

**НАСТАВНО-НАУЧНОМ ВЕЋУ**

**Предмет:** Подобност теме и кандидата Ђорђа Гроздића, мастер инж. електротехнике и рачунарства, за израду докторске дисертације

Одлуком Наставно-научног већа Електротехничког факултета бр. 5013/11-1 од 14.7.2014. године, именовани смо за чланове Комисије за оцену подобности теме и кандидата Ђорђа Гроздића за израду докторске дисертације и научне заснованости теме „Примена неуралних мрежа у препознавању шапата“.

На основу материјала приложеног уз захтев кандидата Комисија подноси следећи

**ИЗВЕШТАЈ**

**1. Подаци о кандидату**

1.1. Биографски подаци

Ђорђе (Томислав) Гроздић је рођен 1987. године у Београду, где је завршио основну школу и 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“. Мастер студије на Електротехничком факултету у Београду (Смер за аудио и видео технологије) завршио је 2011. године са просечном оценом 10. Тема мастер рада је била „Анализа варијација акустичких обележја у говорном сигналу“.

Ђорђе Гроздић је децембра 2011. године уписао докторске студије на Електротехничком факултету у Београду, модул Телекомуникације, са руководиоцем научно-истраживачког рада проф. др Слободаном Јовичићем. Тренутно је на трећој години докторских студија са положеним свим испитима и просечном оценом 10.

Током основних студија у склопу студентске праксе 2010. године је радио као практикант у Институту за експерименталну фонетику и патологију говора. Радну каријеру је започео 2011. године са стручном праксом у предузећу Београдски водовод, где је шест месеци радио у ИТ сектору као системски инжењер. У периоду од 2011. до 2012. године је радио у фирми Телефонкабл а.д. у сектору за пројектовање и развој. Почетком 2012. године стиче 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 ТП32032 и називом „Е-логопед“. Похађао је и успешно завршио обуку из области Форензике говора „Basic Training course for the subject FAW, including SIS II, Sound Cleaner, EdiTracker, STC-H246“.

Још током основних студија је своје научно интересовање фокусирао на мултидисциплинарно истраживање говора и говорне комуникације и то пре свега у областима дигиталне обраде говорних сигнала, комуникације човек-рачунар, вештачке интелигенције, аутоматског препознавања говора и нетипичних облика говора (емотивни говор, шапат...). До сада има 23 објављених научних и стручних радова.

Ђорђе Гроздић је члан Удружења студената електротехнике Европе. Добитник је награде „Блажо Мирчевски” за најбољи рад младог аутора на међународној конференцији ТЕЛФОР 2012.

## 1.2. Стечено научноистраживачко искуство

Од децембра 2011. године, Ђорђе Гроздић је студент докторских студија на модулу Телекомуникације, Електротехничког факултета у Београду. На докторским студијама Ђорђе Гроздић је одрадио све обавезе и положио испите из следећих предмета:

|                                                         |                  |
|---------------------------------------------------------|------------------|
| Примењена статистика (9 ЕСПБ)                           | оцена 10 (десет) |
| Акустика просторија (9 ЕСПБ)                            | оцена 10 (десет) |
| Говорне технологије (9 ЕСПБ)                            | оцена 10 (десет) |
| Акустичке мерне технике (9 ЕСПБ)                        | оцена 10 (десет) |
| Технике обраде и препознавања говорног сигнала (9 ЕСПБ) | оцена 10 (десет) |
| Робусна обрада говорних сигнала (9 ЕСПБ)                | оцена 10 (десет) |
| Одабрана поглавља из акустике просторије (9 ЕСПБ)       | оцена 10 (десет) |
| Неуралне мреже (9 ЕСПБ)                                 | оцена 10 (десет) |
| Комуникологија (6 ЕСПБ)                                 | оцена 10 (десет) |
| Акустички дизајн просторија (9 ЕСПБ)                    | оцена 10 (десет) |

Предложеној тему за докторску дисертацију кандидат је успешно одбранио на јавној одбрани која је одржана 3.09.2014. године пред комисијом у саставу: проф др Миомир Мијић, проф др Слободан Јовичић и проф др Жељко Ђуровић.

