

ИЗВЕШТАЈ О ОЦЕНИ ПОДОБНОСТИ КАНДИДАТА, ТЕМЕ И МЕНТОРА ЗА
ИЗРАДУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

I ПОДАЦИ О КОМИСИЈИ

Орган који је именовао комисију: Декан ФТН, на основу Одлуке НН већа број 012-199/33-2019 и у складу са Статутом

Датум именовања комисије: 27.06.2019.

Састав комисије:

- | | | |
|--|-------------------|-----------------------------------|
| 1. Др Милан Сечујски | ванредни професор | Телекомуникације и обрада сигнала |
| презиме и име | звање | ужа научна област |
| Факултет техничких наука, Нови Сад | | председник |
| установа у којој је запослен-а | | функција у комисији |
| 2. Др Татјана Грбић | редовни професор | Теоријска и примењена математика |
| презиме и име | звање | ужа научна област |
| Факултет техничких наука, Нови Сад | | члан |
| установа у којој је запослен-а | | функција у комисији |
| 3. Др Јернеја Жганец-Грос | редовни професор | Информациони системи |
| презиме и име | звање | ужа научна област |
| Факултет за индустријски инжењеринг, Ново место, Словенија | | члан |
| установа у којој је запослен-а | | функција у комисији |
| 4. Др Никша Јаковљевић | доцент | Телекомуникације и обрада сигнала |
| презиме и име | звање | ужа научна област |
| Факултет техничких наука, Нови Сад | | члан |
| установа у којој је запослен-а | | функција у комисији |
| 5. Др Бранислав Поповић | научни сарадник | Телекомуникације и обрада сигнала |
| презиме и име | звање | ужа научна област |
| Факултет техничких наука, Нови Сад | | члан |
| установа у којој је запослен-а | | функција у комисији |

II ПОДАЦИ О КАНДИДАТУ

1. Име, име једног родитеља, презиме: Дарко, Јован, Пекар
2. Датум рођења: 06.03.1972. Место и држава рођења: Нови Сад, Србија

II.1 основне студије

година уписа: година завршетка: просечна оцена током студија:

универзитет: Универзитет у Новом Саду

факултет: Факултет техничких наука

студијски програм: Енергетика, електроника и телекомуникације

звање: Дипломирани инжењер електротехнике и рачунарства

наслов завршног рада: Препознавање речи преко парова фонема као јединице препознавања

II.2 мастер или магистарске студије

година уписа: година завршетка: просечна оцена током студија:

универзитет: Универзитет у Новом Саду

факултет: Факултет техничких наука

студијски програм: Енергетика, електроника и телекомуникације

звање: Мастер инжењер електротехнике и рачунарства

научна област: Телекомуникације и обрада сигнала

наслов завршног рада: Препознавање речи преко парова фонема као јединице препознавања

II.3 докторске студије

година уписа:

универзитет: Универзитет у Новом Саду

факултет: Факултет техничких наука

студијски програм: Енергетика, електроника и телекомуникације

број ЕСПБ до сада остварених: просечна оцена током студија:

II.4 приказ научних и стручних радова

Р. бр.	аутори, наслов, издавач, број страница	категорија
1.	Branislav Popović, Marko Janev, Darko Pekar , Nikša Jakovljević, Milan Gnjatović, Milan Sečujski, Vlado Delić (2012) "A Novel Split-and-Merge Algorithm for Hierarchical Clustering of Gaussian Mixture Models", Applied Intelligence, ISSN: 0924-669X, Springer-Verlag New York, Inc., Vol. 37, No. 3, pp. 377-389, DOI: 10.1007/s10489-011-0333-9	M21
<i>кратак опис садржине:</i> У раду је представљена нова метода растављања и спајања за хијерархијско кластеровање модела са Гаусовим смешама, што има тенденцију да унапреди локално оптимално решење добијено иницијалном поставком. Иницијализовано је преко локалних оптималних параметара добијених коришћењем основног приступа скличног k-means алгоритму, и у већој мери се приближава глобалном оптимуму циљне функције кластеровања, тако што итеративно дели и спаја кластере Гаусових компоненти добијених као излаз основног алгоритма. Предложена метода је тестирана како на вештачки генерисаним подацима, тако и на реалном систему за препознавање говора, и у оба случаја је примећено побољшање у односу на основни метод.		
<i>рад припада проблематици докторске дисертације:</i> ДА НЕ <u>ДЕЛИМИЧНО</u>		

