

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата **Јована Галића**, дипломираног инжењера електротехнике - мастера електронике и телекомуникација

Одлуком Наставно-научног већа Електротехничког факултета бр. 5036/16-3 од 25.01.2019. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Јована Галића под насловом

“Препознавање мултимодалног говора засновано на статистичком приступу”.

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

REFERAT

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Кандидат Јован Галић је 13. децембра 2008. године први пут уписао докторске академске студије Електротехнике и рачунарства, модул Телекомуникације, на Електротехничком факултету Универзитета у Београду. Све испите предвиђене наставним планом и програмом докторских студија положио је просечном оценом 9,90 и испунио све nastavne обавезе. Због истека рока предвиђеног за завршетак студија, кандидату је у октобру 2016. одобрен поновни упис уз признавање свих положених испита и испуњених обавеза.

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.

Кандидат је 01.02.2018. године пријавио тему за израду докторске дисертације под насловом “Препознавање мултимодалног говора засновано на статистичком приступу”. За ментора је предложена проф. др Драгана Шумарац Павловић, редовни професор Електротехничког факултета у Београду.

Комисија за студије трећег степена је 06.02.2018. године разматрала предлог теме за израду докторске дисертације и упутила предлог Комисије за оцену подобности теме и кандидата на усвајање Наставно –научном већу Електротехничког факултета.

Наставно-научно веће именovalo је Комисију за оцену услова и прихватање теме докторске дисертације (Одлука бр. 5036/16-1 од 27.2.2018.) у саставу:

1. др Миомир Мијић, редовни професор, Електротехнички факултет у Београду,

2. др Милан Војновић, научни сарадник, Центар за унапређење животних активности, Београд, и
3. др Горан Квапчев, Електротехнички факултет у Београду.

Кандидат је на јавној усменој одбрани предложене теме докторске дисертације, одржаној 07.03.2018. године, успешно положио докторски испит са оценом “задовољно”.

Наставно-научно веће усвојило је Извештај Комисије за оцену услова и прихватање теме докторске дисертације (Одлука бр. 5036/16-2 од 17.4.2018. године). За ментора дисертације именована је Драгана Шумарац Павловић, редовни професор.

Веће научних области техничких наука дало је сагласност на предлог теме докторске дисертације Одлуком бр. 61206-2281/2-18 од 28.05.2018. године.

Дана 27.12.2018. године кандидат је предао докторску дисертацију на преглед и оцену. Комисија за студије трећег степена је 08.01.2019. године потврдила испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације.

На седници одржаној 15.01.2019. године, Наставно-научно веће Факултета именовало је Комисију за преглед и оцену докторске дисертације под насловом “Препознавање мултимодалног говора засновано на статистичком приступу” (број одлуке 5036/16-3 од 25.1.2019. године) у саставу:

1. др Миомир Мијић, редовни професор, Електротехнички факултет у Београду
2. др Јелена Ћертић, доцент, Електротехнички факултет у Београду
3. др Милан Војновић, научни сарадник, Центар за унапређење животних активности у Београду
4. др Жељко Ђуровић, редовни професор, Електротехнички факултет у Београду
5. др Ана Гавровска, доцент, Електротехнички факултет у Београду

1.2. Научна област дисертације

Докторска дисертација кандидата Јована Галића под насловом “Препознавање мултимодалног говора засновано на статистичком приступу” припада научној области Техничке науке - Електротехника, ужа научна област Техничка акустика, за коју је матични факултет Електротехнички факултет Универзитета у Београду. За ментора докторске дисертације одређена је др Драгана Шумарац Павловић, редовни професор Електротехничког факултета Универзитета у Београду која је изабрана у звање редовног професора из уже научне области Техничка акустика. Сви њени објављени радови који је квалификују за ментора су из те области.

1.3. Биографски подаци о кандидату

Јован Галић рођен је 07.11.1981. године у Травнику, Република Босна и Херцеговина. Основну школу завршио је у Бањој Луци са Вуковом дипломом. Средњу електротехничку школу завршио је 2000. године у Бањој Луци. Током школовања је неколико пута учествовао на Републичким такмичењима из Математике. Електротехнички факултет Универзитета у Бањој Луци уписао је 2000. године. Дипломирао је на Одсеку за електронику и телекомуникације у марту 2007. године са просечном оценом положених испита 8,80 (на дипломском испиту оцена 10).