У својству истраживача сарадника, Ђорђе Гроздић у Лабораторији за форензичку акустику и фонетику у Центру за унапређење животних активности, као и у Лабораторији за акустику на Електротехничком факултету у Београду, учествује у разним активностима:

### **- учешће на пројектима:**

1. „Е-логопед”, пројекат ТП32032 Министарства просвете, науке и технолошког развоја Републике Србије, 2011.

### **- техничка решења:**

1. Šarić Z., Subotić M., Jovičić S., Sovilj M., Čabarkapa N., Punišić S., Maksimović S., Bilibajkić R., Stokić M., Đorđević M., Savić M., Grozdić Đ., Buzganović I.: „KSAFA-D digitalni selektivni amplifikator”, Tehničko rešenje.
2. Bilibajkić R., Savić M., Šarić Z., Jovičić S., Sovilj M., Vojnović M., Čabarkapa N., Subotić M., Punišić S., Maksimović S., Stokić M., Grozdić Đ., Nenadović V., Pavković I.: „Logopedski elektronski karton”, Tehničko rešenje.

### **- учешће на курсевима:**

1. „Basic Training course for the subject FAW (Forensic Acoustic Workstation), including SIS II, Sound Cleaner, EdiTracker, STC-H246”, by Speech Technology Center,

performed in Belgrade (Serbia) in the period from 24th of February 2014. till 28th of February 2014.

Током дипломских и постипломских студија, Ђорђе Гроздић је до данас објавио укупно 21 рад, од којих су за предложену тему докторске дисертације релевантни следећи:

**- у домаћим часописима:**

1. Grozdić Đ.T., Galić J., Marković B., Jovičić S.T.: *Application of Neural Networks in Whispered Speech Recognition*, Telfor Journal, Vol. 5, No. 2, 2013, ISSN 1821-3251, 103-106.

**- у монографским студијама:**

1. Grozdić Đ.T., Jovičić S.T.: *Role of speech quality features and accented syllables in strong emotion discrimination*, In S.T. Jovičić, M. Subotić (eds): VERBAL COMMUNICATION QUALITY, Interdisciplinary Research, I; CUŽA, Belgrade, ISBN 987-86-81879-34-4, 19-41, 2011.
2. Grozdić Đ.T., Jovičić S.T., Galić J., Marković B.: *Experiments in whisper recognition using neural networks*. In: S.T. Jovičić, M. Subotić, M. Sovilj (eds): VERBAL COMMUNICATION QUALITY, Interdisciplinary Research, II; CUŽA, Belgrade, ISBN 978-86-81879-46-7, 91-110, 2013.

**- на међународним конференцијама:**

1. Grozdić Đ.T., Marković B., Galić J., Jovičić S.T.: *Primena neuralnih mreža u prepoznavanju govora u šapatu*, Proceedings of Int. Sc. Conf. TELFOR, Beograd 2012., 728-801. (Рад је добио награду „Блажо Мирчевски” као најбољи рад младог аутора на конференцији.)
2. Dimitrijević S.M., Grozdić Đ.T.: *Poređenje izmerenih nivoa zvuka u prostori korišćenjem klasične i swept-sine metode*, Processing of Int. Sc. Conf TELFOR, Beograd, 2012, 1264-1267.
3. Marković B., Galić J., Grozdić Đ., Jovičić S.: *Whispered Speech Database: Design, Processing and Application*, The 16th International Conference TSD 2013, in: Habernal and V. Matousek (Eds.): TSD 2013, LNAI 8082, Springer-Verlag Berlin Heidelberg, 591-598.
4. Grozdić Đ.T., Marković B., Galić J., Jovičić S., Furundžić D.: *Neural Network-Based Recognition of Whispered Speech*, Proceedings of Int. Sc. Conf. Speech and Language, Belgrade 2013, 223-229.
5. Galić J., Marković B., Grozdić Đ., Jovičić S.: *The Influence of Feature Vector Selection on Performance of Automatic Recognition of Whispered Speech*, Proceedings of Int. Sc. Conf. Speech and Language, Belgrade 2013, 258-264.
6. Marković B., Galić J., Grozdić Đ., Jovičić S.: *Application of DTW method for Whispered Speech Recognition*, Proceedings of Int. Sc. Conf. Speech and Language, Belgrade 2013, 308-315.
7. Furundžić D., Jovičić S., Subotić M., Grozdić Đ.: *Evaluation of Phonemes Quality Articulation Using Neural Network Ensembles*, Proceedings of Int. Sc. Conf. Speech and Language, Belgrade 2013, 181-189.
8. Jovičić S.T., Grozdić Đ.: *Arguments for auditory-instrumental approach in forensic speaker recognition*, Proceedings of Int. Sc. Conf. Archibald Reiss Days, Belgrade 2014.
9. Marković B.R., Grozdić Đ.T.: *The LPCC-DTW analysis for whispered speech recognition*, Proceedings of Int. Sc. Conf. ETRAN, Vrnjačka Banja 2014.

