

УДК 004.96

DOI <https://doi.org/10.32782/2663-5941/2026.3.1/25>

**Олексій А.О.**

<https://orcid.org/0009-0006-5354-8098>

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

**Дичка І.А.**

<https://orcid.org/0000-0002-3446-3076>

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

**Олещенко Л.М.**

<https://orcid.org/0000-0001-9908-7422>

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

**Хічко Я.В.**

<https://orcid.org/0000-0002-6455-8498>

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

**Суцук-Слюсаренко В.І.**

<https://orcid.org/0000-0002-6096-3832>

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

## ОГЛЯД МОЖЛИВИХ УДОСКОНАЛЕНЬ НЕЙРОМЕРЕЖЕВОГО МЕТОДУ UWAR-GAN ТА НЕЙРОМЕРЕЖЕВОЇ АРХІТЕКТУРИ U-NET ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРИДУШЕННЯ ФОНОВИХ ШУМІВ

У статті здійснено огляд та проаналізовано можливі вдосконалення методу UWAR-GAN для придушення підводного фонового шуму, який поєднує генератор U-Net з дискримінатором PatchGAN та втратою кореляції. Базова архітектура продемонструвала ефективність при відношенні сигнал/шум до -10 дБ на обмежених наборах даних. Досліджувані архітектурні модифікації включають механізми уваги для вибіркового фокусування на тональних компонентах корабля, гібриди трансформатора-згорткової мережі, такі як CGF-Unet та Speed-UT2-CGAN, для моделювання залежностей на великій відстані, мережі зі складними значеннями, такі як Deep Complex U-Net, для обробки фазової інформації, багатомасштабне вилучення ознак та архітектури комплементарного навчання на основі вейвлетів, прикладом яких є ERCL-AttentionNet. Неархітектурні вдосконалення охоплюють функції втрат на основі сприйняття та стилю, стратегії навчання Noise2Noise, що усувають вимоги до "чистих" даних, методи адаптивного навчання для умов з низьким SNR, представлені NAFSA-Net із зваженою втратою SI-SNR, самодистилляцію з множинними функціями втрат, як у ESTMST-ST, та модулі налаштування ефективною адаптації з обмеженими даними. Кожна модифікація підкріплена рецензованими дослідженнями, що демонструють вимірювані покращення на наборах даних підводної акустики, включаючи ShipsEar та DeepShip. Цей огляд синтезує різноманітні підходи у комплексну структуру для вдосконалення UWAR-GAN, приділяючи особливу увагу подоланню фундаментальних викликів збереження фази, дуже низького значення SNR, узагальнення на невидимі умови та ефективність використання даних. Результати дослідження свідчать, що гібридні архітектури, незалежні від шуму, пропонують найбільш перспективні напрямки для майбутньої експериментальної валідації та практичного впровадження в системах підводного моніторингу в режимі реального часу.

**Ключові слова:** нейронна мережа, придушення підводного фонового шуму, UWAR-GAN, вдосконалення архітектури U-Net, часово-частотна обробка сигналів, глибоке навчання в акустиці, датасет ShipsEar.



**Постановка проблеми.** Визначено науково-прикладну проблему, яка полягає в необхідності підвищення ефективності нейромережових методів очищення гідроакустичних сигналів, зокрема, методу UWAR-GAN, для роботи в реальних умовах експлуатації. Метод UWAR-GAN, який поєднує генератор на основі архітектури U-Net з дискримінатором PatchGAN та кореляційною функцією втрат, продемонстрував значне покращення якості відновлення сигналів суден порівняно з класичними статистичними методами та деякими GAN-підходами, особливо в умовах низького співвідношення сигнал/шум та обмежених навчальних вибірок. Водночас аналіз літератури засвідчує наявність низки обмежень, що стримують подальше вдосконалення цього методу. До них належать: недостатня ефективність роботи в умовах низького співвідношення сигнал/шум, обмежена здатність до узагальнення на невідомі типи суден та шумів, а також залежність від наявності великих обсягів даних. Окрім того, використання лише амплітудної інформації або її спрощена обробка в багатьох існуючих архітектурах призводить до втрати фазових характеристик сигналу, що негативно позначається на якості реконструкції. Розв'язання зазначених проблем є важливим науковим завданням, оскільки дозволяє суттєво підвищити надійність систем пасивної гідроакустики, що використовуються для морської безпеки, навігації, екологічного моніторингу та захисту морських екосистем. Подальший розвиток методу UWAR-GAN шляхом інтеграції сучасних архітектурних рішень та неархітектурних удосконалень відкриває перспективи для створення високоефективних, стійких до шумів та універсальних систем обробки гідроакустичної інформації в режимі реального часу.

