



**Vilniaus
universitetas**

Doktorantė: Justina Ramonaitė
Doktorantės vadovė: Dr. Gražina Korvel
Mokslo kryptis: INFORMATIKA N009
Studijų laikotarpis: 2023-2027
Studijų metai: 2025/2026

Giliuoju mokymusi pagrįstas šnekos signalo gerinimas



2025/2026 I semestro veikla

Visų studijų plano ir jo vykdymo suvestinė

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

Bendrieji gebėjimai 2025/2026 (I pusmetis)	
Veikla	Kreditai (ECTS) [2/3]
16th International Baltic Conference on Digital Business and Intelligent Systems, 2024 06 30 – 2024 07 03, Vilnius, Lietuva	1
Europos ir Lietuvos HPC skaičiavimo resursai mokslui 2025-05-05, 2025-05-12, 2025-05-19	1

Visų studijų plano ir jo vykdymo suvestinė

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	Publikuota	1		
IV (2026/2027)	1				2					
Iš viso:	2	1	1	3	2	1		1	1	

Ataskaitinio pusmečio darbo planas ir jo įvykdymas

Publikacijos 2025/2026 (I pusmetis)			
Planas	Įvykdyta	Būklė	Publikacijos tipas
IEEE Access	J. Ramonaitė, G. Korvel, G. Tamulevičius, „Generative Adversarial Networks in Speech Enhancement: A Survey,“ IEEE Access, vol. 14, pp. 29048-29071, 2026, doi: 10.1109/ACCESS.2026.3667063	Publikuota	Turi citavimo rodiklį (2024 JCR – JIF 3,6)

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
2.	Mokslinio tyrimo vykdymas:			
...	...	...	...	...
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	Atkurtas CMGAN tinklas, numatomas tolesnis modelio tobulinimas įtraukiant triukšmo profiliavimą.
2.3.	Empirinis tyrimas:			
2.3.1.	Sudarytų metodų pritaikymas praktinių uždavinių sprendimui.	2026 m. II ketvirtis	2027 m. I 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 prisitaikantią sistemą, skirtą šnekos signalo kokybei ir suprantamumui gerinti, kai signalas yra paveiktas triukšmų;
- Eksperimentiškai palyginti sukurta šnekos signalo gerinimo sistemą su esamais metodais.

2025/2026 I semestro rezultatai

CMGAN

Išnagrinėti rezultatai, gauti panaudojus generatyvų adversarinį tinklą CMGAN, kuris skirtas šnekos signalo kokybės gerinimui. CMGAN yra konvoliucija augmentuotu transformeriu (konformeriu) ir metrika paremtas generatyvinis adversarinis tinklas (angl. conformer-based metric generative adversarial network).

Gauti rezultatai palyginti su SEGAN rezultatais.

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.

Standartinio testavimo rinkinio PESQ yra **1,97**, originaliame SEGAN straipsnyje pasiektas **2,16**, o CMGAN **3,41**.

Palyginimas, kai skiriasi pradinis SNR

Average PESQ change

SNR	SEGAN	CMGAN
2,5	0,324379465	1,463851964
7,5	0,333097308	1,540058772
12,5	0,200712332	1,463843182
17,5	0,061821666	1,187864976

Number of records with negative PESQ change

SNR	2,5	7,5	12,5	17,5
SEGAN	10	28	62	88
CMGAN	0	0	0	3
Total	207	206	207	204

Palyginimas, kai skiriasi pridėtas triukšmas

Average PESQ change

Noise	SEGAN	CMGAN
<i>bus</i>	0,031350261	1,177744077
<i>cafe</i>	0,296864088	1,393534018
<i>living</i>	0,361699361	1,652433069
<i>office</i>	0,112794917	1,343315051
<i>psquare</i>	0,346723217	1,501881276

Number of records with negative PESQ change

Noise	bus	cafe	living	office	psquare
SEGAN	84	14	8	68	14
CMGAN	3	0	0	0	0
<i>Total</i>	<i>164</i>	<i>164</i>	<i>166</i>	<i>164</i>	<i>166</i>

Palyginimas, kai skiriasi kalbėtojas

Average PESQ change

Speaker	SEGAN	CMGAN
<i>Female</i>	0,289742537	1,50947326
<i>Male</i>	0,165507974	1,310499955

Number of records with negative PESQ change

Speaker	Female	Male
SEGAN	79	109
CMGAN	3	0
<i>Total</i>	431	393

Išvados

- Abiem atvejais PESQ mažiausiai pagerėja su 17,5 dB pradiniu SNR, daugiausiai su 7,5.
- Abiem atvejais PESQ mažiausiai pagerėja esant autobuso triukšmui, daugiausiai esant svetainės.
- PESQ mažiau pagerėja su vyru kalbėtoju, tačiau, kai naudojamas CMGAN, PESQ pablogėjimo atvejai pasitaikė tik su moterimi kalbėtoja. Naudojant SEGAN modelį tokių atvejų buvo daugiau su vyru.

Rezultatai su kitais duomenimis

SNR	file count	average PESQ change
2,5	96	1,258320436
7,5	96	1,634670385
12,5	94	1,835974956
17,5	93	1,871889849

Noise	file count	average PESQ change
Highway	126	1,689031858
Kitchen_hood1	126	1,968559965
Rain	127	1,287700401

2025/2026 II semestro planas

Suplanuoti darbai

- Pasirinkti triukšmo informaciją ar požymius, kuriuos būtų galima prijungti prie modelio vienu iš būdų (kaštų funkcija, papildomas modelis, architektūros papildymas atskira šaka, skirta apdoroti triukšmui ar pan.)



**Vilniaus
universitetas**

Klausimai