

Giliojo mokymo modeliai kalbos signalui gryninti

Parengė: Daniel Zakševski

Darbo vadovas: Doc. Dr. Gintautas Tamulevičius

Doktorantūros studijų metai: 2023-2027

Studijų metai: 4-ieji

Doktorantūros studijų ir mokslinių tyrimų plano bei jo vykdymo suvestinė

Studijų metai	Egzaminai	
	Planas	Ivykdyta
I (2023/2024)	2	2
II (2024/2025)	1	1
III (2025/2026)	1	1
IV (2026/2027)	0	0
Iš viso:	4	4

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

Ataskaitinio pusmečio darbo planas ir jo vykdymas, 2025/2026 – II pusmetis

Egzaminai, 2025/2026 – II pusmetis		
Planas	Ivykdyta	Būklė

Dalyvavimas konferencijose, 2025/2026 – II pusmetis		
Planas	Ivykdyta	Konferencijos tipas

Publikacijos, 2025/2026 – II pusmetis			
Planas	Ivykdyta	Būklė	Publikacijos tipas
Publikacija mokslo leidinyje, turinčiame cituojamumo rodiklį			

Doktorantūros studijų pasiekimai, tarptautiniai renginiai ir publikacijos

Dalyvavimas tarptautinėse konferencijose	
	Aprašas
1.	D. Zakševski, G. Korvel, G. Tamulevičius, „A Decoupled Biophysically-Inspired Architecture for Speech Enhancement“, „6th International Conference on Communications, Information, Electronic and Energy Systems“, 2025 m., lapkričio 26 – 28 d., Rusės m., Bulgarija

Publikacijos su citavimo rodikliu	
	Bibliografinis aprašas

Mokslinių tyrimų ir disertacijos rengimo etapai (1)

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
1.	Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):			1.1., 1.2., 1.3. 1.4 įvykdyti
1.1.	Kalbos signalo gryninimo problematikos apžvalga	2023 m. IV ketvirtis	2024 m. I ketvirtis	
1.2.	Kalbos signalo gryninimo problemos klasikinių sprendimų metodikos analizė	2024 m. I ketvirtis	2024 m. I ketvirtis	
1.3.	Kalbos signalo gryninimo problemos mašininio mokymu grįstų sprendimų metodikos analizė	2024 m. I ketvirtis	2024 m. II ketvirtis	
1.4.	Išankstinio duomenų apdorojimo metodų analizė	2024 m. I ketvirtis	2024 m. III ketvirtis	
2.	Mokslinio tyrimo vykdymas:			2.1.1 įvykdytas
2.1.	Tyrimo metodikos sudarymas:			
2.1.1.	Tyrimų tikslų ir užduočių suformulavimas, tyrimo metodikos sudarymas	2024 m. II ketvirtis	2024 m. III ketvirtis	
2.2.	Teorinis tyrimas:			2.2.1., 2.2.2. įvykdyti
2.2.1.	Klasikinių bei mašininio mokymu grįstų metodų palyginimas kalbos signalo gryninimo uždavinio sprendimui	2024 m. III ketvirtis	2025 m. I ketvirtis	
2.2.2.	Giliojo mokymo metodais grįstų kalbos signalo gryninimo algoritmų apibendrinimas	2025 m. I ketvirtis	2025 m. III ketvirtis	

Mokslinių tyrimų ir disertacijos rengimo etapai (2)

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
2.3.	Empirinis tyrimas:			2.3.1. įvykdytas, 2.3.2 vykdomas
2.3.1.	Skirtingomis architektūromis paremtų giliojo mokymo algoritmų tyrimas	2025 m. III ketvirtis	2026 m. I ketvirtis	
2.3.2.	Giliojo mokymo metodais grįsto kalbos gryninimo algoritmų praktinio taikymo problematikos tyrimas	2026 m. I ketvirtis	2026 m. IV ketvirtis	
2.4.	Gautų duomenų analizė, apibendrinimas, išvadų parengimas:			
2.4.1.	Gautų duomenų analizė, apibendrinimas, išvadų parengimas	2026 m. IV ketvirtis	2027 m. III ketvirtis	
3.	Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų, ir kt.) parengimas:			
3.1.	Literatūros apžvalga	2026 m. IV ketvirtis	2027 m. I ketvirtis	
3.2.	Tyrimų metodikos aprašymas	2027 m. I ketvirtis	2027 m. II ketvirtis	
3.3.	Rezultatų aprašymas	2027 m. II ketvirtis	2027 m. III ketvirtis	
3.4.	Ginamų teiginių ir išvadų aprašymas	2027 m. III ketvirtis	2027 m. IV ketvirtis	
4.	Daktaro disertacijos parengimas ir svarstymas padalinyje	2027 m. IV ketvirtis	2027 m. IV ketvirtis	
5.	Daktaro disertacijos gynimas	2027 m. IV ketvirtis	2027 m. IV ketvirtis	

