

Универзитет у Београду
Електротехнички факултет

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидаткиње Милане Милошевић.

Одлуком Наставно-научног већа Електротехничког факултета бр. 5026/12-3 од 02.07.2020. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидаткиње Милане Милошевић под насловом

Идентификација говорника у условима емотивног говора

После прегледа достављене Дисертације и других пратећих материјала, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Кандидаткиња Милана Милошевић 28.06.2018. године пријавила је тему за израду докторске дисертације под називом „Идентификација говорника у условима емотивног говора“.

Комисија за студије трећег степена разматрала је 04.07.2018. године предлог теме за израду докторске дисертације и предлог Комисије о оцени подобности теме и кандидата и упутила Наставно –научном већу Електротехничког факултета на усвајање.

Наставно-научно веће Електротехничког факултета је на својој седници од 19.07.2018. године, одлуком бр. 5026/12-2, именovalo Комисију за оцену услова и прихватање теме докторске дисертације у саставу др Горан Квашчев, ванредни професор (Универзитет у Београду-Електротехнички факултет), др Зоран Перић, редовни професор (Универзитет у Нишу – Електронски факултет), др Милош Цветановић (Универзитет у Београду – Електротехнички факултет).

Дана 24.08.2018. године кандидаткиња је полагала јавну усмену одбрану теме пред наведеном Комисијом и добила задовољавајућу оцену.

Наставно-научно веће Електротехничког факултета је на својој седници од 11.09.2018. године усвојило Извештај Комисије за оцену услова и прихватање теме докторске дисертације (Одлука бр. 5026/12-2).

Веће научних области техничких наука Универзитета у Београду дало је сагласност на предлог теме докторске дисертације (број 61206-4047/2-18 од 24.09.2018. године).

Дана 11.06.2020. године кандидаткиња је предала докторску дисертацију на преглед и оцену.

Дана 16.06.2020. године Комисија за студије трећег степена потврдила је испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације.

На својој седници од 02.07.2020. године Наставно-научно веће Факултета именovalo је Комисију за преглед и оцену докторске дисертације у саставу др Жељко Ђуровић, редовни професор (Универзитет у Београду-Електротехнички факултет), др Горан Квашчев, ванредни професор (Универзитет у Београду – Електротехнички факултет), др Зоран Перић, редовни професор (Универзитет у Нишу – Електронски факултет), (број одлуке 5026/12-3).

На основу члана 101. Статута Универзитета у Београду, члана 74. Статута Универзитета у Београду-Електротехничког факултета и захтева студента, одобрено је продужење рока за завршетак студија до истека троструког броја школских година потребних за реализацију уписаног студијског програма.

1.2. Научна област дисертације

Докторска дисертација Милане Милошевић под насловом „Идентификација говорника у условима емотивног говора“ припада научној области електротехника и рачунарство, а ужој научној области Обрада сигнала и управљање системима, за коју је матичан Електротехнички факултет Универзитета у Београду.

Ментор докторске дисертације је др Жељко Ђуровић, редовни професор Електротехничког факултета Универзитета у Београду. Проф. Др Жељко Ђуровић се дуги низ година бави научноистраживачким радом у области обраде сигнала, стохастичких система и естимације као и примене различитих техника вештачке интелигенције, што је документовано релевантним научним радовима који су наведени приликом пријаве теме докторске дисертације кандидата.

1.3. Биографски подаци о кандидату

Милана Милошевић рођена је 23.05.1988. у Београду. Била је ђак генерације и Вуковац у основној школи, Вуковац у Математичкој гимназији, и освајала је награде на такмичењима из математике, физике и информатике и представљала Србију на Међународној Олимпијади из Физике 2007. године у Ирану. Електротехнички факултет, одсек за Сигнале и системе, завршила је у року – основне студије са просечном оценом 9.91 и мастер студије са просечном оценом 10. Њен дипломски рад под називом „Сегментација дигиталне слике применом интелигенције колонија“ награђен је трећом наградом од стране удружења БАФА. Активно је

учествовала у раду Факултета као представница студената у Наставно-научном већу, Савету факултета и као Студенткиња продеканка.