**Аналіз останніх досліджень і публікацій.** Популярність та ефективність нейромережової архітектури U-Net для задач шумопригнічення підтверджується численними сучасними дослідженнями, що розширюють її базову концепцію. Наприклад, у роботі, присвяченій придушенню морського шуму за допомогою Deep Complex U-Net (DCUNet), дослідники успішно застосували цю архітектуру для обробки нелінійних та негаусівських сигналів, досягнувши покращення SNR більш ніж на 10 дБ навіть при початковому дуже низькому значенні SNR [2]. Це дослідження також підкреслює важливість використання комплекснозначних даних для збереження фазової інформації, що перегукується з ключовою ідеєю UWAR-GAN. Інший приклад – оптимізо-

вана структура OUNet-JL, яка інтегрує багатозалишкові модулі та просторову багатомасштабну обробку ознак для покращення якості підводних зображень, демонструючи гнучкість U-Net для роботи з різними типами підводних даних [3]. Навіть у суміжних галузях, таких як обробка медичних зображень, модифікації U-Net, такі як U-Tunnel-Net, показують вищі результати завдяки зміні архітектури пулінгу та впровадженню нових згорткових блоків [4].

Таким чином, враховуючи успіх методу UWAR-GAN та загальний високий потенціал архітектури U-Net, підтверджений низкою досліджень, актуальною науково-практичною задачею є подальша модифікація нейромережі U-Net в рамках методу UWAR-GAN. Метою такої модифікації є підвищення якості придушення підводних шумів шляхом інтеграції сучасних рішень, що дозволить ще ефективніше відновлювати корисний сигнал в умовах нестационарних завад та обмежених навчальних вибірок.

**Постановка завдання.** Придушення підводних шумів є критично важливим завданням для забезпечення ефективної роботи гідроакустичних систем моніторингу, навігації та зв'язку. Метод UWAR-GAN, заснований на архітектурі U-Net, продемонстрував високу ефективність у відновленні корисних сигналів суден за умов низького співвідношення сигнал/шум та обмежених навчальних вибірок. Водночас стрімкий розвиток глибокого навчання відкриває нові можливості для подальшого покращення якості шумопридушення, розширення спектру робочих умов та підвищення узагальнювальної здатності моделі.

У зв'язку з цим, метою даного дослідження є вирішення наступних задач.

1. Проведення систематичного аналізу та узагальнення сучасних архітектурних та неархітектурних методів удосконалення нейромереж U-Net, які можуть бути інтегровані в існуючий метод UWAR-GAN для підвищення ефективності придушення підводних шумів.

2. Дослідження та класифікація перспективних підходів, включаючи механізми уваги, трансформер-згорткові гібриди, комплексні нейромережі, вдосконалені функції втрат (Style loss, weighted SI-SNR) та стратегії навчання (Noise2Noise, самодистиляція) з метою визначення їхнього потенціалу для покращення якості відновлення сигналів суден.

3. Оцінка можливості подолання ключових обмежень базової моделі UWAR-GAN за допомогою розглянутих модифікацій, зокрема для

покращення збереження фазової інформації, підвищення стійкості до екстремально низьких співвідношень сигнал/шум (до -20 дБ), покращення узагальнення на невідомі типи суден та шумів, а також зменшення залежності від наявності великих обсягів "чистих" даних для навчання.

4. Формування комплексної основи (framework) для подальших експериментальних досліджень, визначаючи найбільш перспективні комбінації архітектурних рішень та стратегій навчання, здатних забезпечити найвищу якість придушення шумів для практичного застосування в системах реального часу.

5. Обґрунтування актуальності модифікації методу UWAR-GAN для підвищення ефективності вирішення прикладних задач морської безпеки, екологічного моніторингу, навігації та зв'язку.

**Виклад основного матеріалу.** Розглянемо основні напрямки архітектурних модифікацій, які можуть підвищити ефективність придушення підводних шумів у нейромережі U-Net, що є основою генератора методу UWAR-GAN (рис. 1).