Докторске студије уписао је 2008. године на Електротехничком факултету Универзитета у Београду, модул телекомуникације. Положио је све испите предвиђене студијским

програмом са општим успехом 9,90 и испунио све обавезе у вези са студијским истраживачким радовима. Област истраживања током студија обухватала је аутоматско препознавање говора и нетипичних облика говора (шапат), те дигиталну обраду аудио-сигнала и говорну комуникацију.

Јован Галић је аутор преко 30 научних и стручних радова објављених у међународним часописима и зборницима радова међународних и националних конференција. Ангажован је као рецензент у 2 интернационална часописа (од тога један на SCI листи) и коаутор је једног основног универзитетског уџбеника. Члан је професионалног удружења IEEE. Учесник је неколико пројеката Министарства Науке и Технологије Републике Српске и 2 међународна пројекта.

Радну каријеру је започео на Катедри за телекомуникације Електротехничког факултета Универзитета у Бањој Луци, где и сада ради, у звању вишег асистента.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација Јована Галића је има 131 страну. Садржи 55 слика, 15 табела, 63 једначине и 96 библиографске референце. Дисертација садржи насловну страну на српском и енглеском језику, страну са подацима о ментору и члановима Комисије за преглед и оцену, захвалницу, посвету, кратак резиме дисертације на српском и енглеском језику, списак слика и табела, садржај, 7 тематских поглавља, списак референци, биографију аутора, 7 прилога и неопходне изјаве аутора (изјава о ауторству, изјава о истоветности штампане и електронске верзије и изјаву о коришћењу).

Поглавља дисертације су насловљена на следећи начин: 1. Увод, 2. Увод у препознавање мултимодалног говора, 3. Методологија препознавача говора, 4. Опис експерименталне поставке, 5. Резултати експеримената, 6. Резултати експеримената са мултимодном базом за обуку и 7. Закључак.

2.2. Кратак приказ појединачних поглавља

У уводном поглављу су сагледане перформансе, могућности и недостаци савремених система за аутоматско препознавање говора (тзв. ASR системи). С обзиром да су ASR системи превасходно намењени препознавању говора уобичајеног интензитета (нормалног говора), указано је на значај усавршавања система у погледу препознавања других говорних модалитета, пре свега шапата. Посебно је наглашен значај препознавања шаптавог говора, али са моделима обученим на изговорима нормалног говора (означено као „неусаглашени сценарио“). Дефинисани су основни циљеви у вези са унапређењем перформанси у препознавању шапата базираном на статистичком приступу, односно скривеним Марковљевим моделима (НММ). Такође, дат је опис структуре докторске дисертације.

У Поглављу 2 детаљније је описан појам мултимодалног говора. Приказани су основни параметри битни за аутоматско препознавање говора за 5 говорних модалитета: шапат, тихи говор, нормални говор, гласни говор и вика. Посебно су наведени сличности и разлике између нормалног говора и шапата дефинисаних са акустичког аспекта. Пошто су говорне технологије изразито зависне од језика којем су намењене, детаљно је описан процес креирања говорне базе изолованих речи у српском језику - Whi-Spe, која је коришћена у истраживањима. Такође, дат је и осврт на досадашње резултате у препознавању мултимодалног говора.

У Поглављу 3 приказане су методологије препознавача које су заступљене у многим савременим ASR системима, а које су коришћене у овој дисертацији: скривени Марковљеви модели и метода потпорних вектора (SVM). Посебно су описани основни проблеми код препознавања базираног на НММ приступу: проблем евалуације, декодовања и естимације параметара. Приказани су и алгоритми за екстракцију MFCC и PLP вектора обележја, који се често користе у ASR системима.