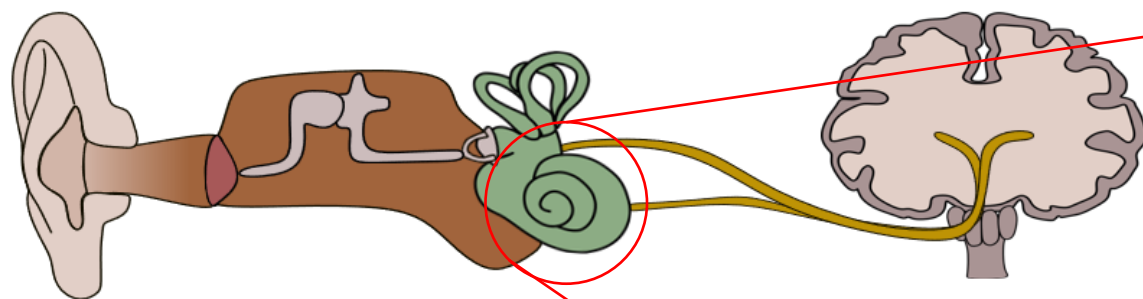
Tyrimo detalės

- **Tyrimo objektas:** šnekos signalas, balso signalo kokybė bei triukšmas.
- **Tyrimo tikslas:** pasiūlyti ir pritaikyti klausos modeliais grįstus giliojo mokymo tinklus kalbos signalui gryninti.
- **Tyrimo uždaviniai:**
 1. Įvertinti giliojo mokymo modelius naudojamus kalbos signalui gryninti.
 2. Apžvelgti ir įvertinti žmogiškosios klausos modelius.
 3. Integruoti klausos modelius, kitas žinias į kalbos signalo gryninimo modelį.
 4. Įvertinti pasiūlytųjų modelių modifikacijų efektyvumą, signalo gryninimo kokybę.

Moksliniai rezultatai, 2025/2026 – II pusmetis (1)

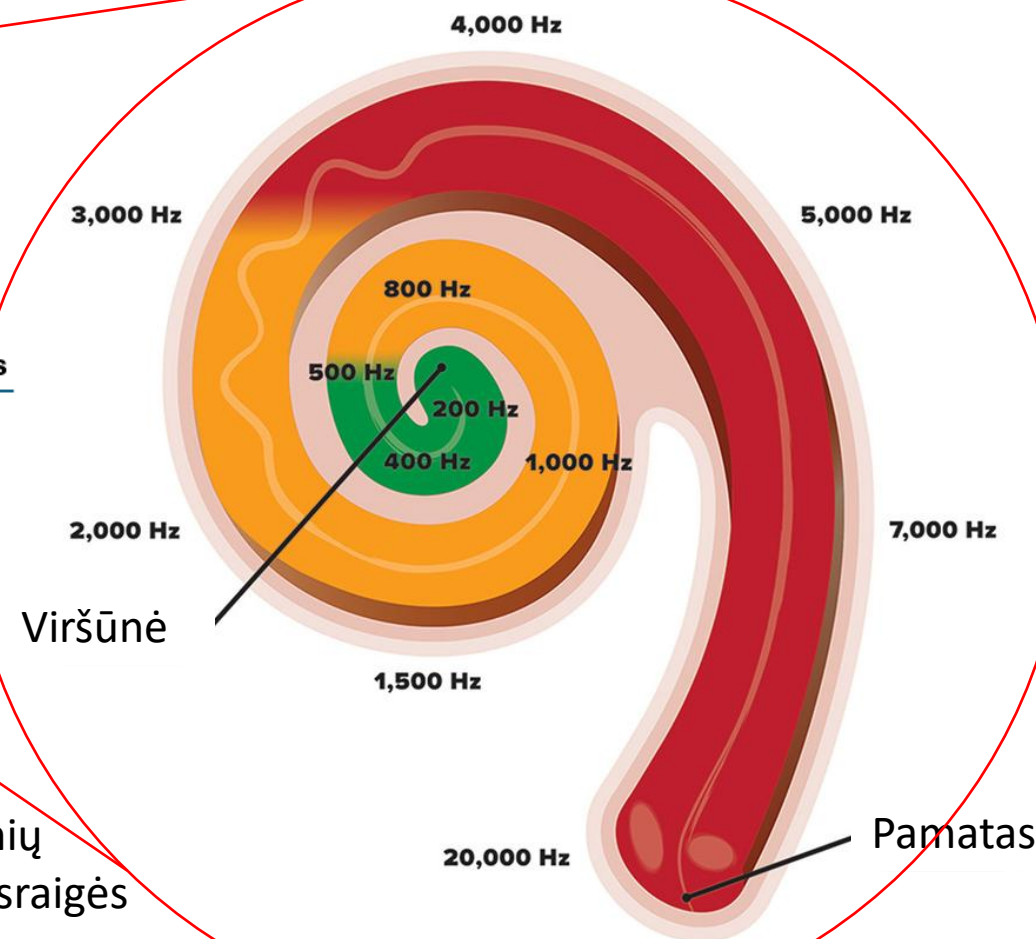
- Pagrindinis kalbos gryninimo (angl. speech enhancement) tikslas – padidinti akustinių kalbos signalų kokybę ir suprantamumą, pašalinant įvairius triukšmus bei jų sukeltus iškreipymus, atkuriant ar išryškinant spektrines kalbos signalo savybes.
- Geriausia šnekos suvokimo sistema yra žmogaus klausa. Savybių, leidžiančių žmogaus klausai efektyviai veikti, integravimas į skaitmeninius šnekos gryninimo algoritmus, potencialiai leistų toliau juos tobulinti.

Moksliniai rezultatai, 2025/2026 – II pusmetis (2)



	Išorinė ausis	Vidurinė ausis	Vidinė ausis	Smegenų kamienas ir smegenys
Funkcija:	Garso surinkimas	Vibracijų sustiprinimas	Dažnių analizė	Garso suvokimas ir interpretavimas

1 pav. Žmogaus klausos sistema, jos dalių funkcijos



2 pav. Sraigė ir jos tonotopinės savybės visame dažnių spektre. Aukšto dažnio garsai labiausiai stimuliuoja sraigės pamatinę dalį, o žemo dažnio garsai – viršūnę [1].

