



**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

II 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

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

Ataskaitinio pusmečio darbo planas ir jo įvykdymas

Dalyvavimas konferencijose 2025/2026 (II pusmetis)		
Planas	Įvykdyta	Konferencijos tipas
2026 Seventh International Conference on Intelligent Data Science Technologies and Applications (IDSTA2026) Data 2026-09-22 – 2026-09-25 Vieta: Bukareštas, Rumunija	Autoriai: Justina Ramonaitė, Gražina Korvel Pranešimo pavadinimas: „Improving Speech Enhancement Performance of CMGAN Using SI-SNR-Based Loss“ Konferencijos pavadinimas: 2026 Seventh International Conference on Intelligent Data Science Technologies and Applications (IDSTA2026) Data: 2026-09-22 – 2026-09-25 Vieta: Bukareštas, Rumunija	Tarptautinė

Tarptautiniai renginiai

Dalyvavimas tarptautinėse konferencijose	
	Aprašas
1.	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
2.	Autoriai: Justina Ramonaitė, Gražina Korvel Pranešimo pavadinimas: „Improving Speech Enhancement Performance of CMGAN Using SI-SNR-Based Loss“ Konferencijos pavadinimas: 2026 Seventh International Conference on Intelligent Data Science Technologies and Applications (IDSTA2026) Data 2026-09-22 – 2026-09-25 Vieta: Bukareštas, Rumunija

Publikacijos

Publikacijos (tik su citavimo rodikliu)		
	Bibliografinis aprašas	Būklė
1.	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 (2024 JCR – JIF 3,6)

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
2.2.	Teorinis tyrimas:			
...	...	...	...	...
2.2.2.	Triukšmingų šnekos signalų gerinimui naudojamą mokymosi metodo sukūrimas/modifikavimas ir/ar testavimas.	2025 m. III ketvirtis	2026 m. III ketvirtis	Atkurtas CMGAN tinklas. Pasiūlyta papildyta CMGAN nuostolių funkcija įtraukiant triukšmo lygio informaciją. Tęsiami analogiški eksperimentai su SEGAN bei įtraukta triukšmo tipo bei pradinio SNR informacija.
2.3.	Empirinis tyrimas:			
2.3.1.	Sudarytų metodų pritaikymas praktinių uždavinių sprendimui.	2026 m. II ketvirtis	2027 m. I ketvirtis	Planuojamas gauto nutriukšminto signalo tinkamumo tolimesnėms užduotims (pvz. kalbos atpažinimo) tikrinimas.

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 II semestro rezultatai

Triukšmo įtraukimas apmokant tinklą

Tikslas – pasitelkti triukšmo informaciją gerinant SE modelių rezultatus.

- Noise-aware mokymas – kai naudojama triukšmo informacija, tačiau pats modelis nekinta.
- Noise-adaptive mokymas – priklausomai nuo triukšmo kažkas pakinta modelyje ar jo apmokymo procese.

SI-SNR pridėjimas

Generatoriaus nuostolių funkcija papildyta SI-SNR, t.y. į mokymosi procesą įtraukiama informacija apie triukšmo lygį signalė. Tikrinami skirtingi šio elemento svoriai. Eksperimentai atliekami su CMGAN ir SEGAN.

$$L_{G_new} = L_{G_old} + \gamma L_{SI-SNR}$$

$$L_{SI-SNR} = -10 \log_{10} \left(\frac{\|s_{target}\|^2}{\|e_{noise}\|^2} \right)$$

$$s_{target} = \frac{(\hat{s}^T s)}{\|s\|^2} s$$

$$e_{noise} = \hat{s} - s_{target}$$

s – švarus signalas, $\hat{s}$ – nutirukšmintas signalas.

Rezultatai

CMGAN					
y	PESQ	STOI	CSIG	CBAK	COVL
initial	1.97	0.92	3.35	2.44	2.63
none	2.97	0.95	4.23	3.08	3.61
0.001	3.16	0.95	4.39	3.19	3.79
0.005	3.26	0.96	4.47	3.27	3.90
0.01	3.24	0.96	4.48	3.28	3.89
0.05	3.20	0.96	4.39	3.26	3.82
0.1	3.14	0.95	4.40	3.26	3.80

SEGAN					
y	PESQ	STOI	CSIG	CBAK	COVL
initial	1.97	0.92	3.35	2.44	2.63
none	2.19	0.93	3.53	3.01	2.84
0.1	2.31	0.93	3.63	3.09	2.95
0.5	2.42	0.93	3.74	3.14	3.06
1	2.27	0.93	3.66	2.48	2.95

Pritaikymas pagal pradinį triukšmą

γ – priklauso nuo pradinio triukšmo lygio ir tipo, t.y. $\gamma(x, y) = \alpha f(x)g(y)$, kur x yra pradinis SNR, o y – triukšmo tipas.

$$f(x) = a + \frac{x - x_{min}}{x_{max} - x_{min}}(b - a),$$

$$\gamma(x, y) = \alpha \begin{cases} f(x), & \text{if } y \text{ is } car; \\ cf(x), & \text{otherwise.} \end{cases}$$

Čia $[a, b]$ naujas intervalas SNR reikšmėms, α yra papildomas daugiklis, naudojamas komponento pritaikymui, c yra daugiklis, skirtas SI-SNR komponento stiprinimui pagal triukšmo tipą.

Triukšmų intensyvumas ir spektrinės savybės

Set	Type	Average of Intensity	STD of Intensity	Average of SC	Average of SBW	Average of SF	Average of SR
Test	DLIVING	38.04	1.01	7242	7846	0.02754	18083
	OOFFICE	48.67	0.58	5379	7373	0.00478	15544
	SCAFE	53.69	2.41	3508	5138	0.00758	8536
	SPSQUARE	54.80	2.85	3121	5380	0.00327	7510
	TBUS	68.78	6.47	2221	4257	0.00114	4928
Train	DKITCHEN	52.19	0.68	8306	5706	0.04549	14177
	OMEETING	38.72	2.71	7376	7670	0.04021	17714
	PCAFETER	49.58	2.30	4159	5678	0.01182	10072
	PRESTO	49.42	1.33	4486	5086	0.02313	9820
	PSTATION	57.27	0.37	2542	4180	0.00317	5562
	STRAFFIC	58.47	3.01	3398	4935	0.00497	8497
	TCAR	71.42	2.98	833	3275	0.00004	348
	TMETRO	64.73	2.05	2358	4031	0.00200	5257

Apmokymo rinkinys

Noise type	File count	PESQ
babble	1158	1.26
cafeteria	1156	1.33
car	1159	2.41
kitchen	1152	1.53
meeting	1155	1.38
metro	1159	1.56
restaurant	1157	1.24
ssn	1168	1.26
station	1156	1.28
traffic	1152	1.39

Rezultatai

SEGAN		
SNR interval	Type (when not car)	PESQ
0.1-1	-	2.37
0.1-0.5	2	2.36
0.1-0.25	4	2.44

CMGAN		
SNR interval	Type (when not car)	PESQ
0.1-1	-	3.20
0.1-0.25	4	3.33

2026/2027 I semestro planas

Suplanuoti darbai

- Užbaigti tyrimus su nuostolių funkcijos papildymu.
- Pereiti prie modelio pritaikymo praktiniams tikslams (pvz. ASR).



**Vilniaus
universitetas**

Klausimai