Докторске студије је уписала 2012 године на Електротехничком факултету на одсеку Управљање системима и обрада сигнала. Положила је све испите са просечном оценом 10 и тренутно је посвећена изради докторске дисертације. Течно говори енглески, шпански и немачки језик.

Милана Милошевић усмерена је на истраживања у области машинске обраде говора – детекцију емотивног стања говорника и препознавање говорника у условима емотивног говора. Област њеног интересовања су примене ових истраживања у системима за аутентификацију, безбедносним системима, паметним аутомобилима као и у форензичке сврхе.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација кандидаткиње Милане Милошевић под насловом „Идентификација говорника у условима емотивног говора“ је написана на 126 страна, односно на 134 са прилозима. Организована је у три дела, осам поглавља и садржи 50 слика и 15 табела. Наслови поглавља су: 1. Увод, 2. Препознавање говорника, 3. Модели генерисања сигнала говора, 4. Карактеристике сигнала говора, 5. Класификација и моделирање говорника у условима емотивног говора, 6. Одређивање броја Гаусових мешавина *subtractive* кластеризацијом, 7. Закључак и 8. Коришћена литература. У наставку дисертације су приказани додаци, кратка биографија кандидаткиње, и на крају се налазе прилози.

2.2 Кратак приказ појединачних поглавља

Докторска дисертација кандидаткиње Милане Милошевић састоји се од уводног поглавља, три поглавља прегледног типа, затим два истраживачка поглавља, праћена закључцима, прегледом литературе и два додатка.

У уводном поглављу кандидаткиња Милана Милошевић говори о актуелности и сложености проблема идентификације говорника. Последњих година је ова област поново добила на значају захваљујући развоју паметних, персонализованих система. Осим примене у свакодневном животу, идентификација говорника је од интереса у форензичким истраживањима, телефонским сервисним центрима, центрима за хитне случајеве, у трансферу поверљивих података итд. Такође је у музици једна од тема аутоматске анализе снимака и препознавање певача.

Друго поглавље је посвећено задацима препознавања говорника, изазовима у овој области са посебним освртом на варијабилности у гласу. Феномен емотивног говора ретко је моделован у досадашњем истраживању препознавања говорника. Гледано из перспективе емотивног стања говорника, кључни изазов у системима за препознавање говорника је разлика у емотивном стању говорника приликом обуке система и тренутка када је потребно извршити препознавање. Карактеристике гласа мењају се под утицајем емоција, што у

системима за идентификацију говорника може довести до грешке која даље изазива фрустрацију корисника, а у хитним случајевима, нарочито када су војне и безбедносне примене у питању, може довести до озбиљних последица. Ови примери илуструју важност задатка препознавања говорника и развој поузданих система који на њега могу одговорити. Циљ у дизајну ових система је робусност, а да се при томе користи што мање података за обуку.

У трећем поглављу су описани физиолошки и психо-лингвистички модел генерисања сигнала говора. Описани су механизми генерисања речи од мисаоних форми, као и говорни органи, њихова физиномија и акустика. Психоллингвистички модел говора подразумева моделирање процеса стварања сигнала говора од намере до импулса: шта и како изговорити. То подразумева осмишљавање концепта, састављање реченице и њене интонације. Утицаји су бројни: језик којим се говори, култура и регион одакле потиче говорник, професија, образовање и социјални статус, али и тренутно расположење и емотивно стање говорника. Физиолошки модел говора подразумева сам процес артикулације осмишљене реченице. Елементи овог модела су сви органи који учествују у артикулацији гласа, њихова међусобна интеракција и акустика. Утицај на овај модел превасходно имају биолошка својства говорних органа говорника. Она зависе од старости говорника, пола, али и емоција и тренутног здравственог стања говорника, чиме се даје основа за опис даљег истраживања.

