

## НАСТАВНО-НАУЧНОМ ВЕЋУ

**Предмет:** Реферат о урађеној докторској дисертацији кандидата Ђорђа Гроздића.

Одлуком бр. 5013/11-3 од 25.4.2017. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Ђорђа Гроздића под насловом

### “Примена неуралних мрежа у препознавању шапата“

После прегледа достављене дисертације и других пратећих материјала и разговора са кандидатом, Комисија је сачинила следећи

## РЕФЕРАТ

### 1. УВОД

#### 1.1. Хронологија одобравања и израде дисертације

Ђорђе Гроздић је уписао докторске студије у пролећном семестру школске 2011/12. године и пријавио је тему за израду докторске дисертације 12.6.2014. године. Комисија за студије трећег степена је на својој седници дана 18.6.2014. године разматрала поднету пријаву теме докторске дисертације и свој предлог о оцени подобности теме и кандидата упутила Наставно-научном већу на усвајање. Наставно-научно веће је на својој седници дана 14.7.2014. године именовало Комисију за оцену услова и прихватање теме ове докторске дисертације. Наставно-научно веће је на седници одржаној дана 28.10.2014. године усвојило поднети извештај Комисије. Веће научних области техничких наука својом одлуком број 61206-5281/2-14 од 24.11.2014. године је дало сагласност на предлог теме докторске дисертације под насловом: “Примена неуралних мрежа у препознавању шапата“.

Кандидат је дана 30.3.2017. године предао урађену докторску дисертацију на преглед и оцену. Комисија за студије трећег степена на својој седници одржаној дана 4.4.2017. године потврдила је испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације. На основу тога Наставно-научно веће Факултета је својом одлуком број 5013/11-3 од 25.4.2015. године именовало Комисију за преглед и оцену докторске дисертације у сасатаву који је наведен на крају овог извештаја.

#### 1.2. Научна област дисертације

Ова докторска дисертација припада области Техничких наука Електротехника, ужа научна област Електроакустика. За ментора је одређена др Драгана Шумарац Павловић, ванредни професор. Она је изабрана у звање ванредног професора за област Електроакустика и сви њени публиковани радови који је квалификују за ментора су из те области.

### 1.3. Биографски подаци о кандидату

Ђорђе (Томислав) Гроздић је рођен 20.01.1987. године у Београду, где је завршио основну школу и Трећу београдску гимназију са Вуковом дипломом. Дипломирао је 2010. године на Електротехничком факултету у Београду на Одсеку за телекомуникације и информационе технологије. Тема дипломског рада је била „Мултидимензионална анализа акустичких обележја у говорном сигналу“, под менторством проф. др Слободана Јовичића. Мастер студије на Електротехничком факултету у Београду (Смер за аудио и видео технологије) је завршио 2011. године. Тема мастер рада је била „Анализа варијација акустичких обележја у говорном сигналу“, такође под менторством проф. Слободана Јовичића.

Радну каријеру је започео 2011. године у фирми Телефонкабл а.д. у сектору за пројектовање и развој телекомуникационих система. Децембра 2011. године, наставља студије на Електротехничком факултету и уписује докторске студије, модул Телекомуникације, са руководиоцем научно-истраживачког рада проф. Слободаном Јовичићем. Почетком 2012. године стиче научно звање истраживач-сарадник и започиње радни однос у Центру за унапређење животних активности у Лабораторији за форензичку акустику и фонетику, где је ангажован на технолошком пројекту Министарства просвете, науке и технолошког развоја Републике Србије, под бројем ТР32032 и називом „е-логопед“. Одласком проф. Јовичића у пензију 2014. године, менторство у изради докторске тезе преузима проф. др Драгана Шумарац Павловић. Подручје истраживачког рада Ђорђа Гроздића обухвата мултидисциплинарно истраживање говора и говорне комуникације и то пре свега области: дигиталне обраде говорних сигнала, комуникације човек-рачунар, вештачке интелигенције, аутоматског препознавања говора и нетипичних облика говора (попут емотивног говора, шапата...), као и аутоматског препознавања говорника. Од новембра 2016. године, Ђорђе Гроздић је запослен на позицији софтверског инжењера у компанији Fincore Ltd.

