzitet Crne Gore, Prirodno-matematički fakultet, prosječna ocjena: 10.00 ❑ 01.09.2008-20.05.2012 gimnazija, SMŠ Braća Selić Kolašin, prosječna ocjena: 5.00
Radno iskustvo	❑ 03.11.2016-danas Univerzitet Crne Gore, Prirodno-matematički fakultet, saradnik u nastavi ❑ 10.10.2016-31.12.2019 Datum Solutions doo, konsultant
Popis radova	1. Kosta Pavlović, Slavko Kovačević, Igor Đurović, *Speech watermarking using Deep Neural Networks*, 28th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2020 2. Savo Tomovic, Kosta Pavlovic, Milija Bajceta, *Aligning document layouts extracted with different OCR engines with clustering approach*, Egyptian Informatics Journal, 2020 3. Savo Tomovic, Kosta Pavlovic, *Long life learning system for document understanding: Document understanding in cognitive manner*, ISBN-13: 978-6138921714, Scholars' Press, January, 2020 4. Savo Tomovic, Kosta Pavlovic, *Cognitive Approach in Document Indexing*, Eastern European Journal for Regional Studies (EEJRS), Volume 4, 2018 5. Kosta Pavlovic, Aleksandar Popovic, *Jezik MetaR*, Informacione Tehnologije, Žabljak, 2018

	6. Kosta Pavlović, Goran Šuković, *Deep Learning Techniques for Classification of Handrwitten Digits*, Informacione Tehnologije, Žabljak, 2017
NASLOV PREDLOŽENE TEME	
Na službenom jeziku	Umetanje vodenog žiga u digitalne audio signale korišćenjem dubokih neuronskih mreža
Na engleskom jeziku	Digital audio watermarking using deep neural networks

Obrazloženje teme

Intelektualnom svojinom smatra se bilo kakva duhovna tvorevina u nauci ili umjetnosti, bilo u pisanom, govornom ili nekom drugom obliku i nju je, zbog njene nematerijalne prirode, mnogo teže zaštititi od tradicionalne svojine. Zaštita intelektualne svojine u savremenom, digitalnom svijetu sve više dobija na značaju jer je ekspanzijom Interneta krađa intelektualne svojine postala učestalija. Takođe, pojavom dubokih neuralnih mreža javila se mogućnost generisanja vještačkih podataka (takozvanih dipfejkova), koji imaju za cilj stvaranje falsifikata, širenje dezinformacija, diskreditacije javnih ličnosti, krađu identiteta itd. Kreatori audio sadržaja, bilo da se radi o muzičarima, glumcima, političarima ili nekim drugim ljudima koji nijesu javne ličnosti na meti su ovakvih napada i neophodno je da njihova prava budu zaštićena. Iz ovih razloga, umetanje vodenog žiga u audio signale potrebno je ne samo u cilju zaštite intelektualne svojine, već i u cilju zaštite osjetljivih informacija dostupnih u vidu javnih govora i sl.

Pregled istraživanja

Umetanje vodenog žiga u audio signale je aktivna oblast istraživanja više od dvadeset godina. Ono predstavlja proces u kojem se analogni, odnosno digitalni audio signali označavaju upotrebom vodenog žiga kako bi se sačuvala autorska prava i autentičnost. Vodeni žig koji se umeće u signal nosioc informacije najčešće je binarna poruka određene dužine i ne mora imati precizno definisan oblik, već može biti generisana na slučajan način.