Основою методу є архітектура GAN зі змалювальним навчанням генератора та дискримінатора. Як дискримінатор використовується PatchGAN – згортковий бінарний класифікатор для оцінки якості очищених сигналів, тоді як генератор реалізовано на базі U-Net, що виконує шумопригнічення за допомогою згорткового автоенкодера з пропусковими з'єднаннями. Вхідними даними є магнітудні та фазові спектрограми зашумлених сигналів, які після обробки перетворюються у часову форму за допомогою оберненого перетворення Фур'є.

Метод ґрунтується на використанні згорткової мережі Conv1D для вилучення ознак із «сирого» сигналу, які перетворюються у багатовимірне

представлення та структуруються у тензор для подальшої обробки. Основою моделі є двошляхова архітектура трансформера, що поєднує внутрішньо- та міжпослідовні механізми для моделювання відповідно локальних та глобальних залежностей у сигналі. Відбувається виділення релевантних ознак та пригнічення шуму, доповнюючись позиційним кодуванням, залишковими з'єднаннями та нормалізацією для стабілізації навчання. Після обробки ознак реконструкція сигналу виконується за допомогою транспонованої згорткової нейронної мережі з використанням методу перекриття-додавання, що дозволяє відновити очищений сигнал у часовій області (рис. 2).

У дослідженні використовувались датасети з різними класами кораблів з записами, що забезпечують достатній обсяг даних для навчання. Датасет ShipsEar містить дані про близько 80 суден, поділених на чотири класи, а також окремі записи фонових шумів (вода, вітер, дощ, хвилі), що дозволяє формувати різні шумові комбінації. Для навчання було сформовано близько 2000 комбінованих зразків із середнім SNR  $\approx -10$  дБ. Для забезпечення якості сигналів відбиралися зразки з достатньою амплітудою (понад 0,03). Представлення даних включало спектрограми, нормалізовану потужність, логарифмічний спектр та амплітуду сигналу. Фонові шуми сегментувалися на відрізки тривалістю 2 с і комбінувалися між собою для розширення вибірки. Це дозволило підвищити різноманіття шумових умов та сформувати репрезентативний набір даних для навчання моделей шумопригнічення. Продуктивність оцінювалася на датасетах ShipsEar та DeepShip [15], що містять підводні сигнали з різними рівнями шуму. Результати показали суттєве покращення показників SNR та SI-SNR, підтверджуючи ефективність моделі у шумопригніченні та підвищенні якості сигналу.

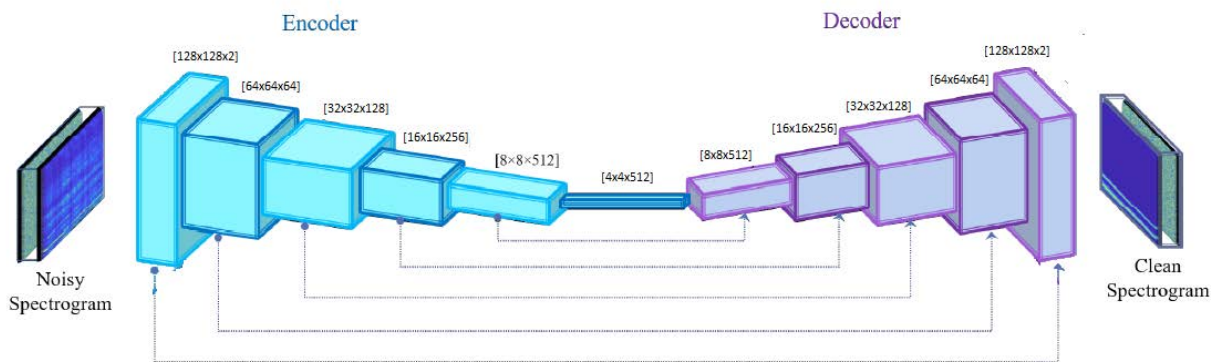


Рис. 1. Архітектура генератора U-Net.