Четврто поглавље даје преглед карактеристика говора, њиховог значаја за препознавање говорника, начина за израчунавање датих карактеристика, као и утицаја емоција на њих. Обухваћене су карактеристике говора у фреквенцијском домену, кепстралном домену, предиктивним коефицијентима, квалитативно итд. У овом истраживању фокус је стављен на карактеристике које квалитетно описују говорнике, као и емоције изражене у говору. Особине идеалних карактеристика за задатак препознавања говорника појављују се природно и често у нормалном говору, лако су мерљиве, разликују се што је више могуће од говорника до говорника, а константне су за истог говорника, не мењају се са временом и не зависе од здравственог стања говорника, не погађа их разуман ниво позадинске буке, нити на њих има утицај канал преноса, не мењају се и поред напора говорника да измени глас, или бар покушај маскирања гласа и имитације на њих нема утицај. На крају поглавља дат је опис утицаја појединачних емоција на вредности и промене карактеристика говорника.

У петом поглављу је дата анализа успешности алгоритама примењиваних у идентификацији говорника са аспекта емоција, пола, количине обучавајућих снимака. Целокупно истраживање спроведено је на четири нивоа: анализа алгоритама за моделирање и класификацију говорника, варијација садржаја говора за обуку модела, варијација конфигурације обуке модела говорника, варијација структуре модела говорника. Анализа алгоритама за моделирање и класификацију говорника обухватила је досада познате методе класификације и њихов начин у препознавању говорника. Теоријски су обрађене технике Гаусових мешавина, скривених Марковљевих модела, машина потпорних вектора, и-вектора, дубоких неуралних мрежа и х-вектора. Експериментално су евалуиране метода стандардног модела Гаусових мешавина, као технике која је основ модерног препознавања говорника, и техника и-вектора која се сматра савременом техником распрострањеном у комерцијалној примени. Варијација садржаја говора за обуку модела говорника односи се на коришћење и емотивног говора за обуку модела говорника - различит број реченица изговорених у одређеном емотивном стању и различит укупан број реченица. Један модел говорника

обучаван је са реченицама неутралног говора, али и са реченицама емотивног говора - радости, беса, туге и страха. Варијација конфигурације модела говорника је моделирање говорника са више од једног модела, такође на основу груписања емотивног говора и касније препознавања не основу овако дистрибуираног модела.

Шесто поглавље описује примену *subtractive* кластеризације за одређивање броја компоненти Гаусових мешавина и почетних кластера у моделу говорника, као и резултате добијене на основу примене ове технике. Резултати су анализирани и дискутовани из различитих аспеката. У овом поглављу изведено је аутоматско одређивање параметара кластеризације на основу вредности улазних параметара. У овом поглављу се налази посебан допринос докторске дисертације који је посвећен статистичкој анализи величина које се појављују као параметри метода за *subtractive* кластеризацију. Као резултат овог истраживања наметнуле су се и препоруке на који начин бирати ове параметре зависно од димензионалности вектора обележја, као и од броја расположивих вектора у обучавајућем скупу.

Резиме тезе, као и сублимирани закључци дати су у седмом поглављу. У закључку су систематизовани мотиви који су постојали на почетку овог истраживања, проблеми који су били у фокусу изучавања, решења која су се искристалисала и ограничења на која је кандидаткиња наишла, а која се могу видети и као оријентири за даљи научни рад. На крају, дат је преглед коришћене литературе.

Први додаток садржи детаље о базама емотивног говора и базама за препознавање говорника. Други додаток садржи независно истраживање о препознавању говорника на основу супрасегменталних карактеристика.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1 Савременост и оригиналност

Иако су системи и алгоритми за обраду говора актуелни од седамдесетих година, последње четири деценије је ова област изузетно жива и богата великим бројем значајних, па често и револуционарних идеја. Разлога за овако активно интересовање многих научника у свету за ову интердисциплинарну област има много, али генерално говорећи, могу се издвојити три најважнија. Први разлог који сведочи о актуелности ове теме јесте све већа количина паметних уређаја и потреба да се оствари интеракција са људима на природан начин.