Као резултат досадашњег научно-истраживачког рада, Ђорђе Гроздић има преко 40 објављених стручних и научних радова, од тога: 3 рада у међународним часописима са SCI листе, 1 рад у часопису националног значаја, 6 поглавља у монографијама међународног значаја, 19 радова на међународним конференцијама, 9 радова на националним конференцијама и 5 техничких решења. Ђорђе Гроздић је члан Удружења студената електротехнике Европе. Добитник је награде за најбољи рад младог аутора на међународној конференцији ТЕЛФОР 2012. године, а 2016. Године је награђен за најбољи рад младог истраживача на конференцији ЕТРАН.

## **2. ОПИС ДИСЕРТАЦИЈЕ**

### 2.1. Садржај дисертације

Докторска дисертација Ђорђа Гроздића је написана на 174 стране. На почетку се налазе насловна страна и кратак резиме на српском и енглеском језику, као и садржај. Сам текст рада садржи једанаест поглавља и преглед коришћене литературе. Поглавља су организована у следећем редоследу: 1 Увод, 2 Увод у аутоматско препознавање говора, 3 Увод у вештачке неуралне мреже, 4 Шапат, 5 Креирање и анализа корпуса шапата, 6 Креирање MLP система за препознавање шапата, 7 Креирање тандем DNN-HMM система за препознавање шапата, 8 Експерименти са MLP системом, 9 Експерименти са тандем DNN-HMM системом, 10 Компарација резултата препознавања шапата, и на крају 11 Закључак. На самом крају дисертације, као додаток, налазе се обавезни прилози: биографија аутора и неопходне изјаве аутора.

### 2.2. Кратак приказ појединачних поглавља

Материја изложена у докторској дисертацији се може логички поделити на две целине. У првој целини је изложен теоријски увод у аутоматско препознавање говора, затим опис неких од најпопуларнијих система аутоматског препознавања говора (тзв. ASR система) укључујући и поглавље о неуралним мрежама, као и посебно поглавље са теоријским освртом на акустичке карактеристике шапата које га чине посебним и проблематичним у препознавању помоћу актуелних

ASR система. Друга целина дисертације садржи приказ самосталног рада кандидата и састоји се из седам поглавља, укључујући и закључак као резиме постигнутог.

У првом поглављу, кроз уводну причу о тренутним перформансама, могућностима и манама актуелних ASR система у препознавању говора, објашњена је потреба за њиховим даљим усавршавањем пре свега у погледу њихове адаптације на различите говорне модалитете међу које спада и шапат. Уз јасно дефинисане мотиве и циљеве докторске дисертације, у овом поглављу је дат и осврт на досадашња истраживања и постигнуте резултате у аутоматском препознавању шапата.

Друго поглавље даје кратак увод у основе аутоматског препознавања говора. Најпре су дефинисани: основни концепт аутоматског препознавања говора, неопходни кораци у креирању једног ASR система, као и статистичка формулација проблема доношења одлуке. Описане су две широко примењене технике у ASR системима: тренутно актуелни скривени Марковљеви модели (HMM) и нешто старија али још увек присутна техника од великог теоријског значаја - динамичко временско усклађивање (DTW). У поглављу су објашњене предности и мане сваке од ових техника, чиме је направљена основа за даље поређење са неуралним мрежама.

Треће поглавље пружа најбитније информације о вештачким неуралним мрежама, почевши од њиховог историјског развоја, различитих типова и архитектура мрежа, преко њиховог начина процесирања, обуке, па све до везе са статистичким моделима. Посебна пажња је посвећена *feedforward* типу неуралних мрежа, тачније вишеслојним перцептронима (MLP), *Backpropagation* методи обуке и дубинским аутоенкодерима који су коришћени у експериментима ове дисертације. Неуралне мреже су коментарисане и са аспекта њихових перформанси, могућности и ограничења у препознавању речи, са посебним освртом на додирне тачке и предности у односу на остале технике аутоматског препознавања говора.

Четврто поглавље описује основне карактеристике и природу шапата. Наведене су и објашњене разлике између шапата и нормалног говора, као и начин на који те разлике утичу на аутоматско препознавање шапата. Са физиолошког аспекта су детаљно описани начин генерисања шапата, специфичан рад артикулационих органа и њихов утицај на акустичка својства шапата, пре свега на: енергију, спектрални нагиб и положај форманата. Дискутоване су могућности преноса информација шапатом и ограничења у перцепцији шапата. На крају поглавља је дат преглед резултата и могућности до сада тестираних ASR система у препознавању шапата.