У наставку дисертације приказани су самостални резултати истраживачког рада кандидата. Тако су у Поглављу 4 описани експериментална поставка и пратећи алати коришћени у 3 начина препознавања изолованих речи: фонема независних од контекста (монофона), фонема зависних од контекста (трифона) и целих речи. Препознавач базиран на НММ је реализован у програмском окружењу MATLAB и користи скуп алата НТК. Препознавач базиран на SVM је реализован у програмском језику Python. Поред тога, детаљно је описан предлог алгоритма за екстракцију вектора обележја са кепстралним коефицијентима и модификованом фреквенцијском скалом мапирања (тзв. μ -фреквенцијска скала). Развијени алгоритам представља оригинални допринос аутора проблематици препознавања шаптавог говора при обуци на изговорима нормалног говора.

У Поглављу 5 је приказан скуп реализованих експеримената у препознавању нормалног говора и шапата у различитим обука/тест сценаријима. Највећа успешност у неусаглашеном сценарију је добијена за моделе монофона. Експерименти су урађени за случај препознавања зависно од говорника (изговори тестираног говорника се користе у обуци) и независно од говорника (изговори тестираног говорника не учествују у обуци). Анализиран је допринос иницијализације модела коришћењем мануелне и аутоматске анотације успешности препознавања НММ препознавача. Урађена је упоредна анализа перформанси НММ и SVM препознавача. На крају су приказани резултати експеримената са предложеним векторима обележја који користе μ -фреквенцијску скалу. Експерименти су показали приметно побољшање перформанси у препознавању шапата док су истовремено задржане добре перформансе у препознавању нормалног говора. Статистички тестови су потврдили значај побољшања у оба начина препознавања, зависно и независно од говорника.

У Поглављу 6 су приказани резултати експеримената када су за обуку препознавача на располагању изговори оба говорна мода. Извршена је упоредна анализа перформанси НММ препознавача између препознавања са измешаном базом за обуку (Multi-Style) и класификацијом мода говора, зависно и независно од говорника. Анализирана је успешност класификације класификатора базираног на односу сигнал/шум, првом кепстралном коефицијенту и основној фреквенцији говорног сигнала.

Закључак, анализа доприноса истраживања спроведених у докторској дисертацији и правци даљих истраживања наведени су у Поглављу 7.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Савремени ASR системи у контролисаним условима показују изузетно добре перформансе у препознавању нормалног говора. У литератури је показано да се у препознавању читаног текста постиже успешност боља од човекове. Међутим, у пракси су се ови системи показали далеко од идеалних услед различитих фактора који маскирају говорни сигнал (амбијентална бука, реверберација, итд.), различитих емотивних стања говорника и разних модалитета изговора. Из тих разлога се значајан део савремених истраживања усмерава ка препознавању разних модалитета говора. Због своје специфичности у начину продуковања (одсуство глоталних вибрација) шапат од свих говорних модалитета представља највећи изазов са становишта задовољавајуће успешности код препознавања

независно од говорника. На тај начин препознавање шапата се препознаје као актуелна истраживачка тема. Такође, због низа добрих особина у односу на друге технике препознавања, а које се огледају пре свега у значајно мањем времену потребном за обуку и одсуству захтева за великом говорном базом за обуку, примена скривених Марковљевих модела у препознавању говора не губи на актуелности.

Оригиналност рада презентованог у оквиру докторске дисертације се огледа и у развијеном алгоритму за екстракцију вектора обележја који су показали значајну робустност у препознавању шапата. Статистички тестови су потврдили већу успешност нових вектора обележја у односу на конвенционалне попут MFCC и PLP вектора обележја. Такође, чињеница да до писања докторске дисертације у литератури није био доступан развој ASR система базираног на методи потпорних вектора иде у прилог оригиналности ове дисертације. Може се закључити да су истраживања у оквиру ове дисертације резултовала новим и оригиналним системом за препознавање нормалног говора и шапата.

3.2. Осврт на референтну и коришћену литературу

Кандидат је у току израде докторске дисертације широко истражио постојећу доступну стручну литературу. У дисертацији је наведено 96 релевантних референци које на адекватан начин покривају област којом се дисертација бави. Такође, дисертација покрива широк спектар референци, од старијих до савремених. Коришћена литература, између осталог, обухвата научне радове из међународних часописа, са међународних и националних конференција на којима је кандидат први аутор, а који су проистекли из рада на дисертацији.