Други разлог, због кога је ова област већ деценијама врло активна, лежи у технологији која прати све сложеније захтеве. Док су у почетку рачунари били врло слабе процесорске снаге, са малим могућностима за комплексна израчунавања захтевана сложенијим моделима, у међувремену су се, као последица технолошког развоја, развили и рачунари за које су ова израчунавања могућа у прихватљивом времену, што је коначно омогућило експерименталну верификацију неких теоријских концепата. Наведене чињенице сведоче о савремености теме која је предмет докторске дисертације кандидаткиње Милошевић. Она је у самом језгру наведених активности које су усмерене ка повећавању ефикасности и поузданости система за моделирање и идентификацију говорника. Поузданост оваквих система под различитим

условима и за различита стања говорника као што су емоције је кључна за примену оваквих система.

Са друге стране, у последњих десетак година дошло је до експлозије доступности огромне количине података коју је често потребно аутоматски анализирати. Тиме се и појавила још једна реална потреба за применом препознавања говорника, нарочито ако се узме у обзир да су људи инфериорни у односу на машине када је овај задатак у питању.

Феномен емотивног говора ретко је моделован у досадашњем истраживању препознавања говорника. Гледано из перспективе емотивног стања говорника, кључни изазов у системима за препознавање говорника је разлика у емотивном стању говорника приликом обуке система и тренутка када је потребно извршити препознавање. Карактеристике гласа мењају се под утицајем емоција, што систем лако може довести у заблуду. Један од разлога за малу пажњу од стране истраживача је делом и захваљујући недостатку снимака емотивног говора за различите говорнике.

По сазнању аутора тезе и чланова Комисије, у литератури тренутно не постоји метод који сваки од наведених проблема адресира на овако ефикасан начин.

3.2 Осврт на референтну и коришћену литературу

Списак коришћене литературе броји 173 наслова и она се, грубо говорећи, може поделити у три скупа.

Први од њих представља базичну литературу из које се могу сагледати основне структуре система за препознавање говорника, карактеристика и моделирања гласа, алгоритама машинског учења примењених у овој области, као и проблеми који се приликом пројектовања таквих система доминантно појављују. Иако је број референци у овој области изузетно велики, свакако као њихове репрезенте треба поменути следеће наслове: J. H. Hansen and T. Hasan, "Speaker recognition by machines and humans: A tutorial review," IEEE Signal processing magazine, vol. 32, no. 6, pp. 74-99, 2015, D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian Mixture speaker models," IEEE transactions on speech and audio processing, vol. 3, no. 1, pp. 72-83, 1995, D. Reynolds, Gaussian Mixture Models, pp. 827-832. Boston, MA: Springer US, 2015.

Другу групу референтних наслова чине резултати који су скоријих година објављени и који сведоче како о актуелности проблема који се разматра тако и о различитим приступима на који се они могу решавати. У првом реду се мисли на препознавање говорника у неидеалним условима (специфично под условима емоција у говору). У овој групи коришћене и реферисане литературе свакако треба поменути следеће наслове: A. Alarifi, I. Alkurtass, and A. Alsalman, "SVM based Arabic speaker verification system for mobile devices," 2012 International Conference on Information Technology and e-Services, pp. 1-6, 2012, T. Wu, Y. Yang, and Z. Wu, "Improving speaker recognition by training on emotion-added models," in International Conference on Affective Computing and Intelligent Interaction, pp. 382-389, Springer, 2005, D. Li and Y. Yang, "Emotional speech clustering based robust speaker recognition system," in 2009 2nd International Congress on Image and Signal Processing, pp. 1-5, IEEE, 2009, D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur, "X-vectors: Robust DNN embeddings for speaker recognition," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5329-5333, IEEE, 2018.