Пето поглавље објашњава поступак креирања првог и тренутно јединог корпуса шапата за српски језик, *Whi-Spe*, на коме су вршени сви експерименти ове дисертације. Најпре је дат преглед постојећих и до сада у литератури познатих корпуса шапата (за јапански, енглески и кинески језик), а затим дизајн, снимање и обрада *Whi-Spe* корпуса. Описане су специфичне манифестације шапата током снимања, као и методе контроле квалитета снимака. Поглавље садржи и детаљно описује поступак евалуације и акустичке анализе *Whi-Spe* корпуса, пре свега у погледу: анализе таласних облика сигнала и њихових спектрограма, спектралног нагиба и кепстралних карактеристика. У наставку поглавља је објашњена улога инверзног филтрирања у креирању базе псеудо-шапата, као и кепстрална анализа постигнутих резултата.

Шесто поглавље систематски приказује поступак креирања MLP система за аутоматско препознавање изолованих речи у нормалном говору и шапату. У првом делу поглавља описан је такозвани *front-end* део ASR система, задужен за предобраду говорних сигнала (сегментација, временско усклађивање, филтрирање и прозоровање говорних сигнала). Затим се описује детаљан математички поступак екстракције MFCC, TECC и TEMFCC говорних обележја, уз посебан акценат и тумачење карактеристика и поступка израчунавања *Teager* енергије. Следи објашњење формирања делова базе за обуку, валидацију и тестирање неуралних мрежа. На самом крају поглавља описано је креирање вештачке неуралне мреже (MLP), одређивање њене оптималне структуре, као и поступак њеног обучавања.

Седмо поглавље описује поступак креирања тандем DNN-HMM система за аутоматско препознавање изолованих речи у нормалном говору и шапату. На почетку поглавља је презентован *front-end* део система чију основу чини дубинска неурална мрежа (DNN) – тачније дубински *denoising* аутоенкодер

(DDAE). У овом делу су поред предобrade сигнале и екстракције говорних обележја детаљно описани: структура, креирање и обука дубинског *denoising* аутоенкодера, који у *front-end* делу тандем DNN-HMM система има улогу секундарног екстрактора робустних говорних обележја. У наставку текста је описан *back-end* део система кога чини HMM препознавач, као и поступак обуке тако формираног тандем система.

Осмо поглавље даје преглед експерименталних резултата аутоматског препознавања изолованих речи у нормалном говору и шапату остварених помоћу MLP система. Презентоване су анализе различитих обука/тест сценарија, поређење перформанси вишеслојних перцептрона и њихових успеха у класификацији речи у зависности од коришћења три типа кепстралних коефицијената. Приказана је анализа конфузије у препознавању речи заједно са спектралним тумачењем конфузија посебно истакнутих критичних парова речи. На основу добијених резултата, у овом поглављу је постављена хипотеза о маскирању одређених говорних обележја услед звучности, која представља главни узрок деградираног препознавања речи у неусаглашеним обука/тест сценаријима. Ради доказа хипотезе и побољшања успеха препознавања речи, пре свега у шапату, предложена је измена *front-end* система где је у фази предобrade сигнала имплементирано инверзно филтрирање. Тако измењен систем је назван MLP-IF, а остварени експериментални резултати у побољшању успеха препознавања речи су описани на крају поглавља.

Девето поглавље представља експерименте са тандем DNN-HMM системом у аутоматском препознавању изолованих речи. На почетку поглавља су табеларно приказани усредњени резултати мушких и женских говорника у усаглашеним обука/тест сценаријима при коришћењу сва три типа кепстралних коефицијената, а затим су у наставку поглавља презентовани и постигнути резултати у неусаглашеним обука/тест сценаријима.

Десето поглавље врши компарацију максималних остварених резултата препознавања речи у усаглашеним и неусаглашеним обука/тест сценаријима помоћу MLP, MLP-IF и DNN-HMM система. Такође, поглавље даје и додатно поређење перформанси ових система са још неким од постојећих ASR системима из литературе (DTW и НТК-HMM) који су тестирани на истој бази података (Whi-Spe).

Једанаесто поглавље сумира постигнуте резултате, дискутује их, објашњава научни допринос ове дисертације и предлаже могуће правце даљег истраживања.