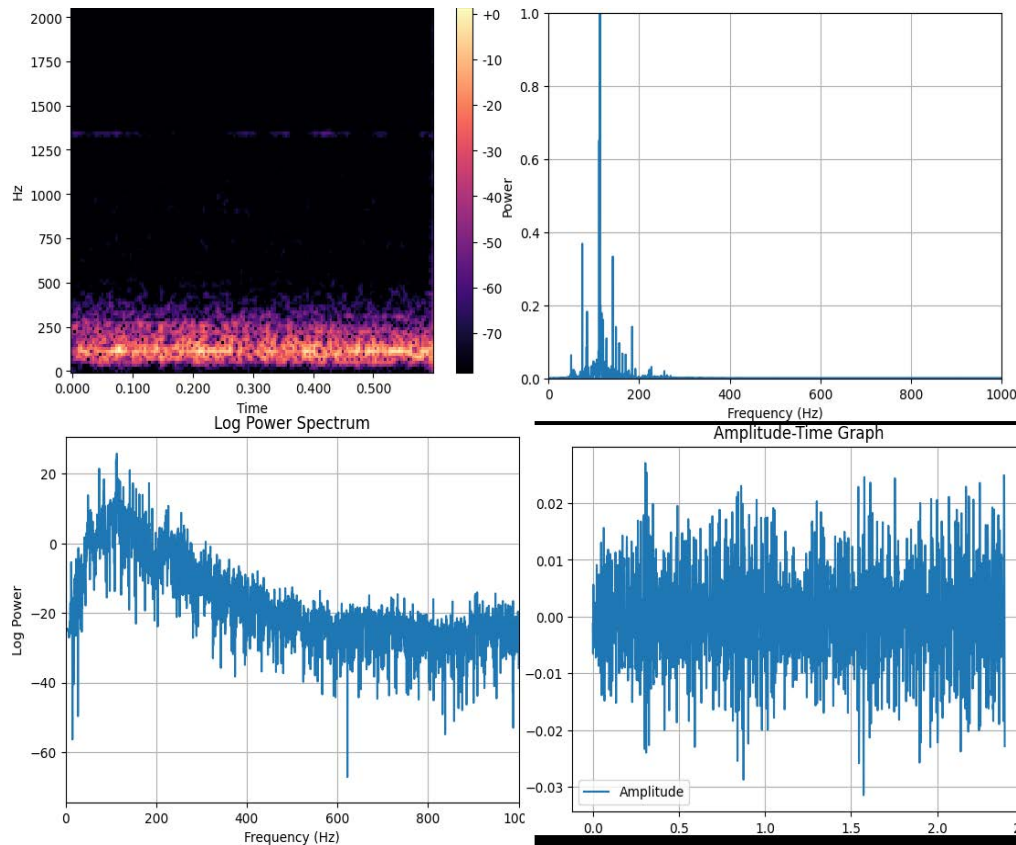


Рис. 2. Приклад представлень семпла акустичного сигналу

Аналіз сучасних досліджень дозволяє виділити дві ключові категорії вдосконалень, які вже довели свою ефективність у суміжних задачах обробки сигналів та зображень: інтеграція механізмів уваги та впровадження багатомасштабних архітектурних рішень. Перспективним напрямком модифікації є інтеграція в архітектуру U-Net механізмів уваги (attention mechanisms), які дозволяють моделі вибірково фокусуватися на найбільш інформативних ділянках вхідних даних. У контексті задачі придушення підводних шумів це означає здатність мережі краще виділяти тональні складові сигналу судна на фоні широкосмугового шуму. Дослідження, присвячене алгоритму придушення гідроакустичних сигналів на основі U-Net, пропонує інтеграцію механізму уваги, що дозволяє мережі селективно зосереджуватися на вмісті та локалізації в картах ознак [5]. Результати моделювання як у середовищі білого гаусівського шуму, так і в умовах реального океанічного шумового фону підтвердили, що запропонований алгоритм перевершує традиційні методи шумопригнічення та класичні алгоритми придушення шуму в зображеннях [5]. Інша робота, присвячена сегментації зображень гідролокатора переднього

огляду, представляє модель Feature Pyramid U-Net with Attention (FPUA), яка використовує піраміду ознак у поєднанні зі структурою уваги для покращення здатності моделі вивчати глибоку семантичну та поверхневу детальну інформацію [6]. Це дослідження також демонструє, що запропонована модель працює краще за інші для класів із малими об'єктами та обмеженою кількістю зразків, що є особливо актуальним для задачі придушення шумів, де корисний сигнал може займати лише невелику область у частотно-часовому просторі [6].