Трећу групу референтних радова чине управо резултати који се налазе у самој сржи ове докторске дисертације, односно истраживањима и напорима да се нађу ефикасне методе за естимацију броја Гаусових мешавина употребом *subtractive* кластеризације. Из ове групе наслова свакако треба поменути следеће: X.-x. Jing, L. Zhan, H. Zhao, and P. Zhou, "Speaker recognition system using the improved gmm-based clustering algorithm," in 2010 International Conference on Intelligent Computing and Integrated Systems, pp. 482-485, IEEE, 2010, X. Cui, S. Liu, and L. Jia, "An improved method of semantic driven subtractive clustering algorithm," in 2015 IEEE 5th International Conference on Electronics Information and Emergency Communication, pp. 232-235, IEEE, 2015, K. Demirli, S. Cheng, and P. Muthukumaran, "Subtractive clustering based modeling of job sequencing with parametric search," Fuzzy Sets and Systems, vol. 137, no. 2, pp. 235-270, 2003.

Анализирајући списак коришћене и реферисане литературе, Комисија је имала задовољство да констатује да је кандидаткиња користила најактуелнију светску библиографију како у смислу макро-приступа проблему, који подразумева структуру система за идентификацију, избор варијабли које ће бити елементи вектора обележја, тако и у микро-приступу који се односи на детаље какви су избор броја Гаусових компоненти, или кернел функција приликом примене методе потпорних вектора, и томе слично. Такође, Комисија има потребу да истакне још један утисак који је стекла читајући предметну докторску дисертацију. Наиме, област препознавања говора и идентификације говорника у условима његовог емотивног стања је врло актуелна, али истовремено и врло затворена због тога што научно истраживање у овој области подразумева поседовање систематске, обимне и професионално снимљене базе. Оне су обично врло скупе и недоступне и због тога до њих није лако доћи. Тиме је резултат који је у овој докторској дисертацији видљив утолико већи.

3.3. Опис и адекватност примењених научних метода

Током израде своје докторске дисертације кандидаткиња Милана Милошевић је од научних метода истраживања доминантно користила технике анализе, апстракције, спецификације, синтезе и генерализације.

С почетка, улазећи у научну област обраде говорног сигнала и специфично, идентификације говорника, кандидаткиња је била посвећена прегледу постојећих резултата из ове области са посебним акцентом на препознавању емоција у говору. У том контексту је посебна пажња била посвећена анализи и поређењу ефикасности различитих приступа. Овај се сегмент истраживања доминантно ослонио на методе анализе.

Касније, улазећи дубље у област истраживања, кандидаткиња се суочавала са озбиљним ограничењима у примени појединих техника, недостацима у коришћењу појединих обележја, нумеричким теретом и рачунарском неизводљивошћу појединих алгоритама, те се стога, користећи технику синтезе, посветила поправкама и решењима која неке од наведених проблема може да реши. У том смислу, свакако треба поменути њен допринос у светлу примене технике *subtractive* кластеринга, где је вршећи статистичку анализу минимума и максимума Еуклидског одстојања у коначном скупу тачака произвољне димензије, синтезом и генерализацијом дошла до предлога за избор параметара који дефинишу ову врсту кластеринга.

Коначно, техникама спецификације, анализе и синтезе, кандидаткиња је пажљивом организацијом експеримената, уређивањем садржаја и обима обучавајућих скупова, бирајући структуре класификатора које су понајвише примерене проблему идентификације говорника

у условима емотивног говора, постављајући резерве према базама снимљеног говора, зависно од условима под којима су формиране, систематизовала резултате и дала јасне препоруке о томе на који начин треба пројектовати један савремени, ефикасан систем за идентификацију говорника у присуству емоција. Ове препоруке садрже, након примене методе генерализације, врло опште инструкције какве су обим базе и избор класификатора, али и врло конкретне детаље, добијене техником спецификације, као што су избор броја компоненти у Гаусовим мешавинама или избор параметара у техникама кластеризације.