10. Galić J., Jovičić S.T., Grozdić Đ.T., Marković B.: *HTK-based recognition of whispered speech*, Proceedings of the 16th International Conference Speech and Computer, SPECOM 2014, Novi Sad 2014.
11. Pantelić F., Grozdić Đ., Bjelić M., Ristanović I., Šumarac Pavlović D., Mijić M.: *Analysis of sound field at low frequencies - empty vs. furnished room*, Forum Acusticum, Krakow 2014, Proceedings.
12. Grozdić Đ.T., Jovičić S., Galić J., Marković B.: *Inverse filtering in whisper recognition*, NEUREL 2014, Belgrade. (predat rad)

**- на домаћим конференцијама:**

1. Grozdić Đ.T., Jovičić S.T.: *Multidimenzionalna analiza akustičkih obeležja u govornom signalu iz mobilnog telefona*, Zbornik radova sa konferencije DOGS, Kovačica 2010, 33-36, 2010.
2. Grozdić Đ.T., Jovičić S.T., Rajković M.: *Uticaj kvaliteta glasa na verbalnu ekspresiju emocija*, Zbornik radova sa konferencije TELFOR, Beograd 2011, 663-666.
3. Grozdić Đ.T., Jovičić S.T.: *Primena ANOVA i PCA statističkih metoda u analizi akustičkih obeležja govora*, Zbornik radova sa konferencije DOGS, Kovačica 2012, 182-185.
4. Galić J., Popović M., Marković B., Grozdić Đ.T., Jovičić S.T.: *Primjena skrivenih Markovljevih modela u prepoznavanju govora u šapatu*, Zbornik radova sa konferencije INFOTEH, Jahorina 2013, 387-390.
5. Dimitrijević S., Mašović D., Grozdić Đ.T.: *Funkcionalnost softvera za merenje i analizu zvučne izolacije*, Zbornik radova sa konferencije INFOTEH, Jahorina 2013, 971-974.
6. Mijić M., Ristanović I., Grozdić Đ.: *Raspodela zvučnog polja na niskim frekvencijama u realnim prostorijama*, Zbornik radova sa konferencije ETRAN, Vrnjačka Banja 2014.

### 1.3. Оцена подобности кандидата за рад на предложеној теми

Комисија је размотрила све претходно наведене резултате досадашњег рада и стечена искуства кандидата Ђорђа Гроздића, како у својству студента докторских академских студија, тако и у својству истраживача у Центру за унапређење животних активности као и на Електротехничком факултету у Београду. На основу тога Комисија је стекла увид у способности кандидата за научно-истраживачки рад и његове могућности да током израде докторске дисертације оствари резултате који ће бити приказани и оцењени од стране међународне научне заједнице. Став Комисије је да Ђорђе Гроздић испуњава потребне услове да приступи изради докторске дисертације под наведеном темом.