Р. бр.	аутори, наслов, издавач, број страница	категорија
2.	Delić, Tijana V., Siniša B. Suzić, Stevan J. Ostrogonac, Simona V. Đurić, and Darko J. Pekar . "Multi-style Statistical Parametric TTS.", DOGS2017, Нови Сад, 2017.	M63
<i>кратак опис садржине:</i> Једна од предности параметарских метода синтезе говора је у њиховиј могућности да лако мењају карактеристике говорника и стилова. У раду је представљена једноставна метода која укључује стилове у синтетизовани говор преко кодова стилова. Предложена метода је успешно примењена како на скривене Марковљеве моделе, тако и на неуронске мреже. Ефикасност је показана преко одговарајућих тестова слушања. Тестови слушања такође сугеришу да је боља експресивност постигнута коришћењем неуронских мрежа као да је квалитет синтезе стилова упоредив са синтезом у неутралном стилу, иако је база неутралног говора 10 пута већа.		
<i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
3.	T. Delić, S. Suzić, M. Sečujski, D. Pekar , "Rapid Development of New TTS Voices by Neural Network Adaptation," in 2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH), 2018, pp. 1–6.	M33
<i>кратак опис садржине:</i> У раду се описује једноставна, али успешна метода креирања нових TTS говорника, базираних на неуронским мрежама, коришћењем мале количине аудио материјала. Говорни подаци циљаног говорника се користе у сврху адаптације неуронске мреже већ трениране на изворном говорнику, радије него да се тренира мрежа од нуле. Овај приступ смањује количину адаптационог материјала неопходног за продукцију висококвалитетне синтезе новог говорника. Резултати су евалуирани како објективним тако и субјективним методама и показано је да адаптирање параметара постојећег модела са само 10 минута говора новог говорника генерише синтетички говор упоредивог квалитета са оним који би се добио коришћењем 3 сата материјала и конвенционални тренинг.		
<i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
--------	--	------------

4.	Suzić, Siniša, Tijana Delić, Vladimir Jovanović, Milan Sečujski, Darko Pekar , and Vlado Delić. "A comparison of multi-style DNN-based TTS approaches using small datasets." In MATEC Web of Conferences, vol. 161, p. 03005. EDP Sciences, 2018.	M33
<i>кратак опис садржине:</i> Једна од главних мана TTS система је недостатак експресивности у поређењу са природним људским говором. Увођење параметарских метода довело је до могућности да се лако могу мењати карактеристике говорника и стилова говора. У раду је предложен једноставан метод увођења тзв. кодова стила. Предложени метод захтева свега неколико минута говора у циљаном стилу и умерену количину неутралног говора. Тестови слушања потврдили су да се већа експресивност постиже коришћењем неуронских мрежа у поређењу са скривеним Марковљевим моделима. Такође се показало да је квалитет синтезе стилова упоредив са синтезом у неутралном стилу, иако је база неутралног говора 10 пута већа. <i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
5.	Siniša Suzić, Tijana Delić, Vladimir Jovanović, Milan Sečujski, Darko Pekar , and Vlado Delić. "A comparison of multi-style DNN-based TTS approaches using small datasets." In MATEC Web of Conferences, vol. 161, p. 03005. EDP Sciences, 2018.	M33
<i>кратак опис садржине:</i> Главна препрека у креирању експресивне синтезе говора је та да количина материјала потребна за тренирање таквих система углавном није довољна. Овај рад упоређује различите приступе у синтези са више стилова, са фокусом на случајеве где је расположива само мала база по сваком стилу. Описане методе су оригинално биле предложене за ефикасно моделовање више говорника са ограниченом количином материјала по говорнику. Међу предложеним методама, приступ базиран на кодовима стила се показао као најбољи, без обзира на циљани стил говора. <i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
6.	E. Pakoci, B. Popović, D. Pekar , "Using Morphological Data in Language Modeling for Serbian Large Vocabulary Speech Recognition", Computational Intelligence and Neuroscience, Advanced Signal Processing and Adaptive Learning Methods, vol. 2019, 2019, pp. 1-8	M22
<i>кратак опис садржине:</i> У раду су приказани резултати примене нових модела језика за српски у систему за аутоматско препознавање говора на великим речницима, који су базирани на <i>n</i>-грамима, односно на дубоким рекурентним неуронским мрежама. Нови алати за креирање језичких модела на бази неуронских мрежа допуштају коришћење и додатних лексичких и морфолошких обележја у моделу, осим самог идентитета речи. Описан је начин за одређивање жељених морфолошких категорија у корпусу за обуку модела, и један начин инкорпорирања додатних морфолошких обележја у сам модел. Резултујући систем омогућава превазилажење дела честих грешака при препознавању говора као што су грешке у падежу, граматичком роду или броју, за морфолошки богате језике као што је српски. Систем је тестиран на највећем постојећем тест скупу за српски, и показао је значајно побољшање у односу на референтни систем. <i>рад припада проблематици докторске дисертације:</i> ДА <u>НЕ</u> ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
7.	M. Sečujski, D. Pekar , R. Mak, E. Pakoci, N. Jakovljević, V. Delić, "A Comparison of Decision Tree Based Approaches to the Modelling of Prosodic Features in Serbian", Verbal Communication Quality, Interdisciplinary Research II, 2013, pp. 75-90	M14