### 3. ОЦЕНА ДИСЕРТАЦИЈЕ

#### 3.1. Савременост и оригиналност

Рад на овој докторској дисертацији инициран је најновијим кретањима науке у области аутоматског препознавања атипичних говорних модалитета. Наиме, актуелни ASR системи су достигли такозвани плафон у погледу перформанси у препознавању неутралног говора. Међутим, у пракси ови системи су се показали далеко од идеалних, пре свега због различитих присутних сметњи у виду шума, буке и других интерференција које дуж пропагације маскирају говорни сигнал и значајно деградирају његово препознавање. Посебан проблем представљају атипични облици говора, попут тихог говора, певаног говора, афективног говора, шапата итд. Шапат сигурно представља један од најкомпликованијих говорних модалитета у погледу могућности исправног аутоматског препознавања помоћу актуелних ASR система. Имајући у виду поменути проблематику и значај усавршавања ASR система, истраживање са циљем реализовања успешног препознавања шапата представља савремени и актуелни истраживачки проблем. У овој дисертацији и у предложеном решењу система за аутоматско препознавање шапата примењено је више савремених и оригиналних приступа. Пре свега, чињеница да до тренутка писања докторске дисертације није постојао ни један систем базиран на неуралним мрежама који је дизајниран за потребе препознавања шапата, довољно иде у прилог и потврђује оригиналност ове дисертације. Метода обраде говорних сигнала, попут инверзног филтрирања говорног сигнала као саставног дела процеса креирања псеудо-шапата је такође оригинални приступ који се по први пут користи у обради говорних сигнала (шапата) и помиње управо у публикованим радовима кандидата. Примена савремених дубинских неуралних

мрежа, у виду *denoising* аутоенкодера је такође по први пут анализирана у сврхе препознавања шапата. Такође, у дисертацији се испитују и релативно нова кепстрална обележја која су доста робустнија од традиционалних MFCC обележја а до сада нису тестирана у препознавању шапата. Према томе, резултати који су проистекли из дисертације су дали један нови, оригинални систем за аутоматско препознавање шапата и неутралног говора.

### 3.2. Осврт на референтну и коришћену литературу

Током рада на дисертацији кандидат је користио врло обимну литературу из области која је обрађена. Списак референци дат на крају дисертације садржи 165 наслова. У оквиру тог броја налази се врло широк опсег публикација који укључује књиге, часописе, зборнике са домаћих и међународних конференција. Велики део публикација је новијег датума, мада су укључени и сви најзначајнији историјски наслови да би се употпунила слика о обрађиваној материји. У списку референци налазе се и радови у којима је кандидат аутор или коаутор, а који садрже оригиналне резултате директно произашле из рада на дисертацији.

### 3.3. Опис и адекватност примењених научних метода

У оквиру анализе проблема аутоматског препознавања шапата кандидат је на основу систематизације истраживања која су присутна у литератури, а баве се системима препознавања говора и математичких алгоритама који се тичу дигиталне обраде говорних сигнала, креирао систем за аутоматско препознавање изолованих речи изговорених у неутралном говору и шапату. Током реализације научних резултата, неколико научних метода је било примењено. Прво, кандидат је формирао говорну базу од 10.000 стимулуса која садржи корпус речи неутралног изговора и изговора у шапату (снимано је више говорника, снимци су сегментирани, анализирани и оптимизовани у погледу квалитета снимака). Затим, кандидат је софтверски реализовао предобраду говорних стимулуса, применивши више математичких алгоритама за екстракцију говорних обележја у параметарском домену. Следећи корак је био креирање више разлочитих софтверских модела система за аутоматско препознавање говора, са посебним акцентом у анализи система базираних на неуралним мрежама. Са формираним ASR системима, приступило се софтверској реализацији њихове обуке и тестирања. Кроз читав низ експеримената у којима су се мериле и поредиле перформансе различитих система у различитим сценаријима (првенствено у погледу анализе неусаглашених обука/тест сценарија) одређивана је архитектура ASR система и вршене су дораде како у *front-end* тако и у *back-end* делу система. На самом крају, веома је битна примењена статистичка анализа добијених резултата, пре свега у погледу оцене тачности препознавања речи, као и у статистичкој анализи препознавања и конфузије стимулуса (анализа конфузија матрица). Кандидат је за своја софтверска решења користио програмски језик MATLAB а за статистичку анализу SPSS програм.