3.3. Опис и адекватност примењених научних метода

Методологија истраживања у оквиру докторске дисертације састојала се од следећих активности:

- Преглед доступне научне литературе и анализа постојећих техника за машинско препознавање мултимодалног говора,
- Формирање говорне базе неопходне за истраживања – контрола квалитета снимака нормалног говора и шапата уз сегментацију изолованих речи. База у почетној форми садржи 10.000 изговора од 10 говорника (5 женских и 5 мушких) у укупном трајању 2 часа. За ту намену је коришћен софтверски алат Adobe Audition.
- Развој софтвера за препознавање говора – имплементација препознавача је урађена у програмском окружењу MATLAB коришћењем скупа алата НТК.
- Анализа резултата успешности НММ препознавача у иницијалном експерименту за 3 јединице за моделовање и одређивање најпогоднијег броја мешавина.
- Усавршавање система у погледу препознавања шапата помоћу иницијализованих модела са мануелном и аутоматском анотацијом и анализа доприноса успешности препознавања нормалног говора и шапата. За мануелну анотацију изговора је коришћен софтверски алат PRAAT, а за аутоматску НТК.
- Развој алгоритма за екстракцију вектора обележја са кепстралним коефицијентима и μ -фреквенцијском скалом и статистичка анализа успешности истог. Алгоритам и препознавач су развијени у програмском окружењу MATLAB.
- Развој препознавача нормалног говора и шапата базираног на методи потпорних вектора у програмском језику Python.
- Развој и упоредна анализа успешности НММ препознавача са измешаном базом за обуку и препознавача са класификацијом мода говора.

Примењена методологија одговара стандардима научно-истраживачког рада и довела је до остваривања циљева који су постављени пред докторску дисертацију.

3.4. Применљивост остварених резултата

Докторска дисертација се доминантно бави унапређењем перформанси система за препознавање нормалног говора и шапата при обуци на изговорима нормалног говора. Предложени алгоритам екстракције вектора обележја са μ -фреквенцијском скалом мапирања може да има директну примену у ASR системима за препознавање бимодалног говора (комбиновано нормалног говора и шапата) зато што су у односу на MFCC обележја перформансе у препознавању шапата побољшане, а да при том перформансе у препознавању нормалног говора нису деградиране. Пошто су сви савремени системи у пракси намењени препознавању независно од говорника у прилог могућој примени иде и чињеница да су у дисертацији анализирани перформансе оба начина препознавања: зависно и независно од говорника. Нови алгоритам екстракције вектора обележја карактерише једноставност реализације - уместо мел-фреквенцијске скале (која се користи код MFCC вектора обележја) користи се μ -фреквенцијска скала која такође користи логаритамску функцију мапирања. На тај начин комплексност система није повећана. То указује на могућност примене у системима за препознавање континуалног говора, иако је дисертација анализирала успешност препознавања изолованих речи.

Пошто је у пракси количина изговора шапата доступна за обуку углавном ограничена, могућа примена се огледа и у чињеници да је дисертација анализирала раст перформанси у препознавању шапата у зависности од количине изговора у моду шапата расположивих за обуку.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Након прегледа докторске дисертације Комисија процењује да је кандидат Јован Галић показао способност за самосталан научно-истраживачки рад. На основу увида у релевантну литературу из области које покрива дисертација кандидат је показао зрелост да самостално решава практичне проблеме у вези са усавршавањем постојећих система за препознавање говора на српском језику и подизањем нивоа комуникације човек-машина на виши ниво. Резултате истраживања из којих је проистекла дисертација кандидат је као први аутор публиковао у међународним часописима и зборницима међународних и националних конференција.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Основни доприноси који су остварени у оквиру ове докторске дисертације огледају се у следећем:

- Говорне технологије су јако зависне од језика којем су намењене, те се не могу увести и применити за било који језик. Због тога је систематски креирана говорна база Whi-Spe на српском језику која се може користити и у другим врстама истраживања бимодалне експресије говора.
- Развијен је препознавач базиран на статистичком приступу и скривеним Марковљевим моделима. Показано је да између анализираних јединица за моделовање фонеме независни од контекста имају највећу робусност. Одређен је најпогоднији број мешавина у препознавању зависно и независно од говорника.