<i>кратак опис садржине:</i> У поглављу су описани упоредни резултати два приступа за кластеровање преко стабла одлуке примењена у моделовању прозодијских обележја у системима за синтезу говора из текста на српском језику. Прецизније, ради се о моделовању интонације и трајања фонетских сегмената. Први алгоритам је примењен у систему за синтезу говора на бази конкатенације говорних сегмената, а други у новом параметарском систему на бази скривених Марковљевих модела. Рађен је и објективни и субјективни тест (тест слушања). Оба теста су јасно показала предност другог приступа. Резултујућа прозодијска обележја у том приступу нису значајно одступала од обележја природног говора.		
<i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
8.	E. Pakoci, B. Popović, D. Pekar , "A Phonetic Segmentation Procedure Based on Hidden Markov Models", 18 th International Conference on Speech and Computer SPECOM 2016, Budapest, Hungary, Lecture Notes in Artificial Intelligence, Vol. 9811, 2016, pp. 67-74	M33
<i>кратак опис садржине:</i> У раду је описана нова варијанта процедуре за аутоматско временско поравнање произвољних аудио база података, која може бити корисна у разним апликацијама. Испитано је и неколико могућих метода модификације основне процедуре. Експерименти поређења аутоматског поравнања са ручним на датом тест скупу показали су веома добре резултате, са готово 90% фонема у оквирима од 20 милисекунди од ручно одређене позиције.		
<i>рад припада проблематици докторске дисертације:</i> <u>ДА</u> НЕ ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
9.	B. Popović, E. Pakoci, N. Jakovljević, G. Kočiš, D. Pekar , "Voice Assistant Application for the Serbian Language", 23 rd Telecommunications Forum TELFOR 2015, Belgrade, Serbia, 2015, pp. 858-861	M33
<i>кратак опис садржине:</i> У раду је описана апликација говорног асистента на српском језику за мобилне телефоне са Андроид оперативним системом. За ове потребе развијен је C++ интерфејс за декодер за аутоматско препознавање говора базиран на Калди алатима. Приказани су и резултати тачности препознавања речи на тест скупу од око 4500 реченица и речнику од преко 14000 речи, прилагођеним за ове потребе.		
<i>рад припада проблематици докторске дисертације:</i> ДА <u>НЕ</u> ДЕЛИМИЧНО		

Р. бр.	аутори, наслов, издавач, број страница	категорија
10.	E. Pakoci, N. Jakovljević, B. Popović, D. Mišković, D. Pekar , "Speaker Detection Using Phoneme Specific Hidden Markov Models", 16 th International Conference on Speech and Computer SPECOM 2014, Novi Sad, Serbia, Lecture Notes in Artificial Intelligence, Vol. 8773, 2014, pp. 410-417	M33
<i>кратак опис садржине:</i> У раду је презентован систем за детекцију циљаних говорника у телефонским снимцима базиран на скривеним Марковљевим моделима на нивоу фонема у комбинацији са моделима Гаусових мешавина. Детекција говорника се ослања на разлику акустичког скорa у појединим деловима сигнала између универзалног модела и модела адаптираног на жељеног говорника уз помоћ врло мале количине аудио материјала. Неколико варијанти адаптације је испитано, а рађено је и поређење са референтним системом базираном само на Гаусовим мешавинама са дијагоналним коваријансним матрицама, у односу на који је добијено значајно побољшање.		

рад припада проблематици докторске дисертације:	ДА	НЕ	<u>ДЕЛИМИЧНО</u>
---	----	----	------------------

Р. бр.	аутори, наслов, издавач, број страница	категорија
11.	D. Pekar , D. Mišković, S. Suzić, E. Pakoci, V. Delić, “Axon Voice Assistant - an Android-Based Smartphone Application in Serbian”, 11 th International Conference on Electronics, Telecommunications, Automation and Informatics ETAI 2013, Ohrid, North Macedonia, 2013, pp 1-6	M33

кратак опис садржине: У раду је представљен говорни асистент на српском језику реализован за телефоне који користе оперативни систем Андроид. Изложени су сви аспекти развоја ове апликације: дијалошки систем, систем за препознавање говора, као и систем за синтезу говора. Предност предложеног решења је његово лако прилагођавање осталим јужнословенским језицима.

рад припада проблематици докторске дисертације:	ДА	<u>НЕ</u>	ДЕЛИМИЧНО
---	----	-----------	-----------

Р. бр.	аутори, наслов, издавач, број страница	категорија
12.	В. Поповић, Е. Pakoci, N. Jakovljević, D. Pekar , “Experiments With Regression Trees for Speaker Adaptation in Continuous Speech Recognition”, 10 th Conference on Digital Speech and Image Processing DOGS 2014, Novi Sad, Serbia, 2014, pp. 47-50	M63

кратак опис садржине: У раду су упоређене перформансе различитих регресионих стабала одлуке у оквиру адаптације на циљаног говорника у систему за аутоматско препознавање говора на српском језику. Тестиран је утицај броја регресионих класа стабла, врсте трансформационих матрица, врсте мере растојања при кластеровању података и броја одбирака за адаптацију на говорника. Тестови су рађени почевши од универзалног модела независног од говорника обученог са око 12 сати говорног материјала, за по 50 мушких и женских говорника, са речником од око 15000 речи.

рад припада проблематици докторске дисертације:	ДА	НЕ	<u>ДЕЛИМИЧНО</u>
---	----	----	------------------

III ОЦЕНА ПОДОБНОСТИ КАНДИДАТА

III.1 услови дефинисани за кандидата студијским програмом:

На основу чланова 65 и 100 став 8. Закона о високом образовању ("Сл. гласник РС" бр. 88/2017 и 27/2018-др. закон), чланова 62 и 149-156 статута Факултета техничких наука (пречишћен текст 4.12.2015) а у вези са чланом 1. став 2 правила докторских академских студија универзитета у Новом Саду (пречишћен текст од 29.03.2018) као и „Правилника о упису, студирању на докторским академским студијама и стицању звања доктора наука, односно доктора уметности“ Факултета техничких наука, право да пријави тему докторске дисертације стиче студент који је положио све испите одређене студијским програмом) и који је одбранио Теоријске основе докторске дисертације.