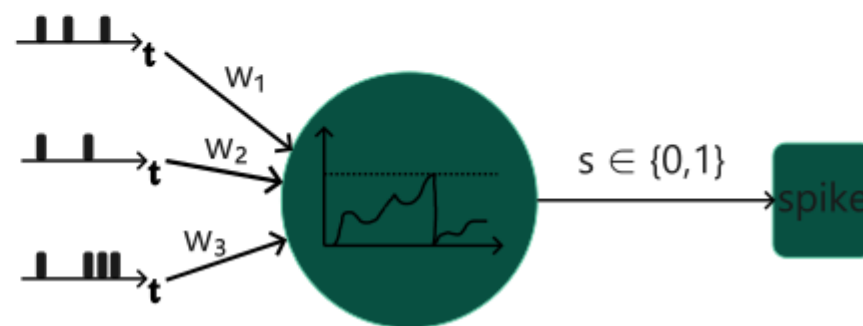
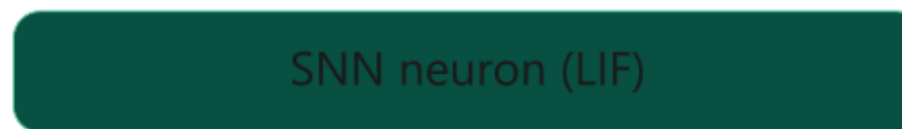
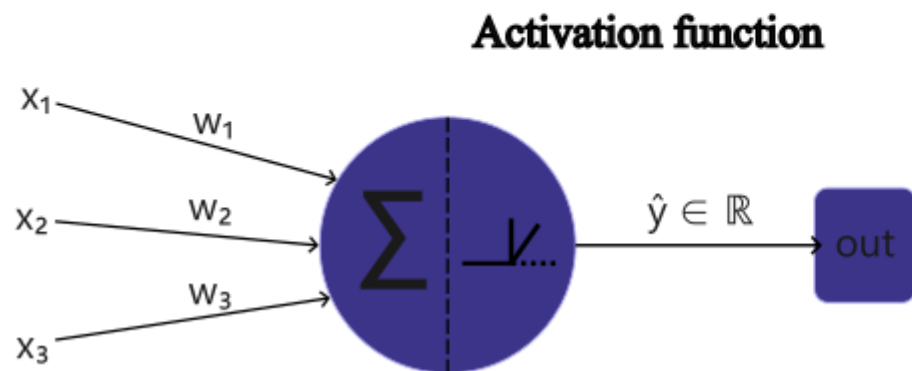
[1] Lahav, Amir & Skoe, Erika. (2014). An acoustic gap between the NICU and womb: a potential risk for compromised neuroplasticity of the auditory system in preterm infants. *Frontiers in Neuroscience*. 8. 381. 10.3389/fnins.2014.00381.

Moksliniai rezultatai, 2025/2026 – II pusmetis (3)

- Remiantis žmogaus klausos sistema, šnekos gryninimui naudojami impulsiniai neuroniniai tinklai (angl. Spiking Neural Network).
- Šie tinklai pasižymi įvykiais grįstu veikimu, vidinės atminties išlaikymu kiekviename neurone, vidine laikine dinamika, bei optimizuotu veikimu specializuotose įrenginiuose.
- Tačiau šie tinklai susiduria su elektrinių impulsų nediferencijuojamumu, ribota vieno impulso informacine verte, mažesniu tikslumu lyginant su standartiniais tinklais bei energiniu efektyvumu būtent tik specializuotoje įrangoje.

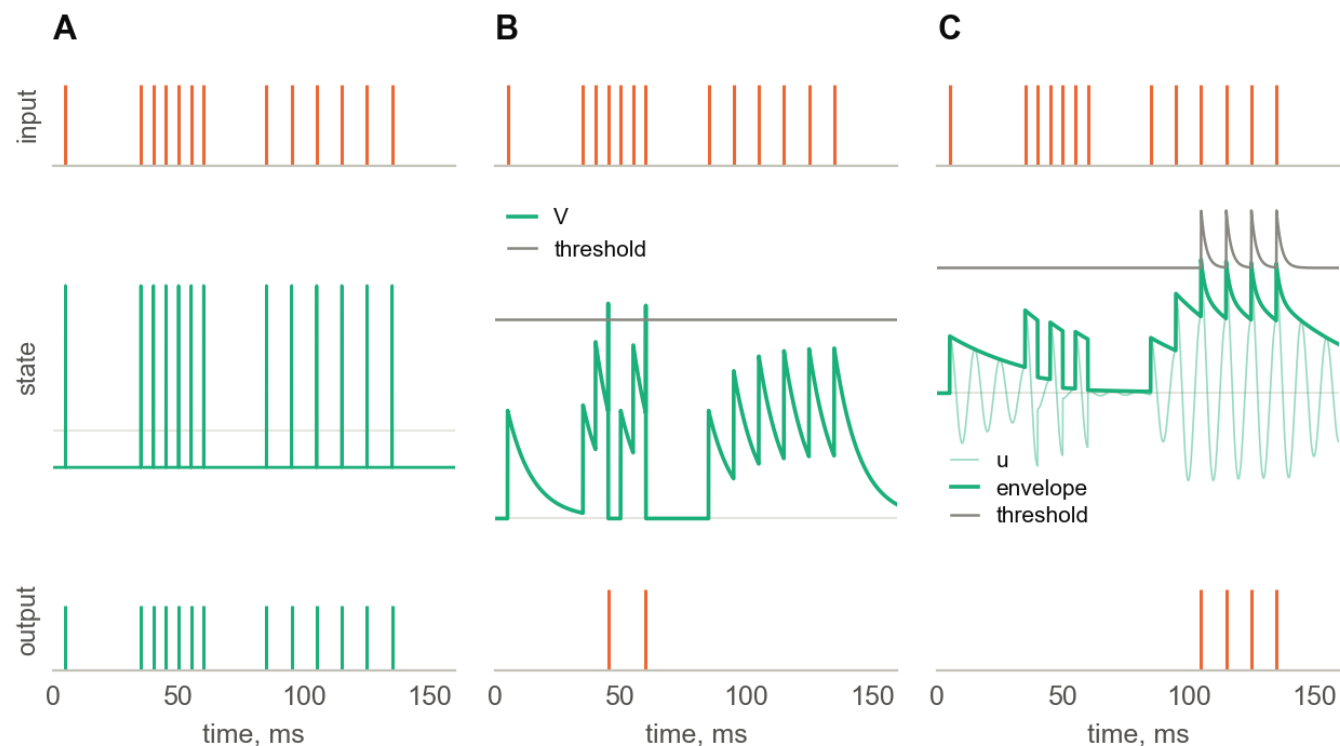
Moksliniai rezultatai, 2025/2026 – II pusmetis (4)