Prvi radovi u ovoj oblasti preuzimali su ideje iz nešto ranije razvijene oblasti umetanja vodenih žigova u slike [1], da bi se kasnije razvijali pristupi isključivo vezani za audio signale. Najčešći način za podjelu ovih metoda je na osnovu domena u kojem se vrši umetanje vodenog žiga. Umetanje može biti izvršeno u vremenskom domenu [2], [3], ovo su jednostavne i efikasne metode koje direktno modifikuju odbirke signala. Međutim, češće se umetanje sprovodi u nekom od transformacionih domena, dobijenih primjenom diskretne Furijeove transformacije (DFT)[4], diskretne kosinusne transformacije (DCT)[5] ili diskretne vejvlet transformacije (DWT)[6]. Razlog ovome je što vlada uvjerenje da je u transformacionim domenima obično lakše ispuniti zahtjeve postavljene prema sistemima za umetanje vodenog žiga. Međutim, u literaturi nedostaju precizno navedeni razlozi iz kojih je jedna od transformacija bolja za korišćenje od drugih u ovom slučaju, a i prednosti transformacionih domena nad vremenskim takođe nijesu na valjan način potpomognute činjenicama. Preporod mašinskog učenja doveo je do toga da se i ove tehnike počnu koristiti, najprije u procesu ekstrakcije vodenog žiga [6], [7], a kasnije, pojavom dubokog učenja, ovi algoritmi počeli su se koristiti i za umetanje vodenih žigova u digitalne signale [8].

Tehnike za umetanje vodenog žiga dijele se i na informisane i neinformisane. Informisane procedure koriste znanje o signalu nosiocu i njegovim svojstvima u procesu umetanja i ekstrakcije vodenog žiga, dok za neinformisane važi suprotno. Takođe, postoji podjela na slijepe

i ne-slijepe tehnike ekstrakcije vodenih žigova. Ne-slijepe tehnike koriste izvorni oblik signala nosioca, kako bi u signalu sa umetnutim vodenim žigom detektovali vodeni žig. Slijepe tehnike, koje ne zahtijevaju ovu informaciju imaju poželjnije karakteristike, pa se u posljednje vrijeme češće razvijaju i koriste u procesu detekcije. U procesu umetanja vodenih žigova češće se srijeću informisane tehnike.

Dva su osnovna zahtjeva koja se postavljaju svakom sistemu za umetanje vodenog žiga u audio signal. Prvi je da umetanje vodenog žiga u audio signal ne utiče na kvalitet signala, tj. da umetanje vodenog žiga bude nečujno, a drugi, da je iz signala sa umetnutim vodenim žigom moguće rekonstruisati sami vodeni žig i time potvrditi autentičnost signala. Kako su ova dva zadatka suprotstavljeni, prilikom kreiranja ovih sistema potrebno je napraviti kompromis između njih.

Od sistema za umetanje vodenog žiga se dodatno traži da je rekonstrukcija vodenog žiga moguća čak i u situacijama kada je signal sa vodenim žigom izobličen ili oštećen, bilo zbog nestručnog ili malicioznog rukovanja ili, u nekim slučajevima, zbog manipulacija koje se nad signalom vrše zbog potreba skladištenja (kompresija sa gubicima) ili prenosa signala odgovarajućim komunikacionim medijumom (A/D ili D/A konverzija). Sistem treba da bude otporan i na promjene u signalu izazvane primjenom različitih efekata, koji su moguće sastavni dio nekog većeg sistema čiji je sistem za umetanje vodenog žiga dio. Sve ove operacije, bile one maliciozne ili ne, sa tačke gledišta sistema da dodavanje vodenog žiga mogu se smatrati napadima. Sistem za umetanje vodenog žiga koji zadovoljava prethodno opisani zahtjev kvalifikuje se kao robustan. Robustnosti se obično daje najveći prioritet kada se dizajnira sistem za umetanje vodenog žiga ukoliko nije moguće ispuniti sve zahtjeve. U literaturi je dizajniran veliki broj napada na ove sisteme, od jednostavnih do veoma sofisticiranih [9], [10]. Ovo je jedan od glavnih razloga zbog kojih robustnost predstavlja najveći izazov za watermarking sisteme.

Mjera očuvanja kvaliteta signala nosioca je takođe veoma važna karakteristika sistema za umetanje vodenog žiga. Kvalitet audio signala može se mjeriti subjektivnim testovima gdje grupa obučenih ljudi upoređuje i ocjenjuje kvalitet audio snimaka na osnovu predefinisane skale. Pošto su ovi subjektivni testovi veoma zahtjevni za izvođenje, jer oduzimaju dosta vremena i potrebno je da ih izvode osobe obučene za taj zadatak, uvedene su i objektivne mjere očuvanja kvaliteta audio signala [11], kao što su odnos signal-šum ili PESQ [12] za kvalitet govora, koje procjenjuju kvalitet procesuiranog signala njegovim numeričkim poređenjem sa originalom ili u određenim slučajevima čak i bez pristupa originalnom snimku. Postoje i druge mjere za ocjenu performansi sistema za umetanje vodenog žiga, a dizajniranje novih je aktivna oblast istraživanja [13].