III.2 Да ли кандидат испуњава услове?

ДА

НЕ

Образложење:

Кандидат је успешно савладао градиво предвиђено програмом за докторске студије и положио све испите са просечном оценом 10,00. Такође је успешно одбранио и Теоријске основе докторске дисертације (квалификациони испит). Кандидат је такође аутор или коаутор преко 80 научних радова, од којих је 6 радова у директној вези са темом дисертације, а још 3 су делимично повезана. Кандидат је током докторских студија као истраживач био ангажован и на неколико пројеката Министарства просвете, науке и технолошког развоја Републике Србије: ТР32035 – „Развој дијалогских система за српски и друге јужнословенске језике“ (од 2013. године), ЦАБУНС – „Централна аудио-библиотека Универзитета у Новом Саду“ (од 2016. године), те на међународном Еурека пројекту ДАНСПЛАТ – „Платформа за апликације говорних апликација на паметним телефонима за језике Дунавске регије“ (од 2016. године до 2019. године)

IV ОЦЕНА ПОДОБНОСТИ ПРЕДЛОЖЕНИХ МЕНТОРА

IV.1 ПОДАЦИ О МЕНТОРИМА

Први ментор, **Владо Делић** рођен је у Новом Саду 1964. год. Дипломирао је на ФТН 1989, магистрирао 1993. на ЕТФ у Београду и докторирао 1997. године на ФТН у Новом Саду, у области обраде сигнала и говорних технологија. У звање редовног професора изабран 2013. године, а од 2015. је шеф Катедре за телекомуникације и обраду сигнала – колектива од преко 40 наставника и сарадника.

Кроз свој научно-истраживачки рад привукао је и окупио тим изузетних младих стручњака и са њима изградио позицију лидера у развоју и примени говорних технологија у региону јужнословенских језика. Његов научни рад праћен је са преко 350 публикација међу којима је више од 20 радова у врхунским часописима и монографијама и 10-так шире примењених техничких решења и патената.

Радови из уже научне области којој припада предлог докторске дисертације:

Р. бр.	аутори, наслов, издавач, број страница	категорија
1.	V. Delić , Z. Perić, M. Sečujski, N. Jakovljević, J. Nikolić, D. Mišković, N. Simić, S. Suzić, T. Delić (2019) "Speech technology progress based on new machine learning paradigm", <i>Computational Intelligence and Neuroscience</i> , Special Issue "Advanced Signal Processing and Adaptive Learning Methods," Volume 2019, Review Article ID 4368036, 19 pages. https://doi.org/10.1155/2019/4368036	M22
2.	S. Suzić, T. Delić, M. Sečujski, D. Pekar, V. Delić (2019) "Style Transplantation in Neural Network Based Speech Synthesis", <i>Acta Polytechnica Hungarica</i> , ISSN:1785-8860, Vol. 16, 2019	M23
3.	D. Mišković, M. Gnjatović, P. Štrbac, B. Trenkić, N. Jakovljević, V. Delić (2017) "Hybrid methodological approach to context-dependent speech recognition", <i>International Journal of Advanced Robotic Systems</i> , ISSN: 1729-8814, SAGE Publications, Vol. 14, No. 1, Feb. 1 st , 2017, pp. 1-12, DOI: 10.1177/1729881416687131	
4.	I. Jokić, V. Delić , S. Jokić, Z. Perić (2015) "Automatic Speaker Recognition Dependency on Both the Shape of Auditory Critical Bands and Speaker Discriminative MFCCs", <i>Advances in Electrical and Computer Engineering</i> , ISSN: 1582-7445, Vol. 15, No. 4, pp. 25-32 DOI: 10.4316/AECE.2015.04004	M23
5.	M. Gnjatović, V. Delić (2014) "Cognitively-inspired representational approach to meaning in machine dialogue", <i>Knowledge-Based Systems</i> , Special issue on Cognitive Infocommunications, Elsevier, ISSN: 0950-7051 (IF2012=4.104), Vol. 71, November 2014, pp. 25-33, DOI: 10.1016/j.knosys.2014.05.001	M21a
6.	S. Sovilj-Nikić, V. Delić , I. Sovilj-Nikić, M. Marković (2014) "Tree-based Phone Duration Modelling of the Serbian Language", <i>Elektronika ir Elektrotehnika</i> , ISSN: 1392-1215, Kaunas University of Technology, Vol. 20, No. 3, pp. 77-82, DOI: 10.5755/j01.eee.20.3.4090	M23
7.	N. Jakovljević, D. Mišković, M. Janev, M. Sečujski, V. Delić (2013) "Comparison of Linear Discriminant Analysis Approaches in Automatic Speech Recognition", <i>Elektronika ir Elektrotehnika</i> , ISSN: 1392-1215, Kaunas University of Technology, Vol. 19, No. 7, pp. 76-79, DOI: 10.5755/j01.eee.19.7.5167	M23
8.	V. Delić , M. Bojanić, M. Gnjatović, M. Sečujski, S.T. Jovičić (2012) "Discrimination Capability of Prosodic and Spectral Features for Emotional Speech	M22

