

ЗДОРЕНКО Валерій

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0000-0001-6508-4290>

alzd123@meta.ua

ДАНИЛОВ Микита

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

e-mail: Nikita.danilov32@gmail.com

ШОЛУДЬКО Кирило

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

e-mail: kirillsholydko@gmail.com

НЕЙРОМЕРЕЖЕВІ МЕТОДИ ВИДІЛЕННЯ УЛЬТРАЗВУКОВОГО СИГНАЛУ З ШУМУ

Розглянуто нейромережеві методи виділення ультразвукового сигналу з шуму в задачах безконтактного вимірювання відстані (TOF), неруйнівного контролю (NDT) та ультразвукової візуалізації. Наведено математичні моделі корисного сигналу та завад (адитивні, корельовані, імпульсні, ревербераційні), сформульовано критерії якості (SNR, MSE/MAE, похибка оцінки часу приходу) та показано, як узгодити функцію втрат нейромережі з практичною метою — мінімізацією похибки TOF і, як наслідок, похибки відстані. Систематизовано сучасні архітектури денойзингу: 1D-CNN з резидуальним навчанням, денойзингові автоенкодера (DAE/CAE), U-Net на часово-частотних поданнях, моделі уваги/трансформери для довгих реверберацій, self-supervised підходи без «чистих» еталонів (Noise2Noise, Noise2Void-подібні) і генеративні апіорі (дифузійні моделі).

Ключові слова: ультразвуковий сигнал, завадостійкість, денойзинг, 1D-CNN, автоенкодер, U-Net, трансформер, Noise2Noise, Noise2Void, дифузійні моделі, STFT, кореляційний аналіз, TOF, похибка вимірювання.

ZDORENKO Valeriy, DANILOV Mykyta, SHOLUDKO Kyrilo

National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute"

NEURAL NETWORK METHODS FOR ULTRASOUND SIGNAL EXTRACTION FROM NOISE

This paper addresses neural network-based methods for extracting ultrasonic signals from noise in applications of time-of-flight (TOF) distance measurement, non-destructive testing (NDT), and ultrasonic imaging. The study is motivated by the limited robustness of conventional filtering and correlation techniques under non-stationary noise, reverberation, and multipath propagation, where even small distortions of the echo waveform may lead to significant TOF estimation errors. Mathematical models of ultrasonic pulse observations are considered, including additive, correlated, impulsive, and reverberation noise components, with particular emphasis on preserving the temporal structure of reflections. Quality criteria relevant to practical ultrasonic measurements are formulated, such as signal-to-noise ratio improvement, mean absolute and squared errors, and, most importantly, the time-of-arrival estimation error directly affecting distance accuracy.

The paper systematizes modern neural denoising architectures applicable to ultrasonic signals. These include one-dimensional convolutional neural networks with residual learning for real-time processing, denoising autoencoders for structured noise suppression, and U-Net-based models operating on time-frequency representations obtained via short-time Fourier transform. Attention-based models and transformers are discussed in the context of long reverberation tails and complex interference patterns. Special attention is given to training strategies in scenarios where clean reference signals are unavailable, including Noise2Noise and blind-spot self-supervised approaches, which enable learning directly from field measurements.

It is shown that optimizing neural networks solely with energy-based losses may lead to excessive smoothing and temporal bias; therefore, loss functions should be aligned with the physical measurement objective, namely minimizing TOF estimation error. Practical recommendations for integrating neural denoising modules into ultrasonic signal processing chains are provided. The results demonstrate that properly designed neural network denoisers can significantly enhance noise immunity while preserving echo timing, enabling more accurate and reliable ultrasonic measurements in challenging environments.

Key words: ultrasonic signal, noise immunity, denoising, 1D-CNN, autoencoder, U-Net, transformer, Noise2Noise, Noise2Void, diffusion models, STFT, correlation analysis, TOF, measurement error.