Інше дослідження, присвячене покращенню якості підводних зображень, пропонує архітектуру OUNet-JL, яка інтегрує багатомасштабне виділення ознак через просторовий багатомасштабний модуль з каналною увагою (SMFM), що використовує розширені згортки з різними коефіцієнтами для захоплення контекстної інформації на кількох масштабах, а також модуль відновлення ознак із посиленням-операцією-відніманням (SOSFM) для підвищення співвідношення сигнал/шум<sup>33</sup>. Експерименти на наборах даних UIEB та UFO-120 показали, що OUNet-JL перевершує сучасні методи за показниками PSNR та SSIM,

демонструючи потенціал багатомасштабних підходів для задач покращення якості зображень у водному середовищі [3].

Одним з найбільш перспективних напрямків є інтеграція в архітектуру U-Net трансформерних механізмів, що дозволяє поєднати переваги згорткових нейромереж для виділення локальних ознак із здатністю трансформерів моделювати довготривалі залежності в сигналі. Дослідження, опубліковане в IEEE Journal of Oceanic Engineering, пропонує архітектуру CGF-Unet, яка стратегічно впроваджує трансформерні блоки під час даунсемплінгу та апсемплінгу, розширюючи доступ до глобальної контекстної інформації та поєднуючи потужне кодування послідовностей трансформера з просторовою інваріантністю згорткових мереж [7]. Експерименти на різних наборах даних гідролокатора бічного огляду продемонстрували виняткові показники mIOU 89.3% та 86.5%, що перевершує існуючі методології, причому метод зберігає стійку продуктивність навіть за наявності шумових збурень [6]. Ще більш спеціалізованим рішенням є архітектура Speed-UT2-CGAN, представлена в Journal of the Acoustical Society of America, яка використовує двогілкову умовну генеративно-змагальну мережу, що поєднує згорткову гілку U-Net для виділення локальних ознак із часових аудіопослідовностей та трансформерну гілку уваги для моделювання довготривалих часових залежностей [8]. Ця архітектура додатково враховує швидкість руху апарату як умовний вхідний параметр для динамічної адаптації до шумових характеристик, що змінюються

залежно від швидкості, досягаючи вихідного SNR 6.6 дБ при вхідному SNR -5 дБ та середнього коефіцієнта кореляції 0.87 [8].

Важливим напрямком досліджень є використання комплексних нейронних мереж для ефективної обробки як амплітудної, так і фазової інформації сигналу. Дослідження, опубліковане в Journal of Applied Geophysics, застосовує Deep Complex U-Net (DCUNet), спеціально перепроектовану з урахуванням особливостей гідроакустичних сигналів, які характеризуються нелінійністю, нестационарністю та негаусівським розподілом [2]. Процес включає перетворення зашумлених даних у LOFAR-спектрограми для подання на вхід DCUNet, причому автори застосували метод навчання Noise2Noise для реконструкції підводного фонового шуму через наскрізну архітектуру. Результати експериментів показали, що метод досягає покращення SNR більш ніж на 10 дБ і здатен відновлювати сигнали з початковим SNR -20 дБ, демонструючи кращу продуктивність порівняно з традиційними алгоритмами шумопопригнічення.

Окремим напрямком є використання вейвлет-перетворення в поєднанні з комплементарним навчанням для придушення реверберації. Дослідження, опубліковане в журналі Oceans, пропонує мережу ERCL-AttentionNet (Echo-Reverberation Complementary Learning Attention Network), яка використовує неперервне вейвлет-перетворення для виділення часо-частотних ознак із прийнятого сигналу, ефективно балансує часову та частотну роздільну здатність [9]. Дійсна та уявна частини часо-частотної матриці об'єднуються