## **2. Предмет и циљ истраживања**

Говорне технологије, препознавање и синтеза говора, развијају се на стандардном модалитету говора а то је неутралан говор, и ту су постигнути задовољавајући резултати и значајне комерцијалне примене. Међутим појава и других модалитета говора у уобичајеној говорној комуникацији значајно умањује проценат препознавања и квалитет синтетизованог говора. Реално се говорна експресија може реализовати кроз различите модалитете почев од шапата, тихог говора, неутралног говора, експресивног (афективног и емотивног) говора, певаног говора и вике [1]. Због тога се последњих година фокус истраживања у говорним технологијама помера ка овим модалитетима, пре свега ка препознавању и синтези емотивног говора [2]. Од свих модалитета шапат се разликује по беззвучности говора. Чињеница да шапат има веома велику разумљивост и да се често користи у говорној комуникацији наметнула је истраживања и у овом правцу. Започета су истраживања у домену реконструкције говора из шапата за медицинске потребе [3],

анализе и препознавања говора шапата [4], детекције сегмената шапата у континуалном говору [5], као и препознавање говорника који говори шапатам [6]. Ова и слична истраживања шапата су на самом почетку и веома мало радова постоје у литератури.

Предмет истраживања у овом докторату је испитивање могућности примене вештачких неуралних мрежа (ANN - *Artificial Neural Network*) у аутоматском препознавању шапата. Примена ANN у препознавању говора је одавно започета [7] и још траје [8], али не и у препознавању шапата [9]. У овом истраживању зацртана су два основна циља. Први се односи на анализу суштинских карактеристика шапата које га чине са једне стране различитим од нормалног говора а са друге стране му обезбеђују велику разумљивост. Анализа је неопходна како би се нашла најпогоднија акустичка обележја за квалитетно препознавање шапата. Други циљ је анализа неусаглашених обука/тест сценарија (када се неурална мрежа обучава на говор а тестира на шапат и обрнуто). Ова анализа треба да приближи поступак препознавања шапата реалним захтевима препознавања комуникативног говора у коме се могу наћи истовремено оба модалитета и говор и шапат.

Истраживање подразумева анализу и тестирање више метода предобrade сигнала говора и шапата у екстракцији различитих обележја за потребе квалитетног препознавања шапата и компарацију различитих архитектура неуралних мрежа. Посебан акценат истраживања је на анализи неусаглашених обука/тест сценарија, јер су прелиминарна истраживања [9] показала да се суштинска разлика говора и шапата највише одражава у експерименталним резултатима ових сценарија.

Значај овог истраживања је вишеструк. Прво, за потребе истраживања је неопходно формирати говорну базу шапата што за српски језик за сада не постоји (прва форма базе је већ формирана [10]). Оваква говорна база треба да буде веома захвалан медијум за будућа експериментална и теоријска истраживања шапата на дубљем плану, све до когнитивног нивоа. Друго, у истраживању ће се применити ANN за препознавање говора/шапата и очекује се да ће се доћи до показатеља успешности препознавања који треба да буду основа за компаративну анализу касније примене и других алгоритама препознавања (на пример НММ, GMM). И треће, истраживање треба да понуди основе за побољшање актуелног препознавања комуникативног говора који поред неутралног говора садржи и сегменте шапата.