- Gauna įvestį, turi svorinius koeficientus, išvestį
- Gali būti jungiami į sluoksnius



3 pav. Dirbtinio ir impulsinio neuronų veikimo principo palyginimas

Moksliniai rezultatai, 2025/2026 – II pusmetis (5)



4 pav. Skirtingų neuronų atsakas į vienodą įvestį. **A** – dirbtinis, **B** – LIF, **C** – RF neuronai.

Algorithm 1 Per-sample update of a single resonate-and-fire neuron.

Require: carried state u, v, q ; input sample x

- 1: $r \leftarrow \rho e^{-cq}$ ▷ effective radius
- 2: $I \leftarrow g_{\text{in}}(1 - \rho)(W_{\text{in}} x)$ ▷ injected current
- 3: $u' \leftarrow r(\cos \theta u - \sin \theta v) + I$
- 4: $v' \leftarrow r(\sin \theta u + \cos \theta v)$
- 5: $\theta \leftarrow v_{\text{th}} + q$ ▷ Adaptive threshold
- 6: $o \leftarrow H(\sqrt{u'^2 + v'^2} + \varepsilon - \theta)$ ▷ spike
- 7: $q' \leftarrow \gamma q + o$ ▷ refractory update
- 8: **return** spike o ; next state u', v', q'

1 algoritmas. Šiame darbe naudojamo RF neurono algoritmas, vieno duomenų taško apdorojimo mechanizmas.

Moksliniai rezultatai, 2025/2026 – II pusmetis (6)

Darbas	Uždavinys	Išvestis
BRF [2]	Sekų klasifikavimas (skaitmenys, EKG, šnekėti skaitmenys)	Klasės žymė
S5-RF [3]	Balso komandų atpažinimas	Klasės žymė
PRF [4]	Ilgo konteksto modeliavimo uždaviniai	Klasės žymė
Ankstesni RF darbai	Įvesties kodavimas, vaizdų ir garso klasifikavimas	Klasės žymė arba požymiai
Šis darbas	Šnekos gryninimas	Skaitmeninis garso signalas

1 lentelė. Darbų, naudojančių RF neuronus, sprendžiami uždaviniai.

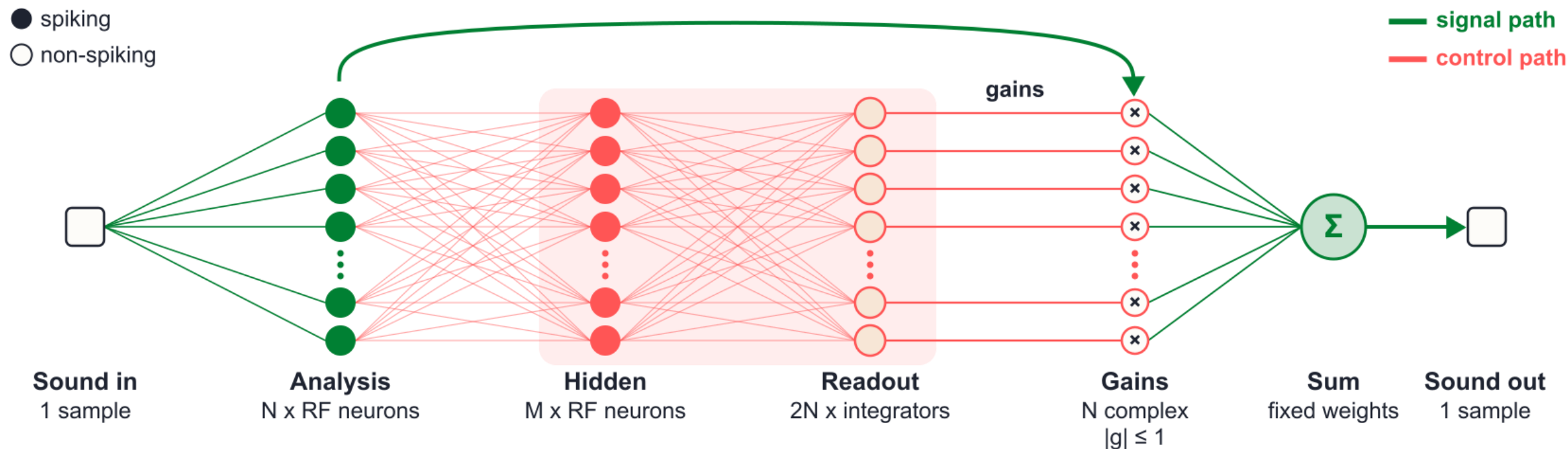
[2] Higuchi, S., Kairat, S., Bohté, S. M., & Otte, S. (2024). Balanced resonate-and-fire neurons. *Proceedings of the 41st International Conference on Machine Learning, ICML'24*.