Таблиця 1

Архітектурні удосконалення архітектури U-Net

№	Назва	Опис	Очікуваний ефект	Джерела
1	Інтеграція механізмів уваги (Attention U-Net)	Мережа вчиться вибірково фокусуватися на найбільш інформативних ділянках (тональних складових сигналу), ігноруючи шум.	Дозволяє моделі краще виділяти корисний сигнал на фоні широкосмугового шуму, особливо в умовах низького SNR.	[5], [6]
2	Трансформер-згорткові гібриди (CGF-Unet, Speed-UT2-CGAN)	Поєднання згорткових шарів (для локальних ознак) з трансформерами (для глобальних залежностей у часі).	Забезпечує краще моделювання довготривалих залежностей у сигналі, що важливо для відновлення структури суднових шумів.	[7], [8]
3	Комплексні нейромережі (Deep Complex U-Net)	Архітектура, яка працює з комплексними числами, дозволяючи ефективно обробляти як амплітудну, так і фазову інформацію.	Зберігає фазову інформацію сигналу, критично важливу для високоякісної реконструкції, особливо при дуже низькому SNR.	[2]
4	Багатомасштабне виділення ознак	Використання паралельних згорткових ядер різного розміру для аналізу сигналу на різних рівнях деталізації.	Дозволяє одночасно відновлювати як низькочастотні, так і високочастотні гармоніки сигналу судна.	[3]
5	Вейвлет-перетворення + комплементарне навчання (ERCL-AttentionNet)	Використання вейвлет-перетворення для балансування часової та частотної роздільної здатності та окреме моделювання сигналу й реверберації.	Підвищує якість виділення цільового сигналу шляхом одночасного вивчення ознак корисного відлуння та завади.	[9]

для генерації представлень уваги, які обробляються мережею. Архітектура складається з двох комплементарних моделей U-Net, які спільно використовують один енкодер та незалежно вивчають ознаки цільового відлуння та реверберації для реконструкції цільового відлуння, причому механізм уваги додатково підвищує продуктивність, фокусуючись на цільовій інформації та пригнічуючи нерелевантні перешкоди [9].

Окремо варто відзначити інтегровані фреймворки, такі як AquaSignal, який об'єднує архітектуру U-Net для шумопригнічення з ResNet18 для класифікації відомих акустичних подій та автоенкодерною моделлю для неконтрольованого виявлення нових або аномальних сигналів [10]. Експериментальні результати показують, що AquaSignal покращує чіткість сигналу та продуктивність виконання задач, досягаючи 71% точності класифікації та 91% точності виявлення новизни, що демонструє потенціал для моніторингу в реальному часі в наукових, екологічних та морських доменах [10].

Таким чином, аналіз літератури підтверджує, що подальше вдосконалення архітектури U-Net для придушення підводних шумів може здійснюватися за декількома напрямками: інтеграція трансформерних механізмів для моделювання довготривалих залежностей, використання комплексних нейронних мереж для ефективної обробки фазової інформації, застосування вейвлет-перетворення в поєднанні з комплементарним навчанням, оптимізація згорткових блоків для багатомасштабного виділення ознак та створення інтегрованих фреймворків, які об'єднують шумопригнічення з подальшими задачами класифікації та виявлення аномалій. Ці модифікації дозволяють мережі краще адапту-

ватися до складних акустичних умов, ефективніше виділяти корисний сигнал на фоні нестаціонарних завад та зберігати важливі структурні особливості сигналу, що є критичним для подальших задач класифікації та ідентифікації суден. У Таблиці 1 перелічені можливі архітектурні покращення для нейромережі U-Net в методі UWAR-GAN, які були розглянуті вище.

Окрім архітектурних модифікацій, існує широкий спектр інших підходів до покращення нейромережі U-Net в методі UWAR-GAN, які зосереджені на стратегіях навчання, функціях втрат, методах обробки даних та підвищенні ефективності навчання в умовах обмежених даних. Ці напрямки є особливо актуальними для задач підводної акустики, де отримання великих обсягів розмічених даних є складним завданням, а рівень шуму часто є екстремально низьким.

Одним із найбільш перспективних напрямків є вдосконалення функцій втрат, які виходять за рамки простого порівняння пікселів і дозволяють мережі навчатися на більш високорівневих ознаках. Дослідження, опубліковане в IEEE Transactions on Geoscience and Remote Sensing у 2025 році, пропонує інноваційну стратегію навчання на основі мульти-втрат та самодистиляції (self-distillation) для моделі Swin Transformerx [11]. Модель ESTMST-ST використовує чотири специфічні функції втрат на різних етапах навчання, що дозволяє значно підвищити стійкість до шуму та ефективність навчання. Експерименти на загальнодоступних наборах даних ShipsEar та DeepShip показали, що цей підхід досягає покращення F1-показників щонайменше на 1.6 та коефіцієнтів Каппа на 2.2 порівняно з п'ятьма сучасними методами.