#### **Релевантни библиографски извори за предложено истраживање:**

- [1] C. Zhang, J.H.L. Hansen, "Analysis and Classification of Speech Mode: Whisper through Shouted," pp. 2289-2292, INTERSPEECH 2007.
- [2] M.E. Ayadi, M.S. Kamel, F. Karray, "Survey on speech emotion recognition: Features, classification schemes and databases," *Pattern Recognition*, vol. 44, pp. 572-587, 2011.
- [3] R. W. Morris, M. A. Clements, "Reconstruction of speech from whispers", *Medical Engineering & Physics*, vol. 24, pp. 515-520, 2002.
- [4] T. Ito, K. Takeda and F. Itakura, "Analysis and Recognition of Whispered Speech," *Speech Communication*, 45, pp.129-152, 2005.
- [5] C. Zhang, J.H.L. Hansen, "Advancement in whisper-island detection with normally phonated audio streams", in: *Proceedings of INTERSPEECH*, Brighton, UK, pp. 860-863, 2009.
- [6] X. Fan, J.H.L. Hansen, "Speaker identification within whispered speech audio stream," *IEEE Transactions on Audio, Speech and Language Processing*, 19(5), 1408-1421, 2011.
- [7] C.J. Wellekens, "Introduction to Speech Recognition Using Neural Networks", in: *Proceedings of ESANN 1998, European Symposium on Artificial Neural Networks*, Bruges (Belgium), 1998.

- [8] D. Gulin, H.S. Murat, "Speech recognition with artificial neural networks", *Digital Signal Processing* (Elsevier), Vol. 20, pp. 763-768, 2010.
- [9] Đ. T. Grozdić, J. Galić, B. Marković, S.T. Jovičić, "Application of neural networks in whispered speech recognition," *Telfor Journal*, Belgrade, pp. 103-106, 2013.
- [10] B. Marković, S. Jovičić, J. Galić, Đ. Grozdić, "Whispered speech database: design, processing and application", in: *The 16th International Conference TSD 2013*, I. Habernal and V. Matousek (Eds.): *TSD 2013*, LNAI 8082, Springer-Verlag, Berlin, Heidelberg, pp. 591–598, 2013.

### 3. Полазне хипотезе

Основне хипотезе предложене докторске дисертације су:

1. Постоји могућност успешног препознавања шапата компарабилног препознавању говора применом вештачких неуралних мрежа.

Шапат је због својих специфичности и битних разлика у односу на говор проблематичан у аутоматском препознавању помоћу традиционалних ASR (*Automatic Speech Recognition*) система, који су углавном засновани на HMM (*Hidden Markov Model*) и GMM (*Gaussian Mixture Model*) алгоритмима. Неуралне мреже, тачније MLP (*Multilayer Perceptron*), могу апроксимирати било коју функцију, што их чини универзалним апроксиматорима. Због својих великих могућности у решавању "pattern recognition" проблема, али и услед аналогije и сличности са човековим нервним системом, постављена је хипотеза да вештачке неуралне мреже могу дати добре резултате у препознавању шапата у поређењу са препознавањем говора. При томе има се у виду чињеница да перцептивни механизам човека са великом вероватноћом препознаје шапат. У доказу ове хипотезе неће се ићи дубље у анализу других облика неуралних мрежа и њихових структура у циљу максимизирања препознавања, што може бити предмет дубљих истраживања.

2. *Teager* говорна обележја представљају боље дескрипторе за аутоматско препознавање шапата од традиционалних MFCC обележја.

Данашњи системи за аутоматско препознавање говора су углавном засновани на коришћењу MFCC (*Mel Frequency Cepstral Coefficients*) говорних обележја која нису показала довољну робустност у препознавању шапата. У овом истраживању су предложена два нова типа кепстралних коефицијената TECC (*Teager Energy Cepstral Coefficients*) и TEMFCC (*Teager Energy based Mel Frequency Cepstral Coefficients*), која су се показала као добра решења у препознавању говора у условима лошег односа сигнал/шум (S/N) и амбијенталне буке. Специфичан начин рачунања енергије помоћу ТЕО (*Teager-Energy Operator*), који даје могућност описивања карактеристичне нелинеарне структуре протока ваздуха дуж вокалног тракта током генерисања нетипичних облика говора, као и употреба аудитивне *Gammatone* банке филтара су довољан разлог да се *Teager* обележја испитају у препознавању шапата.

3. Постоје два неусаглашена обука/тест сценарија: говор/шапат и шапат/говор; деградација препознавања шапата у сценарију говор/шапат је последица маскирања звучношћу одређених обележја у говору.