Cilj i hipoteze

Ciljevi ovog istraživanja su:

- Prikupljanje i organizacija respektabilnog skupa podataka za testiranje sistema za umetanje vodenog žiga u audio signale.
- Pronalazak duboke neuronske mreže, slične U-Net enkoderu [14], koja će vršiti umetanje vodenih žigova, u obliku binarnih poruka, u audio signale, bez narušavanja njihovog kvaliteta.
- Pronalazak duboke neuronske mreže - dekodera, koji će vršiti ekstrakciju vodenih žigova iz audio signala.
- Dizajniranje slojeva za simulaciju napada na audio watermarking sisteme.
- Pronalazak domena u kojem je najpogodnije vršiti pomenute transformacije signala.

- Osmišljavanje procesa obučavanja kako bi postigli konvergenciju konstruisanog sistema.

Postavljene hipoteze su:

- Umetanje vodenog žiga u audio signale je moguće korišćenjem dubokih neuronskih mreža.
- Upotrebom dekoderske mreže moguće je izvršiti ekstrakciju vodenog žiga iz signala nosioca.
- Enkoder mreža će održati kvalitet audio signala na visokom nivou.
- Dekoder će moći da izvrši ekstrakciju vodenog žiga u situacijama kada je signal oštećen određenim vrstama napada.
- Dekoder mreža neće vršiti detekciju pogrešnog vodenog žiga.

Materijali, metode i plan istraživanja

Prvi korak u istraživanju je prikupljanje skupa podataka nad kojima će se vršiti obučavanje dubokih neuronskih mreža i nad kojima će cjelokupni sistem biti testiran. Korpus govornih signala prikupljen je kod Skupštine Crne Gore. Preuzeti su govori sa 90 sjednica održanih tokom 2016., 2017., 2018. i 2019. godine u ukupnom trajanju od oko 868 časova nakon eliminisanja intervala tišine. Ovaj skup može biti proširen vještački generisanom muzikom i zvukom ili određenim korpusima podataka dostupnim na Internetu.

Radi boljeg i efikasnijeg prezentovanja signala neuronskoj mreži i njenim konvolucionim slojevima, neophodno je izvršiti odgovarajuću obradu i preprocesiranje signala. Tradicionalni pristupi koriste vremensko-frekvencijsku obradu signala i tom prilikom se mogu vršiti različite transformacije signala, što će biti polazna tačka ovog istraživanja kako bi se otkrio domen u kojem je najpogodnije rješavati postavljene zadatke. Spektrogram kao najčešći način reprezentacije audio signala u ovom slučaju ne može biti idealan izbor zbog toga što je njegova inverzija gotovo nemoguća. Zbog toga se upotreba kratkotrajne Furijeove transformacije (STFT) nameće kao prvi izbor. Međutim, treba preispitati da li je potrebno uopšte vršiti transformacije signala, jer su tradicionalne konvolucione mreže najbolje rezultate ostvarile u radu sa realnim signalima, što takođe predstavlja mogući pravac istraživanja.

Arhitektura sistema će se sastojati od najmanje dvije neuronske mreže, enkodera i dekodera. Enkoder će biti mreža slična U-Net mreži [14] i sastojaće se od niza konvolucionih slojeva. Konvolucioni slojevi vrše ekstrakciju atributa eliminacijom redudantnih informacija čime se smanjuje dimenzija ulaza. Zadatak ove mreže će biti da ovim smanjenjem ulaznih dimenzija signala, odnosno njegove vremensko-frekvencijske reprezentacije, pronađe određeni nestandardni domen u kojem će biti umetnut vodeni žig u obliku binarne poruke. Dekoder mreža će biti vrsta klasifikatora, koja će se sastojati od konvolucionih i potpuno povezanih slojeva i koja će se nadovezati na enkoder mrežu i pokušati da iz signala u koji je umetnut vodeni žig detektuje svaki njegov bit.