Стаття надійшла до редакції / Received 30.11.2025

Прийнята до друку / Accepted 04.12.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Ультразвукові датчики та перетворювачі широко застосовуються у промислових системах контролю, транспорті та медичних системах. Ключовою перевагою є можливість безконтактного вимірювання та діагностики, однак точність і стабільність таких систем обмежується співвідношенням сигнал/шум та впливом реверберацій, мультишляхових відбиттів і електромагнітних завад. У задачах TOF (time-of-flight) навіть невелике зміщення оцінки часу приходу ультразвукового відлуння призводить до відчутної похибки відстані, оскільки похибка часу прямо масштабується швидкістю звуку.

Традиційні методи фільтрації й детектування (смугові фільтри, виділення огинаючої, кореляційний аналіз) забезпечують простоту реалізації та роботу в реальному часі, проте часто втрачають ефективність у нестационарних умовах, при складній структурі шумів або коли форма імпульсу змінюється через умови середовища, частотні характеристики перетворювача та реверберації. Загальний огляд застосувань глибинного навчання в автоматизованому ультразвуковому NDE підкреслює, що нейромережі можуть бути ефективні на всіх етапах — від препроцесінгу та покращення якості сигналу до детекції та оцінки дефектів — але потребують коректної постановки задачі, даних та процедур валідації.

Актуальність роботи полягає у формуванні узгодженого підходу до нейромережевого денойзингу саме для ультразвукових сигналів: з урахуванням фізичної специфіки (імпульсність, вузькосмуговість, реверберації), з прив'язкою до прикладної мети (мінімізація похибки TOF/відстані, підвищення POD у NDT), з можливістю навчання без “чистих” еталонів у польових умовах.

Об'єктом дослідження є ультразвукові імпульсні сигнали (A-scan у часовій області та/або їх часово-частотні представлення), зареєстровані після аналого-цифрового перетворення, а також методи їх очищення від шуму для подальшого оцінювання TOF та інтерпретації відбиттів. Методи дослідження включають:

1. Аналітичне моделювання корисного сигналу та типових завад (гаусівських, корельованих, імпульсних, ревербераційних) з формулюванням метрик якості.
2. Порівняльний аналіз класичних і нейромережевих методів денойзингу з точки зору збереження часу приходу та форми імпульсу.
3. Систематизацію архітектур нейромереж та стратегій навчання: кероване (supervised), парне зашумлене (Noise2Noise) та самонавчання (Noise2Void-подібні).
4. Рекомендації інтеграції нейромережевого блоку у типову схему трактів обробки (передпідсилення/фільтрація → АЦП → цифрова обробка → TOF/рішення), аналогічно до структурного підходу, що використовується у зразку оформлення.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою дослідження є підвищення точності та завадостійкості ультразвукових вимірювань шляхом застосування нейромережевих методів виділення корисного сигналу з шуму та реверберацій.

Для досягнення мети необхідно:

1. Задати модель спостереження та критерії якості, релевантні TOF/NDT.
2. Розглянути архітектури нейромереж, які придатні для 1D сигналів і/або 2D часово-частотних подань.
3. Визначити функції втрат, що не лише максимізують SNR, а й мінімізують похибку оцінки часу приходу.
4. Описати практичні сценарії навчання з наявністю/відсутністю “чистих” еталонів.
5. Навести рекомендації щодо побудови графіків та протоколів оцінювання.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Базове ультразвукове вимірювання в часовій області описують як:

$$y(t) = x(t) + n(t), \quad (1)$$

де $x(t)$ — корисний відгук (сума відбиттів), $n(t)$ — сумарна завада. У дискретному вигляді:

$$y[n] = x[n] + n[n], n = 0, \dots, N - 1. \quad (2)$$

Для імпульсних систем характерною є модель “суми відбиттів”:

$$x(t) = \sum_{k=1}^K a_k p(t - \tau_k) + b(t), \quad (3)$$

де $p(t)$ — опорний імпульс (віконована несуча), τ_k — затримки (TOF до відповідних відбивачів), a_k — амплітуди, $b(t)$ — повільні компоненти/фонові процеси.

Ця модель корисна тим, що вона одразу показує практичну мету: відновити не “будь-який гладкий сигнал”, а структуру відбиттів і їх часові позиції.