Акустичке карактеристике безвучних гласова су веома сличне у говору и шапату, што се никако не може тврдити и за звучне гласове. Анализа неусаглашених сценарија је показала боље препознавање говора када је неурална мрежа обучена на шапат, него шапата када је мрежа обучена на говор. Овај феномен несиметрије у препознавању указује на могућност маскирања безвучних обележја у говору путем звучних обележја, док овог маскирања у шапату нема. Наиме, говор садржи сва обележја која има шапат плус обележја звучности те када се мрежа обучи на говор а на улаз дође шапат, у коме нема обележја звучности,

тада долази до веће конфузије и смањеног препознавања. Истраживања овог становишта треба да потврде постављену хипотезу.

#### 4. Постоји могућност побољшања аутоматског препознавања бимодалног говора у фази предобrade говорног сигнала

Бимодалност говора у комуникативном смислу подразумева појаву шапата у току нормалног говора, тако да се на улаз препознавача појављују два модалитета говора. То јесте проблем за актуелне препознаваче и питање је колико екстактована обележја у говору могу бити ваљани репрезенти и за шапат. Хипотеза полази од претпоставке да се одређеном предобрадом бимодалног говорног сигнала екстактована обележја у говору и шапату могу довести у позицију тако да не кваре препознавање говора а да побољшавају препознавање шапата. При томе није идеја да се говор адаптира на шапат нити шапат на говор, то је посебан проблем, већ да се максимално искористе заједничка обележја говора и шапата.

#### 5. Постоји могућност побољшања аутоматског препознавања бимодалног говора у фази обуке препознавача.

Поред другачије предобrade говорних сигнала и екстракције одговарајућих говорних обележја, препознавање бимодалног говора (говор-шапат) може се побољшати и посебним методама обуке препознавача. MLP се могу адаптирати на препознавање и говора и шапата унијом обележја шапата и говора у процесу обуке. Поставља се питање колика је минимална потребна количина података шапата, у оваквом процесу обуке, која би допринела довољном побољшању препознавања бимодалног говора.

### 4. Научне методе истраживања

Неколико научних метода биће примењено током реализације научних резултата, у циљу провере наведених полазних хипотеза и остварења задатих истраживачких циљева:

- формирање говорне базе стимулуса која садржи корпус речи нормалног изговора и изговора у шапату (снимање више говорника, сегментација стимулуса, анализа и оптимизација квалитета снимака), планира се обим базе од 10.000 речи,
- софтверска реализација предобrade говорних стимулуса (примена више математичких алгоритама за екстракцију говорних обележја у параметарском домену),
- софтверска реализација симулације неуралних мрежа, оптимизација њихових архитектура и обуке,
- статистичка обрада резултата препознавања неуралних мрежа, статистичка анализа препознавања и конфузије стимулуса на бази матрица конфузија.

Одговарајућа софтверска решења биће реализована коришћењем Matlab програмског језика а за статистичку анализу користиће се SPSS софтвер.

### 5. Очекивани научни допринос

Данашњи системи за аутоматско препознавање говора су осетљиви и непоуздани 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 у циљу постизања боље комуникације човек-рачунар. Препознавање бимодалног говора типа говор-шапат је тек начета истраживачка тема и у литератури постоји свега неколико радова са иницијалним резултатима истраживања.

У складу са тим, очекивани су следећи научни доприноси докторске дисертације:

### 1. Суштинско разумевање својстава шапата у бимодалној говорној комуникацији.

Очекују се теоријска појашњења велике разумљивости коју има шапат а самим тим и интеракције акустичких обележја која карактеришу шапат и говор у бимодалној (говор-шапат) комуникацији.

### 2. Разјашњење слабијих резултата препознавања у неусаглашеним обука/тест (говор/шапат и шапат/говор) сценаријима и њихово побољшање.

Неусаглашене сценарије намеће реалност у примени говорних технологија а експериментални резултати показују слабије препознавање у односу на стандардне усаглашене обука/тест сценарије. Очекује се теоријско разјашњење ових резултата и експериментална верификација идеја за њихово побољшање.