U poznatoj literaturi predložen je i opisan veoma veliki broj napada na audio watermarking sisteme. Ova oblast i dalje je predmet istraživanja. Nove vrste napada se neprekidno dizajniraju i zbog toga se prilikom kreiranja audio watermarking sistema mora odabrati ograničen skup napada i sistem se pripremati i testirati tako da bude otporan barem na te napade. Kako za robustnost ne postoje definisane mjere kao za procjenu kvaliteta audio signala, na dizajnerima audio watermarking sistema je da se potrude taj skup napada na koje njihov sistem nije otporan učine što manjim.

Napadi se razlikuju u zavisnosti od toga kakav efekat pokušavaju ostvariti nad procesom detekcije. Postoje napadi koji pokušavaju da onemoguće detekciju vodenog žiga, napadi koji izazivaju detekciju pogrešnog vodenog žiga i napadi koji dovode do neautorizovane detekcije umetnutog vodenog žiga. Vrsta napada od interesa bira se u skladu sa planiranom namjenom sistema za umetanje vodenih žigova. U prvom dijelu istraživanja sistem će biti obučavan da se suprotstavi napadima koji pokušavaju dovesti do toga da dekoder ne može pravilno detektovati umetnuti vodeni žig. Takođe, u početku se neće pretpostavljati nikakvo znanje o samom vodenom žigu prilikom dizajniranja napada, kao ni o samom postupku umetanja vodenog žiga. U kasnijim fazama istraživanja proširivaćemo ovaj skup i uvodićemo druge tipove napada.

U nekim situacijama, napadač može doći u posjed skupa primjeraka watermarkovanih i originalnih signala. Ovo se može iskoristiti za obučavanje posebne neuronske mreže koja bi na osnovu ovih primjeraka pokušala da nauči kako da onemogući detekciju vodenih žigova u svim audio snimcima.

Napadima se nekada pokušava dovesti do detekcije pogrešnog vodenog žiga. Ovo se ostvaruje tako što se estimira vodeni žig u jednom signalu, a zatim se on iskopira u drugi signal nosioc. Od ovakvog napada se sistem može zaštititi uvođenjem korelacije između vodenog žiga i signala nosioca, a sama estimacija vodenih žigova može se vršiti primjenom dubokih neuronskih mreža.

Procedura obučavanja ovakvog sistema mora biti strogo kontrolisana iz razloga što enkoder i dekoder imaju suprotstavljene zadatke, pa može doći do toga da jedna mreža nadvlada drugu i onemogući njenu konvergenciju tj. obučavanje. Funkcija gubitka ovakvog sistema mora biti pažljivo osmišljena i regulisana tokom obučavanja, kako bi čitav sistem iskovergirao. Pored samih mjera za preciznost rekonstrukcije originalnog signala, odnosno očuvanja kvaliteta signala (zadatak enkodera) i mjere za preciznost ekstrakcije vodenih žigova (zadatak dekodera), u funkciju gubitka eventualno treba dodati i odgovarajuće mjere korelacije signala i vodenog žiga kako bi se postigao veći stepen robustnosti i olakšao zadatak dekoder mreži.

Očekivani naučni doprinos

Primjena dubokih neuronskih mreža je jedan od najmodernijih pristupa za umetanje vodenih žigova u digitalne signale, kao i za njihovu detekciju. Očekuje se da se primjenom ovih tehnika dostignu izuzetno visoke mjere očuvanja kvaliteta signala (SNR iznad 60dB, PESQ iznad 4). Takođe, očekuje se da se postigne visok nivo preciznosti detekcije vodenih žigova u audio signalima od iznad 90% i otpornost na veliki broj različitih napada, između ostalih i napada desinhronizacije koji predstavljaju najveći problem tradicionalnim pristupima za umetanje vodenih žigova.