Практично TOF часто оцінюють узгодженою фільтрацією/кореляцією:

$$r[\tau] = \sum_n y[n] p[n - \tau], \hat{\tau} = \arg \max_{\tau} r[\tau]. \quad (4)$$

Похибка часу $\Delta\tau$ переходить у похибку відстані:

$$\Delta d \approx \frac{c}{2} \Delta\tau, \quad (5)$$

де c — швидкість звуку в середовищі.

Ключове: денойзер має бути “time-faithful”: не зрушувати корисний імпульс і не деформувати його так, щоб максимум кореляції зміщувався [1].

Найчастіше для денойзингу наводять SNR:

Похибка часу $\Delta\tau$ переходить у похибку відстані:

$$\text{SNR}(x, \hat{x}) = 10 \log_{10} \frac{\sum x[n]^2}{\sum (\hat{x}[n] - x[n])^2}. \quad (6)$$

Однак для ультразвуку потрібно оцінювати паралельно:

- $\Delta\tau$ (зміщення/розкид TOF),
- долю “втрачених” слабких відбиттів (пропуски),
- частоту хибних спрацювань (фальш-піки),
- стабільність на зміні умов (температурні режими, інший матеріал тощо).

Смугові фільтри добре працюють тоді, коли сигнал вузькосмуговий, а значна частина шуму лежить поза смугою. Але якщо у смузі присутні реверберації або структурний шум, ефект обмежений. Кореляція з шаблоном є оптимальною для AWGN і стабільного шаблону. Реально шаблон $p(t)$ змінюється (перетворювач, середовище, контакт, геометрія), і фіксований шаблон дає систематичні помилки. Хвильлет-порогування добре для розріджених сигналів, але для ультразвуку з ревербераціями “розрідженість” може втрачатися, а вибір порогу/бази стає критичним. Саме на цьому фоні нейромережі цікаві як “адаптивний фільтр”, здатний вчитись на статистиці реальних сигналів (або на синтезованому домені зі схожими властивостями) та відокремлювати корисну структуру від типових завод.

Для часових сигналів природно використовувати 1D згортки. Ідея проста: мережа отримує на вході фрагмент $y[n]$ і повертає $\hat{x}[n]$. Практично зручно робити резидуальну постановку:

$$\hat{x} = y - g_{\theta}(y), \quad (7)$$

де $g_{\theta}(y)$ оцінює шумову компоненту. Чому це працює добре: мережі легше “вивчити” структуру шуму/артефактів, ніж “відтворювати” весь сигнал з нуля; крім того, це стабілізує градієнти та пришвидшує збіжність.

Сильні сторони 1D-CNN для ультразвуку:

- низька затримка обробки (важливо для реального часу);
- просте впровадження на вбудованих платформах (після оптимізації);
- можливість навчити модель специфічно під свій датчик і тракт.

Типова помилка: навчати лише на MSE так, що мережа “згладжує” імпульсні відбиття — сигнал стає “чистішим”, але слабкі відбиття втрачаються або TOF зміщується.

Автоенкодер стискає сигнал у латентний простір і відновлює назад:

$$\hat{x} = D_{\theta}(E_{\theta}(y)). \quad (8)$$

Інтуїція така: корисний сигнал має обмежену структурну складність (типовий імпульс + кілька відбиттів + затухання), а шум є “надлишковим”. Латентний простір грає роль “пляшкового горла”, не пропускаючи випадкові компоненти. Автоенкодери часто добре працюють у NDT/UGW, де шум формується складними середовищними ефектами, а не просто AWGN. Але для TOF систем треба обережно: занадто агресивне стискання може “з’їсти” перші слабкі компоненти.