	Recognition”, <i>Elektronika Ir Elektrotehnika</i> , Vol. 18, No. 9, November of 2012, pp. 51-54, DOI: 10.5755/j01.eee.18.9.2806	
9.	B. Popović, M. Janev, D. Pekar, N. Jakovljević, M. Gnjatović, M. Sečujski, V. Delić (2012) “A Novel Split-and-Merge Algorithm for Hierarchical Clustering of Gaussian Mixture Models”, <i>Applied Intelligence</i> , Springer-Verlag New York, Inc., ISSN 0924-669X, Published Online First, January of 2012, Vol. 37 No. 3, pp. 377-389, DOI:10.1007/s10489-011-0333-9	M21
10.	M. Gnjatović, M. Janev, V. Delić (2012) “Focus Tree: Modeling Attentional Information in Task-Oriented Human-Machine Interaction”, <i>Applied Intelligence</i> , Springer-Verlag New York, Inc., ISSN 0924-669X, Published Online First, December of 2011, Vol. 37, No. 3, DOI: 10.1007/s10489-011-0329-5	M21

Други ментор, **Марко Јанев**, рођен је 31.05.1977. године у Новом Саду. Завршио је Основну школу "Жарко Зрењанин" у Новом Саду, а потом и средњу Електротехничку школу "Михајло Пупин" у Новом Саду. Дипломирао је 2003. године на Електротехничком одсеку Факултета техничких наука, Универзитет у Новом Саду, на смеру за аутоматско управљање системима, са просечном оценом 8.63. По завршетку основних студија, уписује постдипломске студије на Факултету техничких наука, Универзитет у Новом Саду, на смеру Математика у техници. Магистарску тезу под називом "Уклањање шума са слике коришћењем пуне фракционе неізотропске дифузије", одбранио је на истом смеру, 2009. године. Докторску дисертацију под називом "Примена фракционог рачуна у обради слике" одбранио је на Факултету техничких наука, УНС, 2011. године, под менторством академика Теодора Атанацковића и академика Стевана Пилиповића.

Од 2008. до 2011. био је запослен на Факултету техничких наука, УНС, ангажован као истраживач на пројекту "Интеракција човек-машина" Министарства науке Републике Србије. Од 2011. до данас запослен је на Математичком институту САНУ, у сталном радном односу, и ангажован на пројектима "Развој дијалогских система за српски и друге јужнословенске језике" (TR32035), као и "Интегрисани системи за детекцију и естимацију развоја пожара праћењем критичних параметара у реалном времену" (III44003), Министарства науке и технолошког развоја Републике Србије. Изабран је у звање виши научни сарадник 20.12.2017. године.

Радови из уже научне области којој припада предлог докторске дисертације:

Р. бр.	аутори, наслов, издавач, број страница	категорија
1.	Teodor Atanacković, Marko Janev , Sanja Konjik, Stevan Pilipović, Dušan Zorica, “Expansion formula for fractional derivatives in variational problems“, <i>JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS</i> , Elsevier, (2014), vol. 409 br. 2, str. 911-924, DOI: 10.1016/j.jmaa.2013.07.071	M21
2.	Teodor Atanacković, Marko Janev , Stevan Pilipović, Dušan Zorica, “An expansion formula for fractional derivatives of variable order“, <i>CENTRAL EUROPEAN JOURNAL OF PHYSICS</i> , Springer Verlag, (2013), vol. 11 br. 10, str. 1350-1360, DOI: 10.2478/s11534-013-0243-z	M23
3.	Milan Gnjatović, Marko Janev , Vlado Delić, “Focus tree: modeling attentional information in task-oriented human-machine interaction“, <i>APPLIED INTELLIGENCE</i> , Springer Verlag, (2012), vol. 37 br. 3, str. 305-320, DOI: 10.1007/s10489-011-0329-5	M21
4.	Branislav Popović, Marko Janev , Darko Pekar, Nikša Jakovljević, Milan Gnjatović, Milan Sečujski, Vlado Delić, “A novel split-and-merge algorithm for hierarchical clustering of Gaussian mixture models“, <i>APPLIED INTELLIGENCE</i> , Springer Verlag, (2012), vol. 37 br. 3, str. 377-389, DOI: 10.1007/s10489-011-0333-9	M21
5.	Teodor Atanacković, Marko Janev , Stevan Pilipović, Dušan Zorica, “Complementary variational principles with fractional derivatives“, <i>ACTA MECHANICA</i> , Springer Verlag, (2012), vol. 223 br. 4, str. 685-704, DOI: 10.1007/s00707-011-0588-6	M22