### 3.4. Применљивост остварених резултата

Резултати који су приказани у овој докторској дисертацији имају директну примену у области аутоматског препознавања говора. Предложени систем за аутоматско препознавање изолованих речи омогућава знатно побољшање препознавања шапата уз задржање високих перформанси у препознавању неутралног говора. Самим тим, предложено решење може послужити као основа за развој напредних ASR система који подржавају препознавање бимодалног (комбинованог говора и шапата) или чак мулти-модалног говора. Демонстриране карактеристике *Teager* обележја, инверзног филтрирања, и *denoising* аутоенкодера указују на могућност њихове примене у аутоматском препознавању не само шапата већ и других говорних модалитета, различитих облика деградираниог говора и на самом крају у препознавању патолошког говора. Посебно је интересантна могућност примене *denoising* аутоенкодера у сврху препознавања и реконструкције патолошких облика говора. Иако су експерименти ове дисертације вршени на задацима аутоматског препознавања изолованих речи, предложена решења су демонстрирала релативно брзо и једноставно процесирање у реалном времену што указује на могућност њихове имплементације у системима за препознавање континуалног говора.

### 3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат Ђорђе Гроздић је кроз одабир и систематичан преглед актуелне литературе, теоријску анализу, формирање, обраду и анализу говорне базе, имплементацију постојећих и развој оригиналних алгоритама за обраду и препознавање говорних сигнала, али и кроз друге сегменте свог рада показао висок степен самосталности у истраживачком раду. Томе треба додати и чињеницу да је највећи део радова на међународним и националним скуповима самостално презентовао.

## **4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС**

### 4.1. Приказ остварених научних доприноса

Данашњи системи за аутоматско препознавање говора су осетљиви и непоздани у препознавању било ког нетипичног облика говора. Из тог разлога аутоматско препознавање шапата, самим тим и ова докторска дисертација, представља битну и актуелну истраживачку тему са циљем постизања боље комуникације човек-рачунар. У складу са тим, постигнути су следећи научни доприноси докторске дисертације:

- Дисертација је омогућила суштинско разумевање својстава шапата у бимодалној говорној комуникацији, појаснила је и на одговарајући начин квантификовала разлику акустичких обележја која карактеришу шапат и говор у бимодалној (говор-шапат) комуникацији;
- Теоријски су разјашњењи и експериментално испитани слабији резултати препознавања у неусаглашеним обука/тест (говор/шапат и шапат/говор) сценаријима;
- Предложена су нова, робустнија обележја од традиционалних MFCC, која захваљујући карактеристикама *Teager* енергије и *Gammatone* банке филтара много боље акустички моделују шапат и на тај начин побољшавају препознавање у неусаглашеним обука/тест сценаријима;
- Дисертација је по први пут демонстрирала примену неуралних мрежа у аутоматском препознавању шапата. Тестирана су два система базирана на неуралним мрежама –вишеслојни перцептрони (MLP) и тандем DNN-HMM систем. Перформансе ових система у препознавању шапата су упоређене са још неким постојећим ASR системима (DTW и GMM-HMM).
- Предложена је нова метода компензације нежељених разлика акустичких обележја у неусаглашеним обука/тест сценаријима, заснована на инверзном филтрирању, која додатно побољшава препознавање бимодалног говора (говор-шапат). Инверзно филтрирање је уједно послужило и као нови метод креирања псеудо-шапата, чиме се пружају могућности брзог и једноставног креирања великог броја узорака вештачког шапата чиме се решава проблем непостојања одговарајуће базе реалних снимака шапата неопходних за адаптацију ASR система.
- Предложен је нови ефикасни систем за препознавање шапата, базиран на тандем DNN-HMM систему који поседује низ предности у поређењу са другим ASR системима попут: (1) значајног побољшања тачности препознавања речи у неусаглашеним обука/тест сценаријима, (2) изузетно високих перформанси у усаглашеним обука/тест сценаријима, (3) могућности ефикасне реконструкције кепстралних карактеристика нормалног говора из узорака шапата у реалном времену, (4) брзе и једноставне обуке система помоћу узорака псеудо-шапата чиме је отклоњена потреба за реалним узорцима шапата.