Якщо шум і сигнал мають різні часово-частотні “підписи”, зручно перейти до STFT-спектрограми і очистити її в 2D просторі. U-Net завдяки skip-зв’язкам добре відновлює дрібні деталі, не втрачаючи глобальний контекст. Ультразвуковий імпульс зазвичай має енергію навколо несучої, тоді як шум може бути широкосмуговим або містити смуги перешкод. У 2D поданні це часто стає виразнішим. Self-attention корисний, коли потрібно враховувати далекі залежності: наприклад, довгі ревербераційні хвости або повторювані артефакти в записі. Проте ресурсна вартість зростає, тому для простих TOF систем це не завжди

виправдано. У реальних вимірюваннях часто немає “чистого” x , і тоді виникає проблема supervised-навчання. Є два практичні сценарії:

- Noise2Noise: якщо можна отримати два вимірювання одного й того ж сигналу з незалежним шумом, тоді вчимо мережу перетворювати y_1 у y_2 . Це працює, якщо шум має нульове середнє і незалежний між вимірюваннями.
- Blind-spot (Noise2Void-ідея): мережа прогнозує відлік по його оточенню, не “бачачи” самого відліку, уникаючи тривіального копіювання шуму.

Для ультразвуку це особливо привабливо: можна збирати дані прямо з датчика в робочих умовах, не потребуючи лабораторного “еталону”.

Почнемо з того, що практично цікавить інженера: як виглядає сигнал до/після і чи видно відбиття. На Рис. 1 показано три реалізації: чистий $x(t)$, зашумлений $y(t)$ та відновлений $\hat{x}(t)$. Типова картина: шум “піднімає підлогу”, слабкі імпульси стають невиразними, з’являються хибні піки. Денойзер зменшує фон і “витягує” імпульси, що критично для коректної кореляції та TOF.

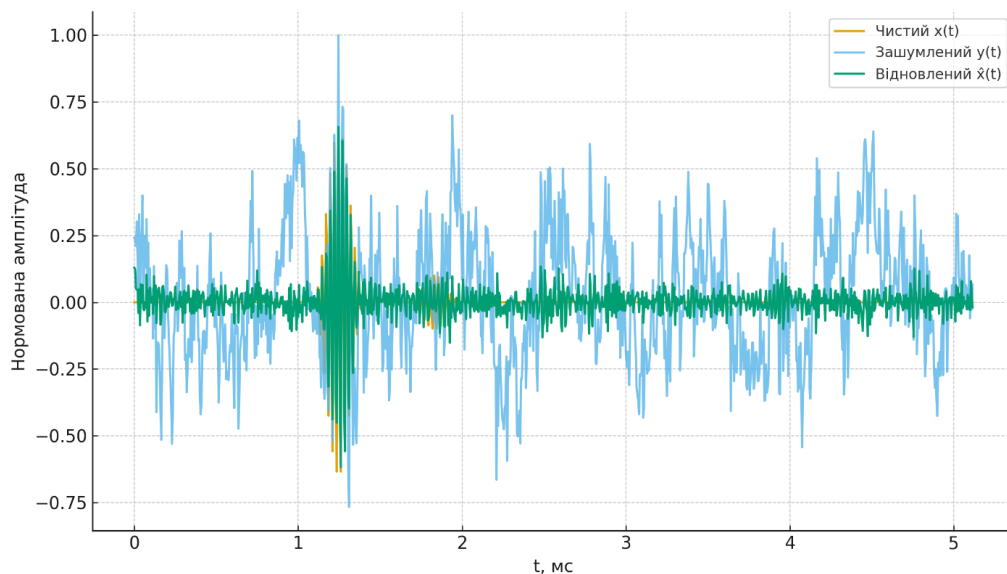


Рис. 1. Часові реалізації: чистий сигнал, зашумлений та відновлений 1D-CNN

Важливий нюанс: навіть якщо амплітуда шуму зменшилась, треба дивитися, чи не “поплив” фронт/положення максимуму [2, 3]. Для цього в експериментах зазвичай обчислюють кореляцію до/після і порівнюють $\hat{t}$.

Спектрограма — дуже інформативний інструмент у ультразвуку, бо імпульс має локалізовану енергію в часі та частоті. На Рис. 2 показано зашумлений сигнал: енергія корисної компоненти частково маскується шумовим фоном. Після денойзингу (Рис. 3) фон зменшується, а частотна структура корисної компоненти проявляється виразніше. Це важливо з практичної точки зору: кореляційний детектор, смугові фільтри та огинаюча працюють стабільніше, коли енергія “зібрана” в типову смугу та часовий інтервал.