[3] Huber, T. E., Lecomte, J., Polovnikov, B., & von Arnim, A. (2025). Scaling Up Resonate-and-Fire Networks for Fast Deep Learning. *Lecture Notes in Computer Science, 15646 LNCS*, 241–258. https://doi.org/10.1007/978-3-031-92460-6_15/FIGURES/3

[4] Huang, Y., Liu, Z., Feng, C., Lin, X., Ren, H., Fu, H., Zhou, Y., Xing, H., & Cheng, B. (2027). PRF: Parallel Resonate and Fire Neuron for Long Sequence Learning in Spiking Neural Networks. *Pattern Recognition, 183*, 114968.

<https://doi.org/10.1016/J.PATCOG.2026.114968>

Moksliniai rezultatai, 2025/2026 – II pusmetis (7)



5 pav. Šiame darbe siūlomos architektūros diagrama. Kiekviename sluoksnyje naudojami RF neuronai.

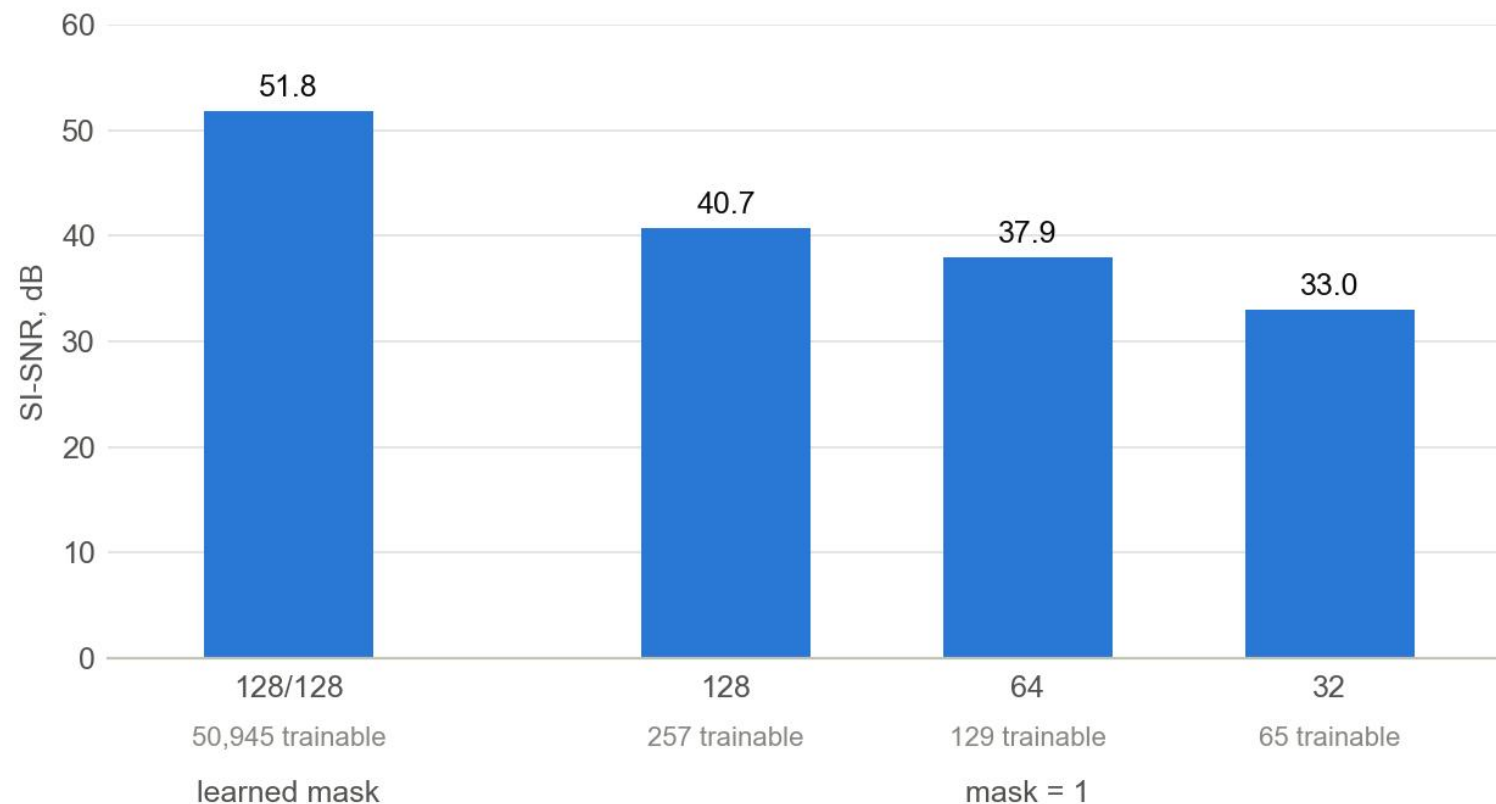
Toliau N/M žymės analizės ir paslėptąjo sluoksnių dydžius (neuronų kiekį)

Moksliniai rezultatai, 2025/2026 – II pusmetis (8)

- Tinklo esmė:
 - Į tinklą vienu metu įeina vienas duomenų taškas (angl. *sample*)
 - Nėra kaupimo, nėra žvalgos priekinė
 - Neuronai turi savo atmintį, kiekvienas neuronas atsakingas tik už savo būseną
 - Veikimo sąlygos ir sprendimai nesikeičia tarp treniravimo ir taikymo

Moksliniai rezultatai, 2025/2026 – II pusmetis (9)

- PESQ bei STOI tarp visų – vienodas.