### 3. Побољшање квалитета комуникације човек-рачунар.

Већи број модалитета говора представља велики проблем за системе аутоматског препознавања говора. Резултати истраживања препознавања бимодалног говора на релацији неутралан говор-шапат треба да разреши бар део овог проблема. Резултати добијени уз помоћ MLP треба да буду основа за даље анализе другачијих структура неуралних мрежа и њиховој оптимизацији у препознавању шапата као и у примени других алгоритама препознавања бимодалног говора.

## 6. План истраживања и структура рада

Предложено истраживање обухватаће следеће активности:

- формирање базе снимака изговора речи у нормалном говору и шапату, сегментација снимака и контрола квалитета снимака,
- анализа акустичких карактеристика шапата и компарација са карактеристикама нормалног говора,
- преглед досадашњих истраживања у области аутоматског препознавања говора применом неуралних мрежа,
- израда софтвера за аутоматску предобраду говорних стимулуса из базе снимака и екстракцију говорних обележја,
- израда софтвера за дизајнирање и оптимизацију неуралних мрежа,
- израда софтвера за обуку неуралних мрежа и њихово тестирање,
- анализа различитих архитектура неуралних мрежа (број скривених слојева неуралне мреже, број скривених неурона),
- анализа различитих метода сегментације речи и њиховог временског усклађивања,
- анализа РСА компресије и примене ансамбла неуралних мрежа,
- оптимизација успеха препознавања изолованих речи додавањем динамичких говорних обележја,
- компарација употребе различитих говорних обележја и њихових резултата у препознавању говора и шапата.
- анализа препознавања говора и шапата у усаглашеним обука/тест сценаријима,
- анализа препознавања говора и шапата у неусаглашеним обука/тест сценаријима,
- анализа грешака у препознавању на бази матрица конфузија,
- анализа критичних парова речи на бази спектрограма и акустичких обележја,
- анализа оптимизације и побољшања препознавања бимодалног говора.

## 7. Закључак и предлог

Након увида у наведене резултате досадашњег рада кандидата Ђорђа Гроздића, Комисија је донела одлуку да је он способан за научно-истраживачки рад и за реализацију

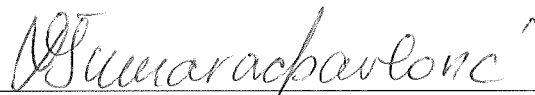
предвиђених истраживања. Одлука је донета имајући у виду његове досадашње стручне резултате остварене током основних и мастер студија, у фази истраживања на докторским студијама, као и његовог истраживачког ангажовања на неколико пројеката и техничких решења рађених у оквиру Центра за унапређење животних активности.

Такође, Комисија је мишљења да је предложена тема докторске дисертације „Примена неуралних мрежа у препознавању шапата“ подобна у смислу наведених непосредних циљева рада, очекиваних резултата и научних доприноса, као и у контексту савремених научних истраживања у предложеној области.

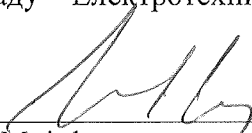
Комисија констатује да тема рада припада области Техничке акустике за коју је Електротехнички факултет у Београду матичан и за ментора рада се предлаже др Драгана Шумарац Павловић, ванредни професор, која је запослена на Електротехничком факултету у Београду. Из тог разлога уз овај Извештај приложен је и списак радова проф. др Драгане Шумарац Павловић који је квалификују за ментора.

Имајући у виду чињенице наведене у овом извештају, Комисија предлаже Наставно-научном већу Електротехничког факултета у Београду да прихвати тему докторске дисертације „Примена неуралних мрежа у препознавању шапата“ и да кандидату Ђорђу Гроздићу омогући да приступи њеној изради.

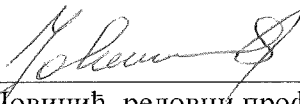
#### ЧЛАНОВИ КОМИСИЈЕ



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



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



др Слободан Јовичић, редовни професор у пензији  
Универзитет у Београду – Електротехнички факултет



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