



**Vilniaus
universitetas**

Doktorantė: Justina Ramonaitė
Doktorantės vadovė: Dr. Gražina Korvel
Mokslo kryptis: INFORMATIKA N009
Studijų laikotarpis: 2023-2027
Studijų metai: 2024/2025

Giliuoju mokymusi pagrįstas šnekos signalo gerinimas



2024/2025 II semestro veikla

Visų studijų plano ir jo vykdymo suvestinė

Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2023/2024)	2	2
II (2024/2025)	2	2
III (2025/2026)		
IV (2026/2027)		
Iš viso:	4	4

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citav.rodikliu			Be citav. rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2023/2024)				3						
II (2024/2025)		1	1						1	Publikuota
III (2025/2026)	1							1		
IV (2026/2027)	1				2					
Iš viso:	2	0	1	3	2	0		1	1	

Ataskaitinio pusmečio darbo planas ir jo įvykdymas (1)

Egzaminai 2024/2025 (II pusmetis)		
Planas	Įvykdyta	Būklė
Gilieji neuroniniai tinklai (2025-09-22)	Gilieji neuroniniai tinklai Data: 2025-09-22	Išlaikytas

Bendrieji gebėjimai 2024/2025 (II pusmetis)	
Veikla	Kreditai (ECTS)
Europos ir Lietuvos HPC skaičiavimo resursai mokslui 2025-05-05, 2025-05-12, 2025-05-19	1

Ataskaitinio pusmečio darbo planas ir jo įvykdymas (2)

Dalyvavimas konferencijose 2024/2025 (II pusmetis)		
Planas	Įvykdyta	Konferencijos tipas
2025 IEEE 12th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), 2025-05-15 – 2025-05-17, Vilnius, Lietuva	Autoriai: Justina Ramonaitė, Faustas Žiliajevas, Gražina Korvel, Gintautas Tamulevičius Pranešimo pavadinimas: „LOUD – Labeled nOise aUdio Dataset“ Konferencijos pavadinimas: 2025 IEEE 12th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE) Data: 2025-05-15 – 2025-05-17 Vieta: Vilnius, Lietuva	Tarptautinė

Ataskaitinio pusmečio darbo planas ir jo įvykdymas (3)

Publikacijos 2024/2025 (II pusmetis)			
Planas	Įvykdyta	Būklė	Publikacijos tipas
Data Science in Applications, Towards AI-Driven Approaches	Ramonaitė, J., Gore, P., Korvel, G., Tamulevičius, G. (2025). Selection of Preprocessing Parameters for Wave-U-Net-Based Speech-Denoising. In: Dzemyda, G., Bernatavičienė, J., Kacprzyk, J. (eds) Data Science in Applications. Studies in Computational Intelligence, vol 1206. Springer, Cham. https://doi.org/10.1007/978-3-031-88486-3_5	Publikuota	Be cituojamumo rodiklio

Visų mokslinių tyrimų ir disertacijos rengimo etapai (1)

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
2.	Mokslinio tyrimo vykdymas:			
2.1.	Tyrimo metodikos sudarymas:			
2.1.1.	Tyrimo metodikos iškeltiems uždaviniams spręsti parinkimas.	2024 m. IV ketvirtis	2024 m. IV ketvirtis	Anksčiau pasirinkta bendresnė generatyvinių tinklų kryptis patikslinta į generatyvinius adversarinius tinklus.
2.1.2.	Teorinio ir empirinio tyrimų suplanavimas pagal pasirinktą metodiką.	2025 m. I ketvirtis	2025 m. III ketvirtis	Suformuluota hipotezė dėl SEGAN modelio tinkamumo užduočiai. Toliau planuojamas duomenų rinkinio paruošimas, modelio pritaikymas bei rezultatų analizė pagal objektyvius ir subjektyvius vertinimo rodiklius.

Visų mokslinių tyrimų ir disertacijos rengimo etapai (2)

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
2.	Mokslinio tyrimo vykdymas:			
2.1.	Tyrimo metodikos sudarymas:			
...	...	...	...	...
2.2.	Teorinis tyrimas:			
2.2.1.	Mokymosi metodų, naudojamų triukšmingų šnekos signalų gerinimui, tyrimas.	2025 m. I ketvirtis	2025 m. II ketvirtis	Susipažinta su skirtingais generatyviniais adversariniais tinklais, vienas jų (SEGAN) detaliai išanalizuotas ir atkurtas.
2.2.2.	Triukšmingų šnekos signalų gerinimui naudojamo mokymosi metodo sukūrimas/modifikavimas ir/ar testavimas.	2025 m. III ketvirtis	2026 m. III ketvirtis	

Tyrimo objektas, tikslas ir uždaviniai

Tyrimo objektas:

- triukšmo signalai, jų požymiai bei algoritmai jiems klasifikuoti (triukšmo profiliavimas),
- užtriukšminti šnekos signalai, giliojo mokymosi algoritmai, skirti triukšmo šalinimui iš šnekos signalo.

Tikslas:

- sukurti giliuoju mokymusi pagrįstą šnekos signalo gerinimo sistemą, kuri padidintų kalbos signalo suprantamumą ir kokybę esant triukšmo trikdžiams.