3.4 Применљивост остварених резултата

Резултати остварени током рада на овој докторској дисертацији су од велике вредности са становишта примењивости, и то на два начина а проистичу из два кључна доприноса која ова дисертација има. Првенствено, исцрпна анализа система за идентификацију говорника у присуству емоција која је приказана у тези је обухватила готово све аспекте који се појављују приликом пројектовања таквог система. Почев од величине и структуре базе, присуство полне и старосне структуре у тој бази, заступљеност појединих емоција, формирање обучавајућег скупа и скупа за верификацију, избор класификатора као и подешавање параметара тих класификатора. Сваки од ових елемената је јасно и систематски анализиран и недвосмислено окарактерисан у овој докторској дисертацији. Сваки потенцијални пројектант таквог система би у резултатима ове анализе могао наћи врло вредне инструкције које би га водиле од нулте тачке формирања базе, до самог краја који се обично завршава формирањем корисничког интерфејса. Без обзира на велики број библиографских јединица доступних у светској литератури, а које се баве предметним истраживањима, није могуће наћи овако систематизовану анализу система за идентификацију говорника у присуству емоција. Посебну вредност овог истраживања даје чињеница да је она симултано спроведена на базама различитих европских језика уз улагање максималних напора да се добијени резултати генерализују.

Са друге стране, Комисија има потребу да се посебно осврне на примењивост резултата који се односи на избор броја компоненти у Гаусовим моделима. Показује се да је то живи проблем који је присутан, не само приликом идентификације говорника, већ у готово свим проблемима класификације случајних вектора, односно у системима за доношење одлука. У литератури постоје различити приступи за решавање овог проблема, али их је мало и углавном су засновани на неким хеуристикама до којих су неки истраживачи дошли, а крију се као сегменти уицијализацији метода за кластеризацију података. Кандидаткиња је проицљиво поставила ограничења примењивости таквих хеуристика, и дошла до закључка да су оне мање или више примењиве зависно од величине и димензионалности вектора у обучавајућем скупу. Применом деликатне статистичке анализе минимума и максимума Еуклидског одстојања између случајних вектора у скупу ограничене кардиналности, а зависно од њихове димензије и њиховог броја, кандидаткиња је дошла до резултата који је примењив за широке класе техника кластеризације, односно одлучивања.

Отуда се очекује да ће предложени метод параметризације *subtractive* приступа у кластеризацији наћи место у сложеним, софистицираним, системима у којима је потребно извршити оптимизацију на основу мале количине улазних података.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Током рада на својој докторској дисертацији кандидаткиња је показала изузетну истраживачку зрелост, која се огледала у неколико аспеката. Као прво, треба поменути њену способности да самостално уочава проблеме, да из доступне литературе препознаје елементе који су од веће или мање важности за њихово решавање. Није занемарљиво ни то да је кандидаткиња показала висок ниво комуникативности са припадницима научне заједнице, што је било од посебног значаја приликом прикупљања говорне базе за различите европске језике. Такође, врло је важно поменути чињеницу да је стекла могућност да критички анализира постојећа решења. Током времена, напредујући кроз рад, испољила је захтевану научну зрелост и оспособљена је за даљи успешан научно-истраживачки рад.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ научних доприноса

Током истраживања и рада на докторској дисертацији кандидаткиња је остварила врло значајне резултате који се могу исказати кроз следеће научне доприносе:

- Извршен је врло детаљан преглед научне области која се бави препознавањем говорника, утицајем емоција на квалитет њихове идентификације, као и карактеризацијом емотивног говорног сигнала, на критичан начин су наведена ограничења и недостаци појединих приступа и на тај начин је отворен простор за развој новог метода који је описан у овој дисертацији.
- Извршена је процена успешности употребе гласа као биометријског податка и система за идентификацију говорника познатих у литератури у условима емотивног говора.
- Установљен је утицај сваке од појединачних емоција радости, беса, туге и страха у односу на неутрално емотивно стање кроз проценат успешности система препознавања.
- Установљен је утицај пола, квалитета снимка и величина базе на резултате система за препознавање говорника у условима емотивног говора.
- Експериментално су упоређени различити модели говорника и последични класификатори: Гаусове мешавине и мел кепстрални коефицијенти, и-вектори и мел кепстрални коефицијенти, као и Гаусове мешавине са бројем Гаусових функција одређеним за сваког од говорника појединачно коришћењем *subtractive* кластеризације.
- Експериментално су упоређени модели различите конфигурације обучени емотивним говором (микс-модел и три-модел) а у зависности од броја тренинг реченица у истим условима тестирања.
- Развијен систем евалуиран је на неколико база емотивног говора, различите величине и језика.
- Исцрпном статистичком анализом изведене су релације које додељују вредности почетним параметрима *subtractive* приступа у кластеризацији, на основу чега је аутоматски одређен број компоненти ГММ за сваког од говорника.

- Експериментално су евалуирани резултати са моделима ГММ чији је број компоненти одређен *subtractive* кластеризацијом.

4.2. Критичка анализа резултата истраживања

Истраживање спроведено и приказано кроз ову докторску дисертацију је од самог почетка имало амбицију да буде свеобухватно и непристрасно. Због тога је у поставци истраживања инсистирано на томе да расположиве говорне базе буду што веће, да буду присутни и мушки и женски говорници, да буде разнолика њихова старосна структура, да базе буду из различитих европских језика, да емотивна стања буду што природнија, да таквих емотивних стања буде што више, да број говорних секвенци по сваком говорнику и у свакој емоцији буде што већи, као и да приступи за идентификацију говорника у условима емоције буду што разноликији како би се стекао један општи увид у стање научног домета, када је овај проблем у питању. Међутим, наишло се на пар ограничења која свакако лимитирају генералност добијених резултата. У жељи да истовремено задржи и генералност анализе али и њен коначни обим, кандидаткиња се определила за рад над базама четири европска језика (српски, руски, енглески и италијански), при чему је још једна база неутралног говора (немачки у швајцарској варијанти) коришћена као репер. Одмах се, грубим освртом на ове базе уочава недостатак њихове неуниформности. Наиме, број говорника није исти, величина изговорених секвенци по говорнику се такође разликује, покривене су различите емоције, при чему је база руског језика са природним емоцијама, док су остале базе снимљене од стране професионалних глумаца. Све ове неуједначености између појединих база су свакако утицале на то да добијени и приказани резултати имају одређену меру несигурности. Комисија сматра да је овај исказ најрелевантнији сегмент критичке анализе предметне докторске дисертације.

4.3. Верификација научних доприноса

У циљу илустрације начина на који је извршена верификација научних доприноса, наводимо списак радова које је кандидаткиња Милана Милошевић публиковала радећи на својој докторској дисертацији:

Категорија радова M23

[1] Nedeljković Ž., **Milošević M.**, Đurović Ž.: Analysis of features and classifiers in emotion recognition systems: Case study of Slavic languages, - Archives of Acoustics, Vol. 45, No. 1, 2019, pp. 129–140 (DOI: 10.24425/aoa.2020.132489).

[2] **Milošević M.**, Nedeljković Ž., Glavitsch U., Đurović Ž.: Speaker Modeling Using Emotional Speech for More Robust Speaker Identification, - Journal of Communications Technology and Electronics, Vol. 64, No. 11, 2019, pp. 1256-1265 (DOI: 10.1134/S1064226919110184).

Категорија радова M33:

[1] Nedeljković Ž., **Milošević M.**, Đurović Ž.: Improved HMM for automatic speech emotion recognition, - Proceedings of XXVII Pacific Voice Conference, Krakow 2020, Vol. 3, pp. 10-14.