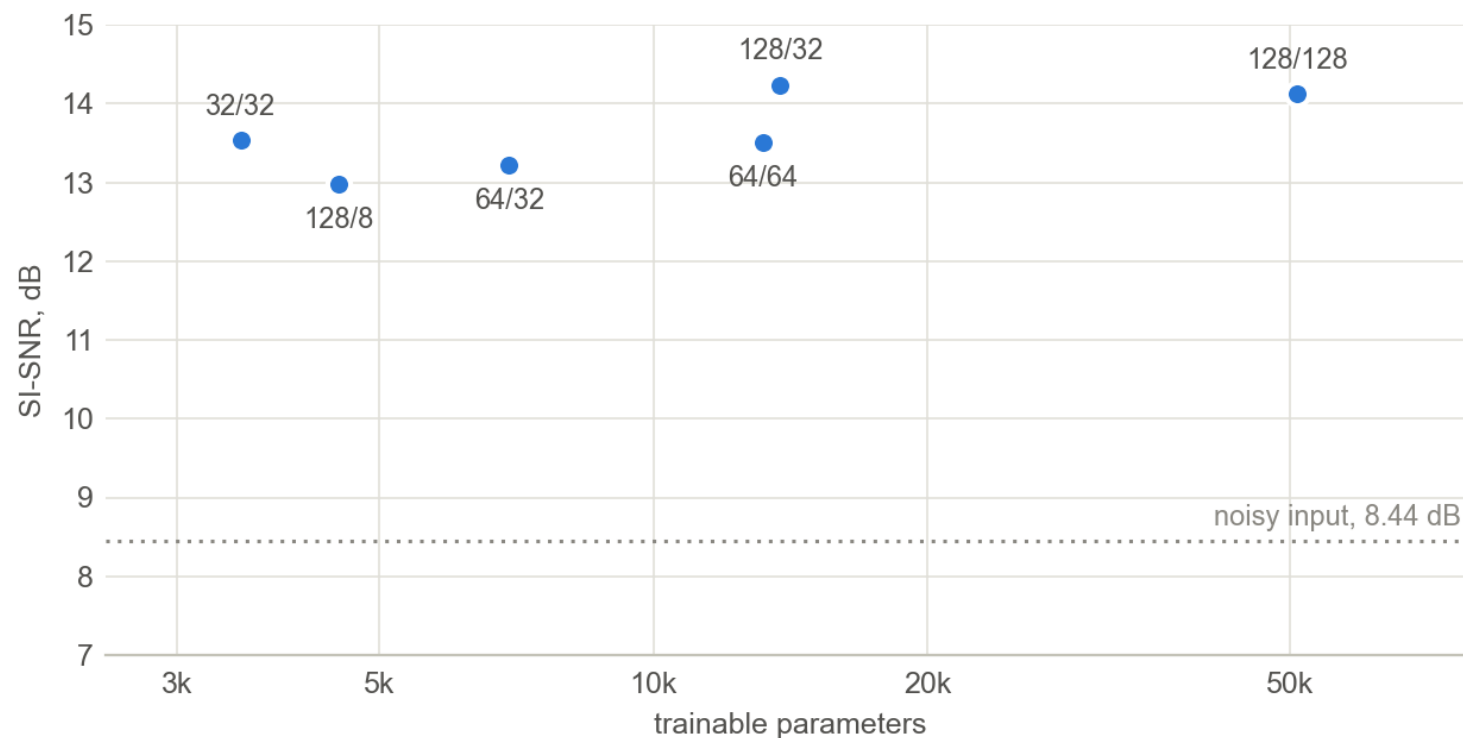
6 pav. Rekonstrukcijos uždavinio (švarus signalas kaip įvestis, švarus signalas kaip išvestis) rezultatai.

Moksliniai rezultatai, 2025/2026 – II pusmetis (10)

Įvertis	Ką matuoja?	Triukšmingas	Šis darbas	Pokytis
SI-SNR, dB	Signalų klaida	8.44	14.13	+5.69
PESQ	Šnekos kokybė	1.97	2.10	+0.13
STOI	Šnekos suprantamumas	0.921	0.911	-0.010
DNSMOS SIG	Signalų kokybė	3.32	3.18	-0.14
DNSMOS BAK	Triukšmo lygis	3.11	3.55	+0.44
DNSMOS OVRL	Bendra kokybė	2.68	2.74	+0.06