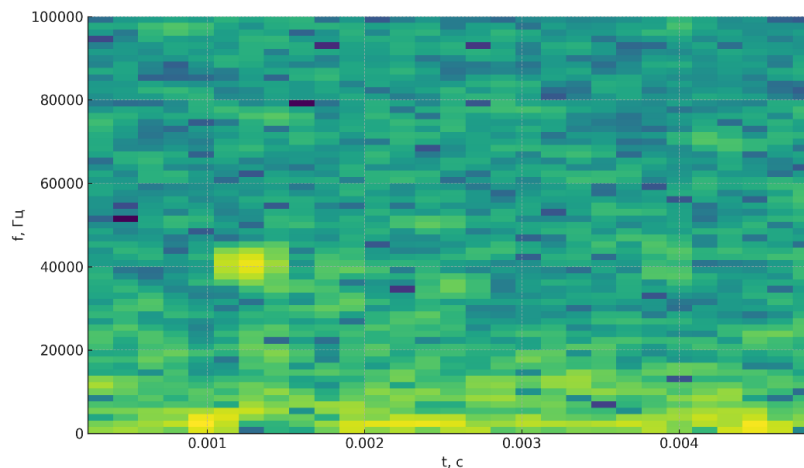


Рис. 2. Спектрограма зашумленого сигналу $y(t)$

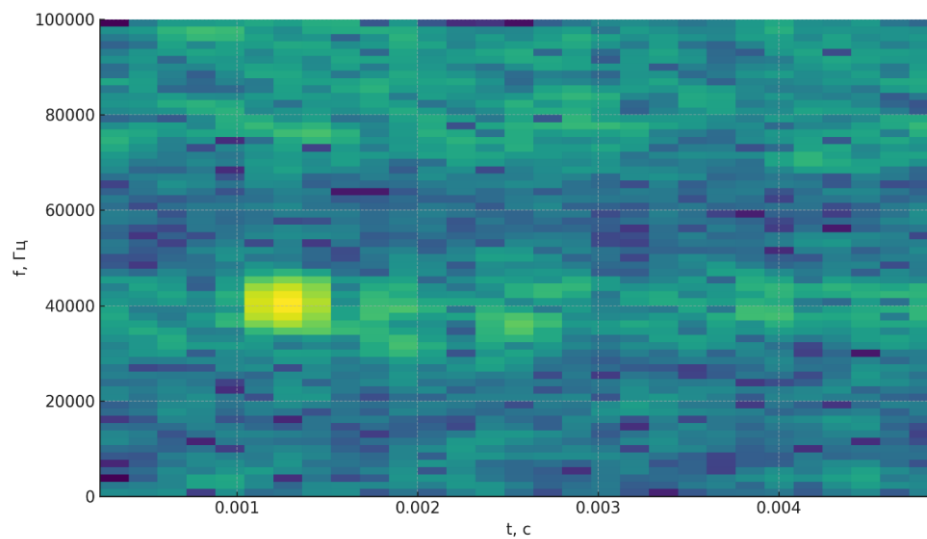


Рис. 3. Спектрограма сигналу після денойзингу $\hat{x}(t)$

Практичне правило: якщо після денойзингу спектр “розмазується” ширше або з’являються штучні смуги, це сигнал проблеми — мережа може “домальовувати” компоненти, що погано для вимірювання.

Найпоширеніша втрата — MSE:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_n (\hat{x}[n] - x[n])^2. \quad (9)$$

Але для ультразвуку MSE може стимулювати згладжування. Тому часто застосовують комбінації:

- L_1 або Huber для робастності до імпульсних завад;
- додаткові спектральні/часово-частотні терміни, щоб зберігати форму імпульсу;
- TOF-орієнтовані штрафи, що прямо зменшують зсув кореляційного піка.