- [2] **Milošević M.**, Glavitsch U. "Combining Gaussian mixture models and segmental feature models for speaker recognition", INTERSPEECH, 20 – 24 August 2017, Stockholm, Sweden
- [3] **Milošević M.**, Nedeljković Ž. "Emotional speech databases in Slavic languages – an overview", 26th IAFPA, 9 – 12 July 2017, Split, Croatia
- [4] Nedeljković Ž., **Milošević M.**, Đurović Ž., "Baseline system for speaker recognition - parameter analysis" – IcETRAN, June 5-8, 2017, Kladovo, Serbia
- [5] **Milošević M.**, Glavitsch U., He L. & Dellwo V. "Segmental features for automatic speaker recognition in a flexible software framework", 25th IAFPA, 24 – 27 July 2016, York, UK
- [6] **Milošević M.**, Nedeljković Ž., Đurović Ž., "SVM Classifier for Emotional Speech Recognition in Software Environment SEBAS" – IcETRAN, June 13 – 16, 2016, Zlatibor, Serbia
- [7] **Milošević M.**, Đurović Ž., "Challenges in Emotional Speech Recognition, " – IcETRAN, June 8-11, 2015, Silver Lake, Serbia
- [8] **Milošević M.**, Đurović Ž., "Feature Extraction for Emotion Classification from Speech Signal, " IcETRAN, June 2-5, 2014, Vrnjačka Banja, Serbia

Категорија радова M63:

- [1] **Milošević M.**, Đurović Ž., "Social Signal Processing: Elements of Analysis and Applications, " ETRAN, June 2013, Zlatibor, Serbia

ЗАКЉУЧАК И ПРЕДЛОГ

Докторска дисертација кандидаткиње Милане Милошевић под називом „Идентификација говорника у условима емотивног говора“ је у целини написана у складу са образложењем наведеним у пријави теме и садржи све битне елементе који се захтевају Правилником о докторским студијама Електротехничког факултета Универзитета у Београду.

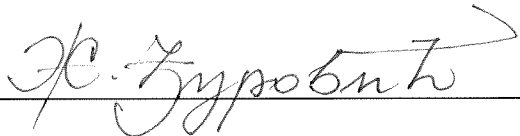
У дисертацији се на систематски и доследан начин анализира проблем идентификације говорника у условима емотивног говора, дефинишу се основни елементи и принципи пројектовања таквог система, и на врло уверљив начин илуструје значај узимања у обзир емотивног говора у креирању модела говорника. Осим тога, дисертација садржи врло вредан резултат иницијализације појединих техника за кластеризацију података, што приказане резултате чини примењивим за читаву класу система за одлучивање.

Кандидаткиња Милана Милошевић је показала способност за самостални научни рад, што потврђује и чињеница је објавила неколико научних радова који су проистекли из истраживања приказаних у оквиру ове тезе, а у којима се појављује као први коаутор. Оцењујући докторску дисертацију, као и чињеницу да је анализирани и решавани проблем актуелан и савремен, као и да садржи научне доприносе, Комисија констатује да је кандидаткиња Милана Милошевић, мастер инжењер електротехнике и рачунарства, испунила све услове предвиђене Законом о високом образовању, Статутом и Правилником о докторским студијама Електротехничког факултета Универзитета у Београду.

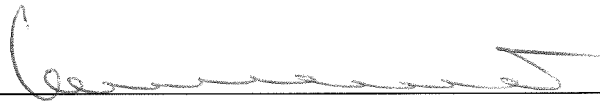
Имајући у виду напред наведено, Комисија са задовољством предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да се докторска дисертација под називом „Идентификација говорника у условима емотивног говора“ кандидаткиње Милане Милошевић прихвати, изложи на увид јавности и упути на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

У Београду, 07.07.2020.

ЧЛАНОВИ КОМИСИЈЕ



Др Жељко Ђуровић, редовни професор
Универзитет у Београду-Електротехнички факултет



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



Др Зоран Перић, редовни професор
Универзитет у Нишу – Електронски факултет