- Повећање робустности препознавача у значајној мери је остварено коришћењем динамичких обележја, уз редукован број монофона и иницијализацију модела са мануелном (зависно од говорника) и аутоматском анотацијом (независно од говорника). При том су повећане перформансе препознавача и за нормални говор.
- Развијен је нови препознавач (први пут примењен у препознавању шапата) базиран на методи потпорних вектора и извршена анализа перформанси у зависности од типа керна и броја преклапајућих прозора.
- Развијен је алгоритам за екстракцију нових вектора обележја заснованих на кепстралним коефицијентима и модификованом скалом мапирања. Пресликавање је извршено према μ -фреквенцијској скали која је прилагођена шапату на тај начин што комбинује резолуцију мел и линеарне фреквенцијске скале. Експерименти су потврдили повећање робустности НММ-препознавача са новим векторима обележја у оба начина препознавања (зависно и независно од говорника).
- Развијен је бинарни класификатор мода говора (нормалан говор/шапат) базиран на односу сигнал/шум, првом кепстралном коефицијенту и основној фреквенцији говорног сигнала. Између анализираних, најмања грешка класификације је добијена за класификатор базиран на основној фреквенцији.
- Извршена је упоредна анализа перформанси НММ-препознавача базираног на измешаној бази за обуку и класификацији мода говора у зависности од удела изговора шапата расположивих у обуци.

4.2. Критичка анализа резултата истраживања

Увидом у постављене циљеве истраживања, полазне хипотезе и остварене резултате Комисија констатује да је кандидат Јован Галић успешно одговорио на сва значајна питања из проблематике којом се дисертација бави. Кандидат је демонстрирао развој решења за аутоматско препознавање нормалног говора и шапата заснованог на статистичком приступу (НММ). Повећање робустности НММ-препознавача у препознавању шапата без нарушавања успешности препознавања нормалног говора представља значајан научни допринос у области којом се дисертација бави.

Комисија констатује да су научни доприноси остварени у дисертацији објављени у научном часопису категорије M23, као и зборницима конференција међународног и националног значаја.

4.3. Верификација научних доприноса

Списак публикованих радова у којима су презентовани научни доприноси кандидата Јована Галића, у складу са нормативима Министарства просвете и науке Србије, наведен је како следи:

Категорија M23:

1. **J. Galić, B. Popović, D. Šumarac Pavlović:** "Whispered Speech Recognition using Hidden Markov Models and Support Vector Machines," Acta Politechnica Hungarica, Vol.15(5), pp. 11-29, 2018, ISSN 1785-8860, IF=0.909.
DOI: 10.12700/APH.15.5.2018.5.2
2. **Đ. Grozdić, S. Jovičić, D. Šumarac Pavlović, J. Galić, B. Marković:** „Comparison of Cepstral Normalization Techniques in Whispered Speech Recognition“, ADVANCES IN

ELECTRICAL AND COMPUTER ENGINEERING, Vol. 17, No. 1, pp. 21-26, 2017, ISSN 1582-7445, IF=0,459. DOI: 10.4316/AECE.2017.01004

3. B. Marković, **J. Galić**, Đ. Grozdić, S. Jovičić, M. Mijić: „Whispered speech recognition based on gammatone filterbank cepstral coefficients“, JOURNAL OF COMMUNICATIONS TECHNOLOGY AND ELECTRONICS, Vol. 62, No. 11, str. 1255-1261, 2017. ISSN 1064-2269, IF = 0,431. DOI: 10.1134/S1064226917110134
4. B. Marković, **J. Galić**, M. Mijić: „Application of Teager Energy Operator on Linear and Mel Scales for Whispered Speech Recognition“, ARCHIVES OF ACOUSTICS, Vol. 43, No. 1, str. 3-9, 2018. ISSN 0137-5075, IF=0,917. DOI: 10.24425/118075

Kategorija M14:

5. Đ. Grozdić, S. Jovičić, **J. Galić**, B. Marković: *Experiments i whisper recognition using neural networks*. In: S. Jovičić, M. Subotić, M. Sovilj (eds): VERBAL COMMUNICATION QUALITY, Interdisciplinary Research, II; CUŽA, Belgrade, ISBN 978-86-81879-46-7, 91-110, 2013.