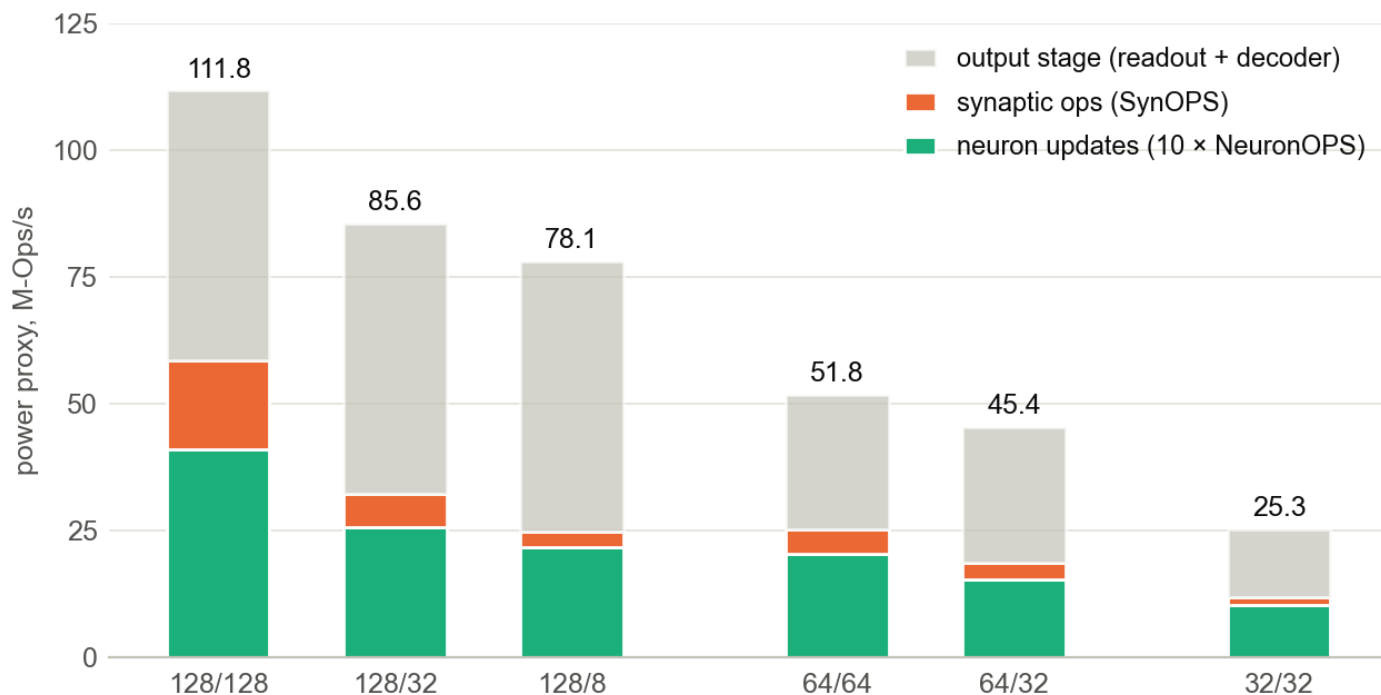
2 lentelė. 128/128 tinklo rezultatai, įverčiai bei jų paaiškinimas.

Moksliniai rezultatai, 2025/2026 – II pusmetis (11)



7 pav. SI-SNR pokytis priklausomai nuo tinklo dydžio/parametrų kiekio.

Moksliniai rezultatai, 2025/2026 – II pusmetis (12)



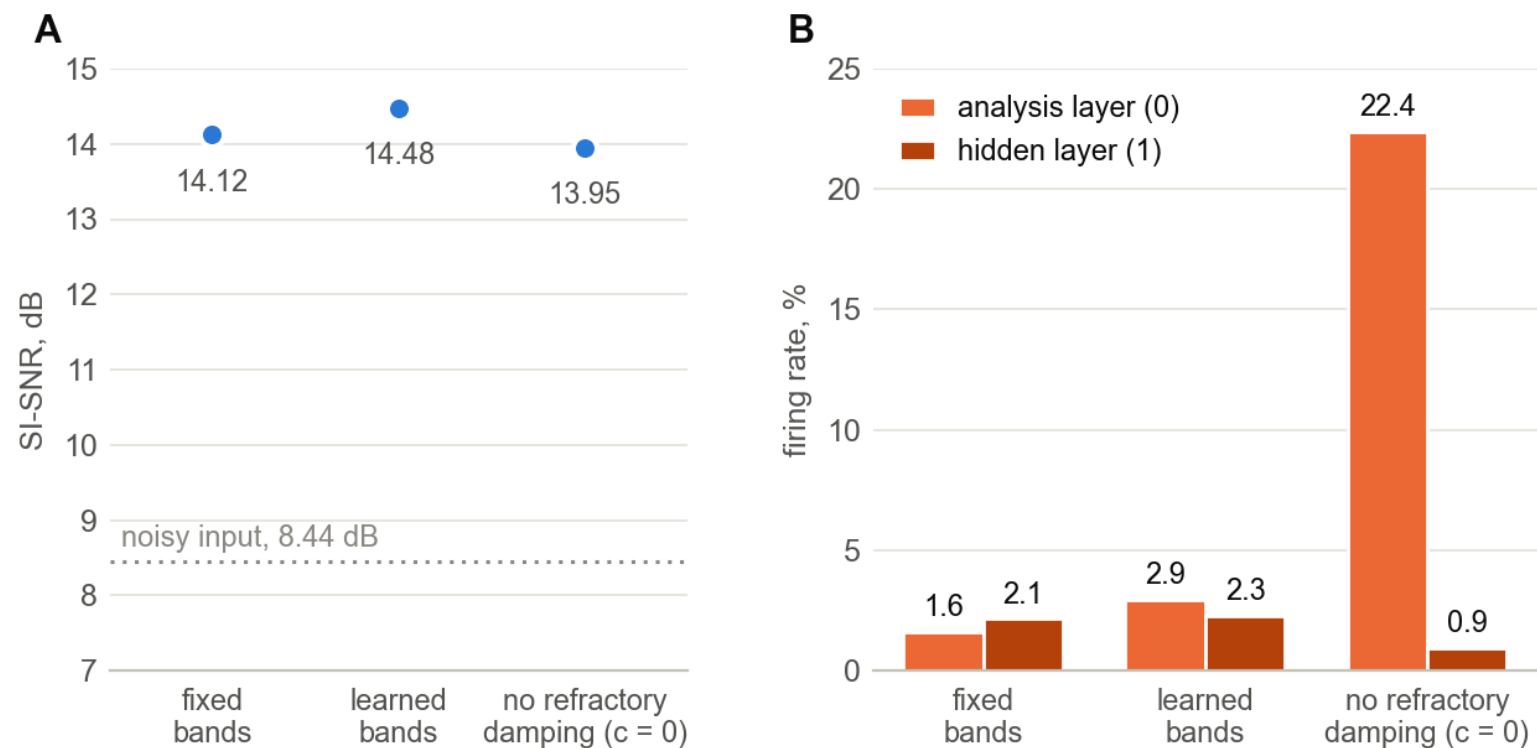
8 pav. Našumo, operacijų skaičius per sekundę, apibrėžtas pagal N-DNS, priklausomybė nuo tinklo dydžio/parametrų kiekio.