### 4.2. Критичка анализа резултата истраживања

На основу систематског истраживања акустичких карактеристика шапата и неусаглашених обука/тест сценарија у процесу аутоматског препознавања, кандидат је демонстрирао развој решења за аутоматско препознавање изолованих речи у шапату применом неуралних мрежа. У квалитативном смислу, већи део дисертације се бави проблемом предобраде говорних сигнала и проналаском начина за смањењем акустичких разлика између неутралног говора и шапата, зарад побољшања успеха препознавања речи у неусаглашеним обука/тест сценаријима. Кандидат је решавању овог проблема

приступио из два правца. Први правац представља проналазак и одабир одговарајућих кепстралних обележја која су довољно робуствна и погодна за аутоамтско препознавање шапата. Други правац представља два начина предобраде говорних сигнала – један помоћу инверзног филтрирања који из стимулуса говора потискује звучност и по акустичким карактеристикама приближава неутралан говор шапату, а други, који ради обратно, помоћу дубинског аутоенкодера реконструише карактеристике неутралног говора из стимулуса шапата. Оба правца су допринела побољшању препознавања шапата, при чему су најбољи резултати постигнути са тандем DNN-НММ системом. Предложена решења и детаљни резултати истраживања су верификовани објављивањем 22 рада у часописима, на међународним и националним скуповима.

#### 4.3.Верификација научних доприноса

Научни допринос докторске дисертације Ђорђа Гроздића верификовани су у следећим радовима (наведени по М категоријама према Правилнику Министарства просвете и науке Србије):

##### Категорија М21а:

- 1 **Ђорђе Т. Гроздић**, Слободан Т. Јовичић, Мишко Суботић, “*Whispered Speech recognition Using Deep Denoising Autoencoder*”, Engineering Applications of Artificial Intelligence, Vol. 59, No. 1, 2016, ISSN 0952-1976, p.15-22, IF=2.368 .  
doi.org/10.1016/j.engappai.2016.12.012

##### Категорија М22:

- 2 **Ђорђе Т. Гроздић**, Слободан Т. Јовичић, Драгана Шумарац Павловић, “*Comparison of Cepstral Normalization Techniques in Whispered Speech Recognition*”, Advances in Electrical and Computer Engineering, Vol. 17, No. 1, 2017, ISSN 1582-7445, p. 21-26, IF=0.459.  
doi: 10.4316/AECE.2017.01004
- 3 Слободан Т. Јовичић, Никола Јовановић, Мишко Суботић, **Ђорђе Т. Гроздић**, “*Impact of Mobile Phone Usage on Speech Spectral Features*”, International Journal of Speech Language and the Law, Vol. 11, No. 1, 2015, ISSN 1748-8885, p. 83-102, IF=0.217.  
doi: 10.1558/ijsl.v22i1.17880

##### Категорија М14:

- 4 **Ђорђе Т. Гроздић**, Слободан Јовичић, Јован Галић, Бранко Марковић, “*Experiments in whisper recognition using neural networks*”. Verbal Communication Quality - Interdisciplinary Research II; Едитори: Слободан Т. Јовичић, Мишко Суботић, Миријана Совиљ, 2013, ISBN 978-86-81879-46-7, LAAC IEPSP, Београд, p. 91-110.
- 5 **Ђорђе Т. Гроздић**, Слободан Јовичић, “*Application of algorithms based on neural networks in whispered speech recognition*”. Specific Applications of Information Technology and Signal Processing in Speech Disorder Diagnosis and Therapy, Едитори: Слободан Т. Јовичић, Зоран Шарић, Мишко Суботић, 2016, ISBN 978-86-89431-14- 8, LAAC& IEPSP, Београд, p. 125-176.