У прикладних проектах добре працює “прагматичний” підхід: спочатку навчити модель на MSE/MAE, а потім робити донавчання з додаванням терміна, який карає розбіжність у $\hat{t}$ (або карає форму кореляційної функції навколо піка).

Навіть проста архітектура може перенавчитися на конкретний шум/тракт. Тому важливо контролювати збіжність і розрив між train/val [4]. Типова крива навчання показана на Рис. 4: якщо train loss падає, а val стоїть або росте — це ознака перенавчання. Для реальних вимірювальних систем це критично: модель, що працює “ідеально” на лабораторних даних, може провалитись у польових умовах.

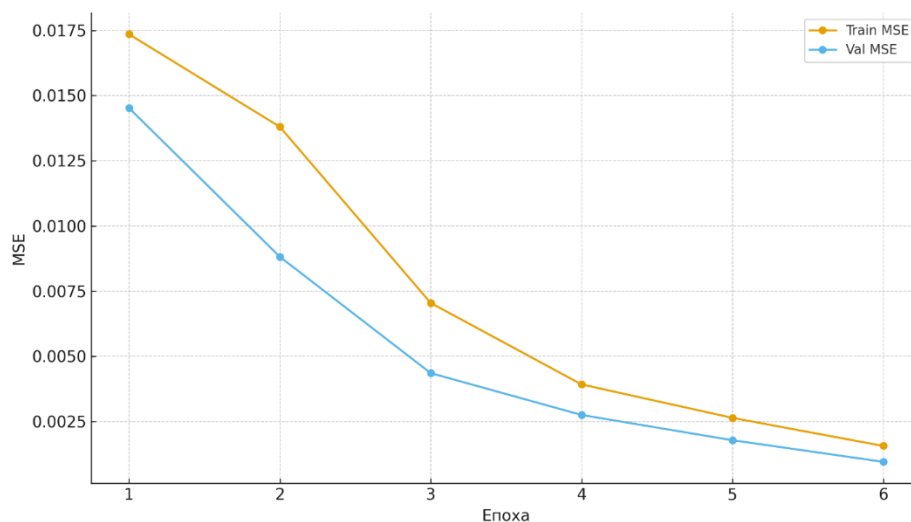


Рис. 4. Крива навчання (Train/Val MSE) на синтетичних даних

Щоб зменшити domain gap між синтетикою та реальністю, застосовують:

- домен-рандомізацію шумів (AWGN + корельовані + імпульсні + ревербераційні компоненти);
- варіації імпульсу (частота, затухання, ширина, фазові зсуви);
- випадкові зміни масштабів амплітуди та нормалізаційних параметрів;
- невеликий fine-tuning на реальних записах.

Один “красивий” приклад на графіку не гарантує стабільності. Тому ефект оцінюють статистично — розподілом приросту SNR:

$$\Delta \text{SNR} = \text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}. \quad (10)$$

Якщо гістограма зміщена вправо, маємо стабільний виграш. Якщо розподіл широкий або має “хвіст” у від’ємні значення, це сигнал, що модель інколи погіршує сигнал (наприклад, на певних типах шуму або при дуже слабких відбиттях) [5].