Power proxy = SynOPS + 10 x NeuronOPS,
kiekvienam tinklo žingsniui kiekvienai įvesčiai

Ryšys su FLOP'ais:

- SynOP: 1 FLOP'as per impulsą (sudėtis), 2 FLOP'ai per įvesties MAC.
- NeuronOP: šio darbo neurono atnaujinimą sudaro 23 FLOP'ai. x10 yra N-DNS įvestas daugiklis
- Išvestis: 32 FLOP'ai kiekvienam įvesties neuronui, 3 kiekvienam žingsniui.

Moksliniai rezultatai, 2025/2026 – II pusmetis (13)



9 pav. **A** – triukšmo šalinimo kokybė, esant fiksuotiems ir apmokomiems centriniams dažniams, bei išjungus slopinimą. **B** – abiejų sluoksnių santykinis impulsų kiekis, esant fiksuotiems ir apmokomiems centriniams dažniams, bei išjungus slopinimą. 128/128 architektūra.

Moksliniai rezultatai, 2025/2026 – II pusmetis (14)

Tinklas	Įvestis/įvesties sistema	Užlaikymas	Parametrų kiekis	PESQ	STOI	OVRL	SIG	BAK
Triukšmas		-	-	1.97	0.92	2.69	3.34	3.12
SDNN (Intel) [5]	STFT kadras	32 ms	525K	2.00	0.91	2.44	3.05	3.09
Spiking-Unet [6]	STFT kadras	≥32 ms	nenurodyta	2.66	0.92	2.81	3.13	3.85
DPSNN [7]	Išmoktas kadras	5 ms	0.6-0.7M	2.37	0.93	2.94	3.27	3.84
Šis darbas (128/128)	Impulsiniai rezonatoriai	0.07 ms	51K	2.10	0.91	2.74	3.18	3.55
Šis darbas (128/32)	Impulsiniai rezonatoriai	0.07 ms	14K	2.12	0.91	2.72	3.16	3.54

3 lentelė. Impulsiniais neuroniniais tinklais gristų sistemų kokybinių įverčių palyginimas. Visi rezultatai gauti naudojant VoiceBank+DEMAND duomenų rinkinį.

[5] Timcheck, J., Shrestha, S. B., ben Dayan Rubin, D., Kupryjanow, A., Orchard, G., Pindor, L., Shea, T., & Davies, M. (2023). The Intel neuromorphic DNS challenge. *Neuromorphic Computing and Engineering*, 3(3), 034005.

<https://doi.org/10.1088/2634-4386/ACE737>

[6] Riahi, A., & Plourde, E. (2023). Single Channel Speech Enhancement Using U-Net Spiking Neural Networks. *Canadian Conference on Electrical and Computer Engineering, 2023-September*, 111–116.

<https://doi.org/10.1109/CCECE58730.2023.10288830>

[7] Sun, T., & Bohté, S. (2024). DPSNN: spiking neural network for low-latency streaming speech enhancement. *Neuromorphic Computing and Engineering*, 4(4), 044008. <https://doi.org/10.1088/2634-4386/AD93F9>

Sekančio pusmečio darbo planas, 2026/2027

– I pusmetis

- Baigti publikacijos rengimą, pateikti publikaciją mokslo leidinyje, turinčiame cituojamumo rodiklį.
- Dalyvavimas „Python“ duomenų analizei mokymuose.
- Dalyvavimas doktorantų žiemos stovykloje, susijusioje su impulsiniais neuroniniais tinklais.
- Tęsti impulsiniu neuroninių tinklų taikymo problematikos tyrimą.

Klausimai

Visi rezultatai, visi įverčiai

Tinklas	Parametrų kiekis	SI-SNR	PESQ	STOI	OVRL	SIG	BAK	M-ops/s
Triukšmas		8.44	1.97	0.921	2.68	3.32	3.11	-
128/128	50945	14.13	2.10	0.911	2.74	3.18	3.55	58.5
128/32	13793	14.24	2.12	0.911	2.72	3.16	3.54	32.3
128/8	4505	12.98	2.03	0.903	2.71	3.13	3.60	24.8
64/64	13185	13.51	2.05	0.907	2.70	3.17	3.48	25.1
64/32	6945	13.22	2.06	0.906	2.75	3.20	3.55	18.7
32/32	3521	13.54	1.98	0.911	2.74	3.24	3.44	11.9

RF neuronas yra juostinis filtras
LIF neuronas yra žemų dažnių filtras