##### Категорија М33:

- 6 **Ђорђе Т. Гроздић**, Бранко Марковић, Јован Галић, Слободан Јовичић, “*Примена неуралних мрежа у препознавању говора у шапату*”. 20. TELFOR, новембар 2012, Зборник радова, Београд, ISBN 978-1-4673-2982-8, p. 728-731.
- 7 Бранко Марковић, Слободан Јовичић, Јован Галић, **Ђорђе Т. Гроздић**, “*Whispered Speech Database: Design, Processing and Application*”. 16th International Conference on Text, Speech and Dialogue TSD2013, 2013, Зборник радова, Чешка република, Плзен, ISBN 978-3-642-40584-6, p. 591-598.
- 8 **Ђорђе Т. Гроздић**, Бранко Марковић, Јован Галић, Слободан Јовичић, Драшко Фурунџић, “*Neural Network-Based Recognition of Whispered Speech*”. 4th International Conference on Fundamental and Applied Aspects of Speech and Language, 2013, Зборник радова, Београд, ISBN 978-86-81879-44-3, p. 223-229.
- 9 Јован Галић, Бранко Марковић, **Ђорђе Т. Гроздић**, Слободан Јовичић, “*The Influence of Feature Vector Selection on Performance of Automatic Recognition of Whispered Speech*”. 4th

- International Conference on Fundamental and Applied Aspects of Speech and Language, 2013, Зборник радова, Београд, ISBN 978-86-81879-44-3, p. 258-264.
- 10 Бранко Марковић, Јован Галић, **Ђорђе Т. Гроздић**, Слободан Јовичић, “*Application of DTW method for Whispered Speech Recognition*”. 4th International Conference on Fundamental and Applied Aspects of Speech and Language, 2013, Зборник радова, Београд, ISBN 978-86-81879-44-3, p. 308-315.
- 11 Бранко Марковић, **Ђорђе Т. Гроздић**, “*The LPCC-DTW Analysis for Whispered Speech Recognition*”, 1. International Conference on Electrical, Electronic and Computing Engineering IcETRAN2014, 2014, Србија, Врњачка Бања, ISBN: 978-86-80509-70-9, АК1.1.1-4.
- 12 Јован Галић, Слободан Јовичић, **Ђорђе Т. Гроздић**, Бранко Марковић, “*HTK-based recognition of whispered speech*”. 16th International Conference Speech and Computer SPECOM, 2014, Зборник радова, Србија, Нови Сад, ISBN 978-3-319-11581-8, p. 251–258.
- 13 **Ђорђе Т. Гроздић**, Слободан Јовичић, Јован Галић, Бранко Марковић, “*Application of inverse filtering in enhancement of whisper recognition*”. Syposium on Neural Network Applications in Electrical Engineering NEUREL, 2014, Зборник радова, Србија, Београд, ISBN 978-1-4799-5887-0, p. 157-161.
- 14 Јован Галић, Слободан Јовичић, **Ђорђе Т. Гроздић**, Бранко Марковић, “*Constrained Lexicon Speaker Dependent Recognition of Whispered Speech*”. 10th International Symposium on Industrial Electronics INDEL, 2014, Зборник радова, Босна и Херцеговина, Бања Лука, ISBN 978-99955-46-22-9, p. 180-184.
- 15 Бранко Марковић, Слободан Јовичић, Јован Галић, **Ђорђе Т. Гроздић**, “*Recognition of the Multimodal Speech Based on the GFCC Features*”. 2. International Conference on Electrical, Electronic and Computing Engineering IcETRAN, 2015, Зборник радова, Србија, Сребрно језеро, ISBN 978-86-80509-71-6; АК11.3.1-5.
- 16 **Ђорђе Т. Гроздић**, Слободан Јовичић, Зоран Шарић, Ирина Суботић, “*Comparison of GMM/UBM and i-vector based speaker recognition systems*”, 5th International Conference on Fundamental and Applied Aspects of Speech and Language, 2015, Зборник радова, Србија, Београд, ISBN: 978-86-89431-07-0, p. 274-281.
- 17 Јован Галић, Слободан Јовичић, Драгана Шумарац Павловић, Бранко Марковић, **Ђорђе Т. Гроздић**, “*HMM-Based Recognition of Normally Phonated and Whispered Speech*”. 3rd International Acoustics and Audio Engineering Conference, 2016, Зборник радова, Србија, Нови Сад, ISBN: 978-86-7892-758-4, p. 62-63.
- 18 Бранко Марковић, Слободан Јовичић, Миомир Мијић, Јован Галић, **Ђорђе Т. Гроздић**, “*Recognition of Whispered Speech Based on PLP Features and DTW Algorithm*”. International Conference on Electrical, Electronic and Computing Engineering IcETRAN, 2016, Зборник радова, Србија, Златибор, ISBN 978-86-7466-618-0, p. АК1.2.1-4.