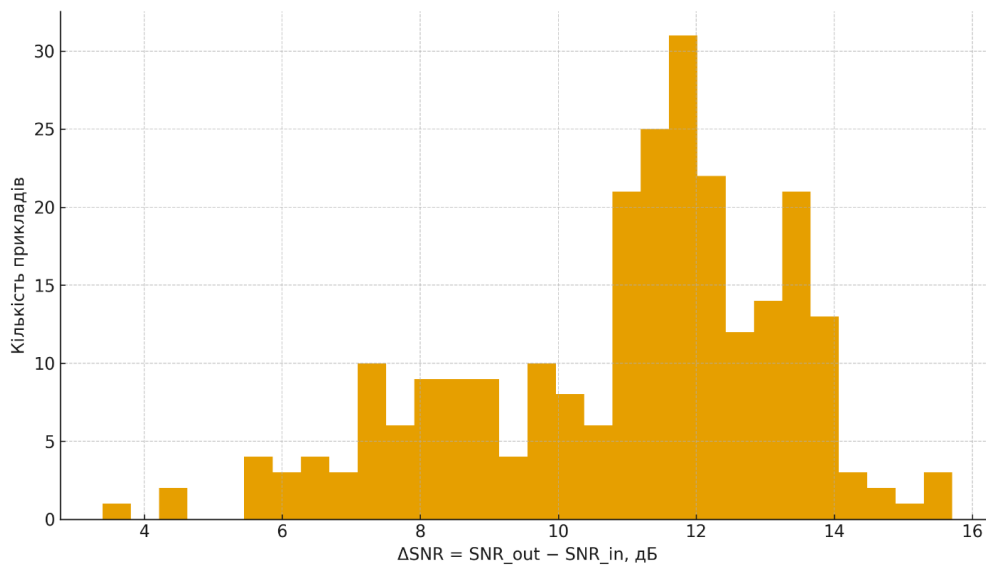


Рис. 5. Розподіл приросту SNR після денойзингу на тестовому наборі

На практиці цього замало: потрібно робити таку ж статистику по Δt_i по частоті пропусків/фальш-піків. Для мікроконтролерів частіше обирають 1D-CNN (мінімум перетворень, прогнозована затримка). Для ПК/потужних SoC у NDT можна дозволити TF-подання. Окреме інженерне питання — реальний час: денойзер має бути реалізований так, щоб обробляти вікно довжини N швидше, ніж надходять дані. Тут допомагають: зменшення числа каналів/шарів, квантування (INT8), оптимізовані бібліотеки та профілювання.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У роботі розглянуто нейромереві методи виділення ультразвукового сигналу з шуму та показано їхню доцільність для задач TOF, NDT і ультразвукової обробки сигналів за умов корельованих, імпульсних завад і реверберацій. На прикладі 1D-CNN денойзера продемонстровано, що нейромережа здатна істотно зменшувати шумовий фон, підкреслювати імпульсні відбиття та покращувати енергетичні показники якості, що підтверджено часовими реалізаціями, спектрограмами та статистикою приросту SNR. Водночас встановлено, що для ультразвукових вимірювань ключовим критерієм має бути не лише MSE/SNR, а й збереження часової структури сигналу, оскільки можливі фазові/часові спотворення денойзером безпосередньо впливають на оцінку часу приходу та похибку відстані.

Показано, що найбільш практичними для обробки A-scan у режимі реального часу є 1D резидуальні згорткові моделі, тоді як для складних спектрально-неоднорідних шумів ефективними є підходи на часово-частотних поданнях (STFT) із U-Net-подібними архітектурами. Для випадків відсутності “чистих” еталонів перспективними є self-supervised стратегії навчання, однак вони потребують перевірки статистичних припущень щодо шуму та обов’язкової валідації на прикладних метриках. Практичним підсумком є рекомендація оцінювати нейромеревий денойзинг комплексно: поряд із приростом SNR аналізувати похибку TOF, стабільність кореляційного піка та ймовірність виявлення слабких відбиттів, що забезпечує

фізично коректне та метрологічно обґрунтоване застосування нейромереж у ультразвукових вимірювальних системах.

Література

1. Li X., Navab N., Jiang Z. *Speckle2Self: Self-Supervised Ultrasound Speckle Reduction Without Clean Data*. 2025.
2. Donoho D. L. De-noising by soft-thresholding. *IEEE Trans. on Information Theory*, 1995.
3. Lehtinen J. et al. *Noise2Noise: Learning Image Restoration without Clean Data*. ICML, 2018.
4. Himanshu Singh, Arif Sheikh Ahmed, Frank Melands, Anowarul Habib. *Ultrasonic image denoising using machine learning in point contact excitation and detection method*. 2023
5. Jing Wu , Yingfeng Yang, Zeyu Lin, Yizhou Lin, Yan Wang, Weiwei Zhang, Hongwei M. *Weak ultrasonic guided wave signal recognition based on one-dimensional convolutional neural network denoising autoencoder and its application to small defect detection in pipelines*. 2025