6.	Marko Janev , Stevan Pilipović, Teodor Atanacković, Radovan Obradović, Nebojsa Ralević, "Fully fractional anisotropic diffusion for image denoising", MATHEMATICAL AND COMPUTER MODELLING, Elsevier, (2011), vol. 54 br. 1-2, str. 729-741, DOI: 10.1016/j.mcm.2011.03.017	M21
7.	Marko Janev , Teodor Atanacković, Stevan Pilipović, Radovan Obradović, "Image denoising by a direct variational minimization", EURASIP JOURNAL ON ADVANCES IN SIGNAL PROCESSING, Springer Verlag, (2011), DOI: 10.1186/1687-6180-2011-8	M22
8.	Marko Janev , Darko Pekar, Nikša Jakovljević, Vlado Delić, "Eigenvalues Driven Gaussian Selection in continuous speech recognition using HMMs with full covariance matrices", APPLIED INTELLIGENCE, Springer Verlag, (2010), vol. 33 br. 2, str. 107-116, DOI: 10.1007/s10489-008-0152-9	M23
9.	Teodor Atanacković, Marko Janev , Sanja Konjik, Stevan Pilipović, Dušan Zorica, "Vibrations of an elastic rod on a viscoelastic foundation of complex fractional Kelvin-Voigt type", MECCANICA, Springer Verlag, (2015), vol. 50 br. 7, str. 1679-1692, DOI: 10.1007/s11012-015-0128-x	M21
10.	Teodor Atanacković, Marko Janev , Ljubica Oparnica, Stevan Pilipović, Dušan Zorica, "Space-time fractional Zener wave equation", Proceedings of the Royal Society A, 471 (2015) 201406141-1-25, DOI: 10.1098/rspa.2014.0614	M21

IV.1 услови дефинисани за ментора у области којој припада докторска дисертација:

Према члану 12. став 1 Правила докторских студија УНС од 25.02.2013, за ментора за израду докторске дисертације односно докторског уметничког пројекта може бити именован **наставник Универзитета** који је у радном односу на факултету који реализује студијски програм докторских студија, и који има потребну научну, односно уметничку способност из области проблематике докторске дисертације.

Према члану 12. став 2 истих Правила, **лице изабрано у научно звање** у складу са законом којим је регулисана научноистраживачка делатност, поред права да учествује у извођењу наставе и испита на докторским студијама, да буде члан комисије за оцену подобности кандидата, теме и ментора и члан комисије за оцену и одбрану докторске дисертације, може бити и ментор ако је у сталном радном односу у научноистраживачкој установи.

Према члану 12. став 4 истих Правила, за ментора може бити именован **наставник** односно **научни радник** који је способан за извођење наставе на докторским студијама и има најмање 5 научних радова објављених или прихваћених за објављивање у научним часописима из одговарајуће области студијског програма са листе министарства надлежног за науку, у последњих 10 година и који, у оквиру научног поља техничко-технолошких наука, испуњава додатни критеријум, а то је да има 5 радова објављених у часописима са ISI листе.

IV.2 Да ли ментори испуњавају услове?

ДА

НЕ

Образложење:

Први ментор, **Владо Делић**, запослен је на Факултету техничких наука Универзитета у Новом Саду, налази се на листи акредитованих ментора докторских дисертација, изводи наставу на докторским студијама и у последњих 10 година има више од 5 радова објављених у часописима са ISI листе из одговарајуће области.

Други ментор, **Марко Јанев**, запослен је на Математичком институту САНУ као виши научни сарадник, у сталном је радном односу, и у последњих 10 година има више од 5 радова објављених у часописима са ISI листе, из одговарајуће области.

V ОЦЕНА ПОДОБНОСТИ ТЕМЕ

V.1 формулација назива тезе (насловa)

Нова метода адаптације на говорника у параметарској синтези говора

Novel method for speaker adaptation in parametric speech synthesis

Наслов тезе је подобан?	ДА	НЕ	ДЕЛИМИЧНО
-------------------------	----	----	-----------

V.2 предмета (проблема) истраживања

Синтеза говора на основу текста је технологија која омогућава рачунарима претварање писаног текста у људски говор. Ова технологија у савременом добу има веома широк спектар примена – од помоћи слепим особама и пацијентима са оштећењима гласа до примене у позивним центрима и говорним асистентима на мобилним уређајима.

За практичну применљивост одређеног система веома је битно да не захтева много улагања приликом прављења новог гласа или стила. До недавно се за ове потребе морала снимати потпуно нова говорна база, често већа од 10х, као и уложити дугатрајни рад на анотацији те базе (фонетској и прозодијској). У овом раду су представљене методе које могу да генеришу нови глас или стил користећи много мању количину снимљеног материјала (чак и испод 10 минута).

У раду се предлаже неколико метода и модела адаптације. Планирају се објективна и субјективна поређења свих метода, презентација и дискусија свих резултата.

Предмет истраживања је подобан?	ДА	НЕ	ДЕЛИМИЧНО
---------------------------------	----	----	-----------

V.3 познавања проблематике на основу изабране литературе са списком литературе

Методе за синтетизовање говора на основу текста могу се поделити у две групе. Прву групу чине конкатенативни приступи [1]. Они се заснивају на повезивању одговарајућих говорних сегмената из раније снимљене говорне базе. Такође, манипулације са синтетизованим гласом су ограничене на промену трајања изабраних говорних сегмената или промену основне учестаности [2], а чак и те промене могу довести до деградације квалитета синтезе. Другу групу метода за синтетизовање говора називамо параметарским методама [3]. Оне користе одговарајуће моделе за предикцију одређених обележја говорног сигнала на основу којих је могуће реконструисати говор.