Spisak objavljenih radova kandidata

1. Kosta Pavlović, Slavko Kovačević, Igor Đurović, *Speech watermarking using Deep Neural Networks*, 28th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2020
2. Savo Tomovic, Kosta Pavlovic, Milija Bajceta, *Aligning document layouts extracted with different OCR engines with clustering approach*, Egyptian Informatics Journal, 2020
3. Savo Tomovic, Kosta Pavlovic, *Long life learning system for document understanding: Document understanding in cognitive manner*, ISBN-13: 978-6138921714, Scholars' Press, January, 2020
4. Savo Tomovic, Kosta Pavlovic, *Cognitive Approach in Document Indexing*, Eastern European

Journal for Regional Studies (EEJRS), Volume 4, 2018

5. Kosta Pavlovic, Aleksandar Popovic, *Jezik MetaR*, Informacione Tehnologije, Žabljak, 2018
6. Kosta Pavlović, Goran Šuković, *Deep Learning Techniques for Classification of Handrwitten Digits*, Informacione Tehnologije, Žabljak, 2017

Popis literature

1. In-Kwon Yeo, Hyoung Joong Kim, *Modified patchwork algorithm: a novel audio watermarking scheme*, IEEE Transactions on Speech and Audio Processing, 11(4):381–386, July 2003
2. Akira Nishimura, *Audio watermarking based on subband amplitude modulation*, Acoustical Science and Technology, 31(5):328–336, 2010.
3. Masashi Unoki, Ryota Miyauchi, *Robust, blindly-detectable and semi-reversible technique of audio watermarking based on cochlear delay characteristics*, IEICE Transactions on Information and Systems, E98,D(1):38–48, 2015.
4. Pranab Kumar Dhar, Tetsuya Shimamura, *An audio watermarking scheme using discrete fourier transformation and singular value decomposition*, In 2012 35th International Conference on Telecommunications and Signal Processing (TSP), IEEE, July 2012.
5. Chao Yin, Shujuan Yuan, *A novel algorithm for embedding watermarks into audio signal based on DCT*, In Lecture Notes in Electrical Engineering, pp. 683–688, Springer London, 2013.
6. Xiaojuan Xu, Hong Peng, Chengyuan He, *DWT-based audio watermarking using support vector regression and subsampling*, In Applications of Fuzzy Sets Theory, pages 136–144, Springer Berlin Heidelberg, 2007.
7. S. Kirbiz, B. Gunsel, *Robust audio watermark decoding by supervised learning*, In 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, volume 5, pages V–V, 2006
8. Heung-Min Mun, Seung-Hun Nam, Haneol Jang, Dongkyu Kim, Heung-Kyu Lee, *Finding robust domain from attacks: A learning framework for blind watermarking*, Neurocomputing, 337:191–202, April 2019.
9. M. Steinebach, F.A.P. Petitcolas, F. Raynal, J. Dittmann, C. Fontaine, C. Seibel, N. Fatès, L. C. Ferri, *StirMark benchmark: audio watermarking attacks*, Proc. International Conference on Information Technology: Coding and Computing, Las Vegas, NV, USA, pp. 49-54, April 2001.
10. Andreas Lang, Jana Dittmann, Ryan Spring, Claus Vielhauer, *Audio watermark attacks: from single to profile attacks*, Proc. of the 7th workshop on Multimedia & Security, MM&Sec 2005, New York, NY, USA, pp. 39-50, August 2005.
11. Philphos C. Loizou, *Speech Quality Assessment, Multimedia Analysis*, Processing and Communications, Studies in Computational Intelligence, Springer, Berlin, Heidelberg, vol. 346, pp. 623-654 2011.
12. Anthony Rix, John Beerends, Michael Hollier, Andries Hekstra, *Perceptual evaluation of speech quality (PESQ) - A new method for speech quality assessment of telephone networks and codecs*, In: Proc. IEEE Int. Conf. Acoust, Speech, Signal Processing, vol. 2, pp. 749–752, 2001.
13. Patrick Bas and Teddy Furon, *A new measure of watermarking security: The effective key length*, IEEE Transactions on Information Forensics and Security , 8(8):1306–1317, 2013.
14. Olaf Ronneberger, Philipp Fischer, Thomas Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015.
15. Lukas Tegendal, Hhh, Master's thesis, Linkoping University, 2019.