#### Категорија М52:

- 19 **Ђорђе Т. Гроздић**, Јован Галић, Бранко Марковић, Слободан Јовичић. “*Application of Neural Networks in Whispered Speech Recognition*”. Telfor Journal, Vol. 5, No. (2), 2013, ISSN 1821-3251, p. 103-106, JBR IF 5=0.400.

#### Категорија М63:

- 20 **Ђорђе Т. Гроздић**, Слободан Јовичић, “*Примена ANOVA и PCA статистичких метода у анализи акустичких обележја говора*”. Конференција Дигитална обрада говора и слике DOGS, 2010, Зборник радова, Србија, Ковачица, ISBN 978-86-7892-439-2: p. 182-185.
- 21 Јован Галић, М. Поповић, Бранко Марковић, **Ђорђе Т. Гроздић**, Слободан Јовичић, “*Примјена скривених Марковљевих модела у препознавању говора у штату*”. 12. научно-стручни симпозијум INFOTEN, 2013, Зборник радова, Босна и Херцеговина, Јахорина, ISBN 978-99955-763-1-8, p. 387 - 390.
- 22 **Ђорђе Т. Гроздић**, Слободан Јовичић, Драгана Шумарац Павловић, Јован Галић, Бранко Марковић, “*Компарација техника нормализације кепстралних коефицијената у аутоматском препознавању шпата*”, 59. Конференција за електронику, телекомуникације, рачунарство, аутоматику и нуклеарну технику ЕТРАН, 2015, Зборник радова, Србија, Сребрно језеро, ISBN 978-86-80509-72-3, АК 1.8. 1-5.



## 5. ЗАКЉУЧАК И ПРЕДЛОГ

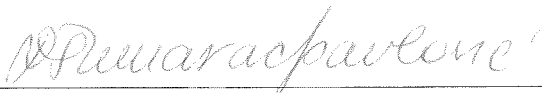
На основу чињеница изложених у овом извештају Комисија је закључила да докторска дисертација Ђорђа Гроздића под насловом “Примена неуралних мрежа у препознавању шапата” испуњава све форме и суштинске услове предвиђене Законом о високом образовању и прописима Универзитета у Београду и Електротехничког факултета.

У дисертацији су истраживачки обрађене теме аутоматског препознавања како говора тако и шапата, акустичке карактеристике шапата и његове разлике у односу на неутралан говор, примене неуралних мрежа у аутоматском препознавању говорних стимулуса, екстракције робустних говорних обележја погодних за препознавање шапата, употребе различитих алгоритама обраде говорних сигнала и адекватних математичких алгоритама са циљем креирања јединственог система за аутоматско препознавање шапата. Кандидат је у оквиру дисертације развио два система базирана на неуралним мрежама за аутоматско препознавање изолованих речи у неутралном говору и шапату. Тиме је предложен нови иновативни систем за аутоматско препознавање шапата који је потврђен детаљним тестирањем са специјалном базом стимулуса нормалног говора и шапата. Кроз истраживачке поступке демонстриране у дисертацији кандидат је показао способност да успешно примењује методологију научно-истраживачко града.

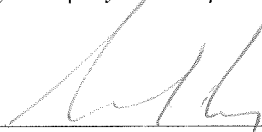
Комисија предлаже Научно-наставном већу Електротехничког факултета да се докторска дисертација под називом “Примена неуралних мрежа у препознавању шапата” кандидата Ђорђа Гроздића прихвати, изложи на увид јавности и упути на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

У Београду, 29.5.2017. године

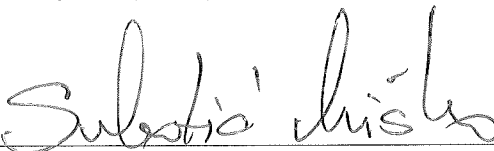
### ЧЛАНОВИ КОМИСИЈЕ



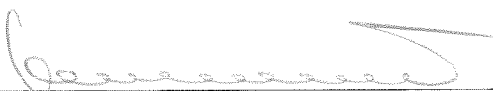
др Драгана Шумарац Павловић, ванредни професор  
Универзитет у Београду – Електротехнички факултет



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



др Мишко Суботић, научни сарадник  
Центар за унапређење животних активности



др Горан Квашчев, доцент  
Универзитет у Београду – Електротехнички факултет