Међу параметарским приступима скоро 20 година били су доминантни статистички приступи на бази скривених Марковљевих модела (енгл. *Hidden Markov Models*, HMM) и модела Гаусових смеша (енгл. *Gaussian Mixture Models*, GMM) [4]. Део проблема присутних код HMM синтезе превазиђен је употребом неуронских мрежа са више слојева (енгл. *Deep Neural Networks*, DNN) [5]. Прве примене адаптације на говорника су биле везане за препознавање говора [9], а када су се и у синтези говора појавили HMM-GMM модели, слична логика је примењена и у овој области [10][11]. Када је у питању DNN приступ, било је покушаја да се искористе искуства стечена адаптацијом HMM-GMM модела, уз евентуалне дообуке система [12]. Још један уобичајени приступ, који није ограничен само на адаптацију, јесте тзв. *transfer learning*, који користи целу или део већ обучене мреже како би је лакше и поузданије обучио за нови задатак [13].

Потреба за висококвалитетном синтезом говора постоји већ дуже време, што се види и из многобројних примена ове технологије. Ова технологија јесте на располагању већ неко време, и то у високом квалитету, али су до сада постојале две главне препреке за њено шире коришћење: одсуство флексибилности, што се огледа у ограниченим могућностима мењања карактеристика говора; скуп и спор процес прављења новог говорника, што је резултирало релативно малим бројем расположивих гласова.

Један од главних проблема приликом креирања новог TTS гласа јесте снимање говорне базе. Процедура снимања и накнадног аотирања базе је дуготрајна и захтева приличне ресурсе, како људске тако и финансијске. Хипотеза која ће бити тестирана јесте да се коришћењем ограничене количине говорног материјала у обуци може добити задовољавајући квалитет синтетизованог говора циљаног говорника. У процесу провере квалитета синтетизованог говора пажња ће бити усмерена на сличност добијене синтезе са циљаним говорником, стабилност синтезе (одсуство повремених артефаката) и наравно природност добијеног гласа.

Проблем адаптације на говорника када је на располагању мала количина материјала се огледа у томе да многе фонетско прозодијске комбинације нису присутне у адаптационом скупу ("нису виђене"). У циљу превазилажења овог проблема предлагане су разне технике, од техника у HMM-GMM приступу [9], преко ослањања на способност DNN да генерализују [15], учења ових релација на бази сачињеној од много говорника [16], примене тзв. *embeddinga*, као и комбинација ових метода.

- [1] A. J. Hunt and A. W. Black, "Unit selection in a concatenative speech synthesis system using a large speech database," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, 1996, vol. 1, pp. 373–376.
- [2] M. Schröder, "Expressive speech synthesis: Past, present, and possible futures," in *Affective information processing*, Springer, 2009, pp. 111–126.
- [3] H. Zen, K. Tokuda, and A. W. Black, "Statistical parametric speech synthesis," *Speech Commun.*, vol. 51, no. 11, pp. 1039–1064, 2009.
- [4] K. Tokuda, Y. Nankaku, T. Toda, H. Zen, J. Yamagishi, and K. Oura, "Speech synthesis based on hidden Markov models," *Proc. IEEE*, vol. 101, no. 5, pp. 1234–1252, 2013.
- [5] H. Ze, A. Senior, and M. Schuster, "Statistical parametric speech synthesis using deep neural networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 7962–7966.
- [6] Y. Qian, Y. Fan, W. Hu, and F. K. Soong, "On the training aspects of deep neural network (DNN) for parametric TTS synthesis," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 3829–3833.
- [7] A. Lazaridis, B. Potard, and P. N. Garner, "DNN-based speech synthesis: Importance of input features and training data," in *International Conference on Speech and Computer*, 2015, pp. 193–200.
- [8] A. van den Oord et al., "Wavenet: A generative model for raw audio," *arXiv Prepr. arXiv1609.03499*, 2016.

- [9] C. J. Leggetter and P. C. Woodland, "Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models," *Comput. Speech Lang.*, vol. 9, no. 2, pp. 171–185, 1995.
- [10] M. Tamura, T. Masuko, K. Tokuda, and T. Kobayashi, "Speaker adaptation for HMM-based speech synthesis system using MLLR," in the third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis, 1998.
- [11] J. Yamagishi and T. Kobayashi, "Average-voice-based speech synthesis using HSMM-based speaker adaptation and adaptive training," *IEICE Trans. Inf. Syst.*, vol. 90, no. 2, pp. 533–543, 2007.
- [12] P. Swietojanski and S. Renals, "Learning hidden unit contributions for unsupervised speaker adaptation of neural network acoustic models," in 2014 IEEE Spoken Language Technology Workshop (SLT), 2014, pp. 171–176.
- [13] H.-C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
- [14] Y. Jia et al., "Caffe: Convolutional architecture for fast feature embedding," in Proceedings of the 22nd ACM international conference on Multimedia, 2014, pp. 675–678.
- [15] C. L. Giles and T. Maxwell, "Learning, invariance, and generalization in high-order neural networks," *Appl. Opt.*, vol. 26, no. 23, pp. 4972–4978, 1987.
- [16] J. Yamagishi, Z. Ling, and S. King, "Robustness of HMM-based speech synthesis," 2008.