Tyrimo uždaviniai:

- Atlikti naujausios mokslinės literatūros, susijusios su užtriukšmintų šnekos signalų gerinimu, triukšmo atskyrimu/šalinimu, apžvalgą;
- Palyginti ir išanalizuoti esamus giliojo mokymosi metodus, taikomus šnekos signalų kokybės ir suprantamumo rodiklių didinimui;
- Išanalizuoti realaus pasaulio triukšmus ir jų sąveiką su šnekos signalais. Sukurti prie triukšmo prisitaikančią sistemą, skirtą šnekos signalo kokybei ir suprantamumui gerinti, kai signalas yra paveiktas triukšmų;
- Eksperimentiškai palyginti sukurta šnekos signalo gerinimo sistemą su esamais metodais.

2024/2025 II semestro rezultatai

SEGAN

Atkurtas pirmasis generatyvinis adversarinis tinklas, skirtas šnekos signalo kokybės gerinimui. SEGAN susideda iš generatoriaus ir diskriminatoriaus, nuostoliui apskaičiuoti naudojama funkcija iš LSGAN tinklo. Modelio įėjimui naudojami laiko srities signalai (angl. waveforms), todėl papildomų požymių išskaičiuoti nereikia.

Gauti rezultatai nagrinėti trimis pjūviais – pradinio SNR, pridėtinio triukšmo ir kalbėtojo.

Įvertinimo metrika

Suvokiamas kalbos kokybės įvertinimas (angl. perceptual evaluation of speech quality, PESQ), kurio vertės kinta nuo -0,5 iki 4,5. Kuo aukštesnė vertė tuo geriau.

Testavimo rinkinio PESQ yra **1,97**, originaliame SEGAN straipsnyje pasiektas **2,16**. Atkurtas SEGAN modelis pasiekė **2,20**.

Palyginimas, kai skiriasi pradinis SNR

SNR	min(PESQ)	max(PESQ)	avg(PESQ)	avg(PESQ change)
2,5	1,071886659	3,021648169	1,741764985	0,324379465
7,5	1,219521165	3,571138382	2,092597373	0,333097308
12,5	1,284722328	3,523057222	2,299979235	0,200712332
17,5	1,263553381	3,599812269	2,663653034	0,061821666

PESQ change	2.5	7.5	12.5	17.5
Negative	10	28	62	88
Positive	197	178	145	116
<i>Total</i>	<i>207</i>	<i>206</i>	<i>207</i>	<i>204</i>

Išvados (SNR)

- Geriausias PESQ pagerėjimas gautas, kai pradinis SNR mažesnis. Vidutinis PESQ pokytis mažiausias, kai SNR vertė didesnė, negu buvo apmokymo rinkinyje.
- Daugiausia atvejų, kai modelis pablogina signalą taip pat pasitaiko, kai SNR vertė didesnė – 12,5 dB ir ypač 17,5 dB.

Palyginimas, kai skiriasi pridėtas triukšmas

Noise	min(PESQ)	max(PESQ)	avg(PESQ)	avg(PESQ change)
bus	1,264960289	3,419785738	2,514415132	0,031350261
cafe	1,071886659	3,1795578	1,785181027	0,296864088
living	1,100660205	3,404481411	1,970227751	0,361699361
office	1,335393429	3,599812269	2,642010493	0,112794917
psquare	1,163126111	3,139425278	2,082048395	0,346723217

PESQ change	bus	cafe	living	office	psquare
Negative	84	14	8	68	14
Positive	80	150	158	96	152
<i>Total</i>	<i>164</i>	<i>164</i>	<i>166</i>	<i>164</i>	<i>166</i>

Išvados (triukšmai)

- Mažiausias vidutinis PESQ pagerėjimas matomas, kai triukšmas yra autobuso arba biuro erdvės.
- Kai signale yra autobuso arba biuro erdvės triukšmo, modelis pablogino signalo kokybę. Ypač tai matosi esant autobuso triukšmui, PESQ vertė sumažėjo daugiau nei pusei įrašų.

Palyginimas, kai skiriasi kalbėtojas

Speaker	min(PESQ)	max(PESQ)	avg(PESQ)	avg(PESQ change)
Female	1,071886659	3,555348873	2,081047905	0,289742537
Male	1,231078506	3,599812269	2,326131512	0,165507974

PESQ change	Female	Male
Negative	79	109
Positive	352	284
<i>Total</i>	<i>431</i>	<i>393</i>

Išvados (kalbėtojai)

- Vidutiniškai PESQ vertė padidėjo daugiau su moterimi kalbėtoja.
- PESQ vertė sumažėjo didesniam kiekiui įrašų, kai kalbėtojas vyras.

Skirtingas įvesties triukšmas z

z	avg(PESQ)
random	2,20
zero	2,18
none	2,12

Išvados (z)

- Atsitiktinio z įtraukimas teigiamai prisideda prie modelio veikimo. Kai z reikšmės lygios nuliui, rezultatai šiek tiek prastesni, o kai jis visiškai išimamas jie krenta dar labiau.

2025/2026 I semestro planas

Suplanuoti darbai

- Išplėsti eksperimentus, įtraukiant daugiau triukšmo tipų ir platesnį signalo ir triukšmo santykio spektrą.
- Patobulinti SEGAN architektūrą įtraukiant triukšmo profiliavimą.
- Įvertinti modelio veikimą pagal subjektyvius klausos testus, įtraukiant klausytojų vertinimus.



**Vilniaus
universitetas**

Klausimai