Kategorija M33:

6. Đ. Grozdić, B. Marković, **J. Galić**, S. Jovičić: „Primena neuralnih mreža u prepoznavanju govora u šapatu“, Zbornik radova konferencije TELFOR 2012, str. 728-731, ISBN 978-1-4673-2984-2, Beograd, Srbija, 2012.
7. **J. Galić**, T. Pešić-Brđanin, „The Voice Fundamental Frequency Statistical Parameters under Noisy Conditions with the Cepstrum Method“, 10th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services TELSIS 2011, str. 769-772, ISBN 978-1-4577-2018-5, sveska 2, Niš, Srbija, 2011.
8. B. Marković, S. Jovičić, **J. Galić**, Đ. Grozdić: „Whispered Speech Database: Design, Processing and Application“, Zbornik radova 16th International Conference, TSD 2013, str. 591-598, ISBN 978-3-642-40584-6, Pilsen, Czech Republic, 2013.
9. B. Marković, **J. Galić**, Đ. Grozdić, S. Jovičić: „Application of DTW Method for Whispered Speech Recognition“, Zbornik radova konferencije SPEECH AND LANGUAGE 2013, str. 308-315, ISBN 978-86-81879-45-0, Beograd, Srbija, 2013.
10. **J. Galić**, B. Marković, Đ. Grozdić, S. Jovičić: „The Influence of Feature Vector Selection on Performance of Automatic Recognition of Whispered Speech“, Zbornik radova konferencije SPEECH AND LANGUAGE 2013, str. 258-264, ISBN 978-86-81879-45-0, Beograd, Srbija, 2013.
11. **J. Galić**, S. Jovičić, Đ. Grozdić, B. Marković: „HTK-Based Recognition of Whispered Speech“, Zbornik radova 16th International Conference on Speech and Computer SPECOM 2014, str. 251-258, ISBN 978-3-319-11580-1, Novi Sad, Srbija, 2014.
12. **J. Galić**, S. Jovičić, Đ. Grozdić, B. Marković: „Constrained Lexicon Speaker Dependent Recognition of Whispered Speech“, Zbornik radova konferencije INDEL 2014, str. 180-184, ISBN 978-99955-46-22-9, Banja Luka, Republika Srpska, 2014.
13. Đ. Grozdić, S. Jovičić, **J. Galić**, B. Marković: „Application of inverse filtering in enhancement of whisper recognition“, Zbornik radova konferencije NEUREL 2014, str. 157-161, ISBN 978-1-4799-5887-0, Beograd, Srbija, 2014.
14. **J. Galić**, D. Šumarac Pavlović, S. Jovičić, B. Marković, Đ. Grozdić: „Prepoznavanje bimodalnog govora bazirano na metodi potpornih vektora“, Zbornik radova konferencije TELFOR, ISBN: 978-1-5386-3072-3, Beograd, Srbija, 2017.

15. B. Marković, G. Stevanović, S. Jovičić, M. Mijić, **J. Galić**, Đ. Grozdić: „Recognition of Normal and Whispered Speech Based on RASTA Filtering and DTW Algorithm“, Zbornik radova konferencije IcETAN, str. AK1.8.2-4, ISBN 978-86-7466-692-0, Kladovo, Srbija, 2017.
16. B. Marković, M. Mijić, **J. Galić**: „APPLICATION OF TEAGER ENERGY OPERATOR ON PLP FEATURE FOR WHISPER SPEECH RECOGNITION“, Zbornik radova konferencije IcETAN 2018, ISBN 978-86-7466-752-1, Palić, Srbija, 2018.