Избор литературе је одговарајући? ДА НЕ ДЕЛИМИЧНО

V.4 циљева истраживања

Један од главних проблема приликом креирања новог гласа за потребе синтезе јесте снимање говорне базе. Процедура снимања и накнадног аотирања базе је дуготрајна и захтева приличне ресурсе. У раду ће бити представљене методе које могу да генеришу нови глас или стил користећи много мању количину снимљеног материјала (чак и испод 10 минута). И поред тако мале количине адаптационог материјала, очекивани резултат је квалитетна и стабилна синтеза.

Циљеви истраживања су одговарајући? ДА НЕ ДЕЛИМИЧНО

V.5 очекиваних резултата (хипотезе)

Квалитет говора синтетизованог предложеним методама биће тестиран објективним и субјективним мерама. Објективне мере су базиране на разлици обележја издвојених у оригиналним реченицама и обележја добијених у току синтезе, и обухватаће између осталог и мел-кепстралну дисторзију, средњу квадратне грешку процене основне учестаности и друге. Све мере су стандарно кориштене у литератури. Субјективна процена квалитета биће заснована на MOS (енгл. *Mean Opinion Score*) и MUSHRA (енгл. *Multiple Stimuli Hidden Reference and Anchor*) тестовима.

Очекује се да предложене методе покажу значајно побољшање у поређењу са конвенционалним, у погледу квалитета и стабилности синтетизованог говора.

Очекивани резултати представљају значајан научни допринос?

ДА

НЕ

ДЕЛИМИЧНО

V.6 план рада

Кандидат је током рада на пројекту МПНТР под називом „Развој дијалогских система за српски и друге јужнословенске језике“ реализовао значајан део истраживања. Целокупан рад би се могао поделити на две главне фазе:

- развој метода за синтезу говора. Објективна и субјективна евалуација говора генерисаног предложеним методама

- развој поступка за адаптацију на новог говорника. Објективна и субјективна евалуација говора генерисаног предложеним методама.

Оријентациони садржај докторске дисертације је следећи:

1. Уводна разматрања
2. Синтеза говора коришћењем скривених Марковљевих модела, дубоких неуронских мрежа и њихово поређење
3. Преглед постојећих техника адаптације на говорника
4. Опис нове технике адаптације на говорника
5. Експериментални резултати
6. Закључак

Крај истраживања, закључно са одбраном дисертације, планиран је за четврти квартал 2019. године.

План рада је одговарајући?

ДА

НЕ

ДЕЛИМИЧНО

V.7 метод и узорак истраживања

Базе које су коришћене у току истраживања биле су доступне захваљујући сарадњи са предузећем Алфанум д.о.о. из Новог Сада. Ради се о базама за енглески језик које поред неутралног говора садрже и по 15-ак минута говора снимљеног у неколико различитих говорних стилова. Поред тога, коришћене су и технике вештачког повећавања базе (енгл. *data augmentation*).

Метод и узорак су одговарајући?

ДА

НЕ

ДЕЛИМИЧНО

V.8 места, лабораторије и опреме за експериментални рад

Део истраживања који укључује имплементацију предложених приступа и њихову субјективну евалуацију спроведен је у Лабораторији за акустику и говорне технологије, на Катедри за телекомуникације и обраду сигнала, на Факултету техничких наука у Новом Саду. Такође, за истраживање је од великог значаја сарадња са предузећем АлфаНум д.о.о. из Новог Сада.

Услови за експериментални рад су одговарајући? ДА НЕ ДЕЛИМИЧНО

V.9 методе статистичке обраде података и осталих релевантних података

Све обуке модела и тестирања су спроведена у оквиру Лабораторије за акустику и говорне технологије, на Катедри за телекомуникације и обраду сигнала, односно у просторијама предузећа АлфаНум и удаљеним серверима доступним одатле.

Све технике и процедуре рађене су уз помоћ јавно доступних алата, као и алата развијених на Катедри и у предузећу АлфаНум. За обуку модела користи се *TensorFlow* окружење за обуку неуронских мрежа.

Предложене методе су одговарајуће? ДА НЕ ДЕЛИМИЧНО

VI ЗАКЉУЧАК

кандидат је подобан	<u>ДА</u>	НЕ	ДЕЛИМИЧНО
ментор је подобан	<u>ДА</u>	НЕ	ДЕЛИМИЧНО
тема је подобна	<u>ДА</u>	НЕ	ДЕЛИМИЧНО

Образложење:

На основу увида у до сада објављене радове, формулацију теме, предмета и циља истраживања, показаног познавања проблематике и очекиваних резултата, Комисија сматра да је тема **„Нова метода адаптације на говорника у параметарској синтези говора“** подобна за докторску дисертацију и да кандидат **Дарко Пекар** испуњава све законске услове за рад на предложеној теми докторске дисертације. Комисија предлаже Наставно-научном већу Факултета техничких наука и Сенату Универзитета у Новом Саду да прихвати као подобну предложену тему и кандидата и да се за менторе именују **проф. др Владо Делић** и **др Марко Јанев**, који испуњавају актуелне услове за вођење докторске дисертације кандидата према Правилнику УНС, и према Правилима за акредитацију у високошколском образовању.

датум: 08.07.2019.

Др Милан Сечујски, ванр. проф.,
председник комисије

Др Татјана Грбић, ред. проф., члан

Др Јернеја Жганец-Грос, ред. проф., члан

Др Никша Јаковљевић, доцент, члан

Др Бранислав Поповић, науч. сарад., члан