Категорија M52:

17. Đ. Grozdić, **J. Galić**, B. Marković, S. Jovičić: *Application of Neural Networks in Whispered Speech Recognition*, Telfor Journal, Vol. 5, No. 2, ISSN 1821-3251, str. 103-106, 2013.
18. **J. Galić**, S. Jovičić, V. Delić, B. Marković, D. Šumarac Pavlović, Đ. Grozdić: “HMM-based Whisper Recognition Using μ -law Frequency Warping”, SPIRAS Proceedings, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, ISSN 2078-9181 (print), ISSN 2078-9599 (online), Issue No. 3(58), 2018. pp. 27-52, DOI: 10.15622/sp.58.2

Категорија M63:

19. **J. Galić**, M. Popović, S.T. Jovičić, B. Marković, Đ. Grozdić: „Primjena skrivenih Markovljevih modela u prepoznavanju govora u šapatu“, Zbornik radova konferencije INFOTEN 2013, str. 387-390, ISBN 978-99955-763-1-8, Jahorina, Republika Srpska, 2013.
20. **J. Galić**, T. Pešić-Brđanin, “Uporedna analiza metoda za estimaciju osnovne frekvencije govornog signala u prisustvu bijelog šuma“, Zbornik radova 55. Konferencije za ETAN, AK2.1-1-4, ISBN 978-86-80509-66-2, Banja Vrućica, Republika Srpska, 2011.
21. Đ. Grozdić, D. Šumarac Pavlović, S. Jovičić, **J. Galić**, B. Marković: „Komparacija tehnika normalizacije kepralnih koeficijenata u automatskom prepoznavanju šapata“, Zbornik radova konferencije ETAN 2015, str. AK1 1.8 1-5, ISBN 978-86-80509-71-6, Srebrno Jezero, Srbija, 2015.
22. **J. Galić**, S. Jovičić, B. Marković, D. Šumarac Pavlović, Đ. Grozdić: „Speaker dependent recognition of whispered speech based on MLLR adaptation“, Zbornik radova konferencije DOGS, str. 29-32, ISBN: 978-86-7892-993-9, Novi Sad, Srbija, 2017.

5. ЗАКЉУЧАК И ПРЕДЛОГ

На основу чињеница изложених у овом реферату Комисија је закључила да докторска дисертација кандидата Јована Галића под насловом “Препознавање мултимодалног говора засновано на статистичком приступу” испуњава све формалне и суштинске услове предвиђене Законом о високом образовању и Правилником о докторским студијама Електротехничког факултета Универзитета у Београду.

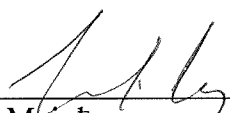
У дисертацији је детаљно анализирана проблематика препознавања мултимодалног говора, и то пре свега нормалног говора и шапата. Иако је тежиште дисертације фокусирано на препознавање засновано на статистичком приступу, развијен је и препознавач заснован на методи потпорних вектора. Перформансе обе технике су испитане и анализирани на реалној говорној бази стимулуса нормалног говора и шапата. Такође, у дисертацији је предложени нови алгоритам за екстракцију робустних вектора обележја. Статистички тестови су

потврдили успешност предложеног алгоритма у препознавању зависно и независно од говорника. Кроз истраживачке активности из којих је проистекла дисертација кандидат је показао способност да успешно примењује методологију научно-истраживачког рада. Резултате проистекле из истраживања спроведеног у оквиру израде дисертације кандидат је публикувао у међународним научним часописима и презентовао на међународним и националним конференцијама.

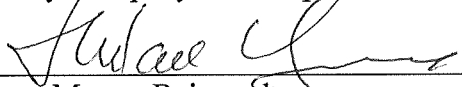
Комисија предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да се докторска дисертација под насловом “Препознавање мултимодалног говора засновано на статистичком приступу” кандидата Јована Галића прихвати, изложи на увид јавности и упуту на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

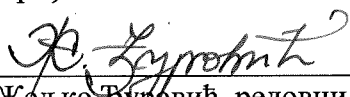
У Београду, 1.03.2019.

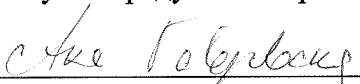
ЧЛАНОВИ КОМИСИЈЕ


др Миомир Мијић, редовни професор
Универзитет у Београду – Електротехнички факултет


др Јелена Терзић, доцент
Универзитет у Београду – Електротехнички факултет


др Милан Војновић, научни сарадник
Центар за унапређење животних активности, Београд


др Жељко Туровић, редовни професор
Универзитет у Београду – Електротехнички факултет


др Ана Гавровска, доцент
Универзитет у Београду – Електротехнички факултет