Таблиця 2

**Неархітектурні удосконалення архітектури U-Net**

№	Назва покращення	Основна суть	Короткий опис	Джерела
1	Стратегія навчання Noise2Noise	Навчання моделі придушення шуму без використання "чистих" цільових сигналів, лише на парі зашумлених зразків.	Вирішує проблему відсутності "чистих" даних, дозволяючи використовувати для навчання наявні зашумлені записи.	[2]
2	Адаптивне навчання для наднизького SNR (NAFSA-Net)	Використання noise-aware підходів та спеціалізованих функцій втрат (weighted SI-SNR) для роботи в умовах екстремально низького SNR.	Підвищує стійкість та ефективність моделі в найскладніших умовах, де сигнал майже повністю прихований шумом.	[12]
3	Самодистиляція та мульти-втрати (ESTMST-ST)	Використання кількох функцій втрат на різних етапах навчання та передача знань від глибшої моделі до меншої.	Значно підвищує точність та стійкість до шуму, що підтверджено покращенням метрик на наборах даних ShipsEar та DeepShip.	[11]
4	Додаткові модулі для налаштування моделі нейронної мережі	Невеликий модуль, що підключається до вже навченої моделі для корекції її поведінки в умовах шуму без повного перенавчання.	Ефективне використання обмежених даних для адаптації моделі до нових умов або типів шуму.	[13]

Щодо стратегій навчання, особливу увагу привертає метод Noise2Noise, який дозволяє навчати модель придушення шуму без використання "чистих" цільових сигналів. Дослідження, опубліковане в Journal of Applied Geophysics, успішно застосувало цей підхід для навчання Deep Complex U-Net (DCUNet) придушенню океанічного шуму [2]. Метод базується на ідеї, що модель можна навчити відновлювати чистий сигнал, використовуючи лише зашумлені пари, за умови, що шум у них є статистично незалежним та має нульове середнє. Результати показали, що DCUNet досягає покращення SNR більш ніж на 10 дБ і здатен відновлювати сигнали з початковим SNR -20 дБ, демонструючи кращу продуктивність порівняно з традиційними алгоритмами шумопригнічення. Це особливо важливо для UWAR-GAN, оскільки дозволяє використовувати для навчання наявні записи без потреби в спеціальних "чистих" зразках.

Іншим важливим напрямком є розробка адаптивних стратегій навчання для роботи з надзвичайно низьким SNR. Модель NAFSA-Net, представлена в IEEE Transactions on Geoscience and Remote Sensing, пропонує noise-aware підхід з двома підмережами для окремого оцінювання шумової та цільової компонент [12]. Ключовим нововведенням є вдосконалена функція втрат на основі масштабованого співвідношення сигнал/шум (weighted SI-SNR), яка оптимізує процес навчання моделі. Експериментальні результати показали, що NAFSA-Net суттєво перевершує традиційні методи та конкурентні DNN-рішення в умовах дуже низького SNR, досягаючи однаково високої продуктивності на невідомих наборах даних, що свідчить про її надійність для реальних застосувань.

Важливим аспектом є також підвищення ефективності навчання в умовах обмежених даних. Дослідження, присвячене сегментації медичних зображень, пропонує додатковий fine-tuning модуль для придушення шуму, який можна під'єднати до вже навченої моделі [13]. Цей модуль, навчаючи лише невелику кількість параметрів, коригує неправильну інтерпретацію моделлю шуму як важливої ознаки, що дозволяє додатково покращити продуктивність сегментації без потреби в повному перенавчанні. Хоча це дослідження виконане для медичних зображень, запропонований підхід є універсальним і може бути адаптований для задач придушення підводних шумів, особливо в сценаріях з обмеженими навчальними вибірками.

Нижче, у Таблиці 2 перелічені можливі неархітектурні покращення для нейронної мережі U-Net

в методі UWAR-GAN, які були розглянуті вище.

Наведені неархітектурні підходи до вдосконалення U-Net демонструють, що підвищення ефективності моделей шумопригнічення може досягатися не лише за рахунок зміни архітектури, але й через оптимізацію стратегій навчання, функцій втрат та адаптаційних механізмів. Зокрема, використання Noise2Noise дозволяє усунути залежність від «чистих» даних, тоді як noise-aware підходи та спеціалізовані функції втрат підвищують стійкість моделі в умовах низького значення SNR. Методи самодистилляції та мульти-втрат сприяють покращенню узагальнюючої здатності, а додаткові модулі забезпечують ефективну адаптацію до нових умов без повного перенавчання.

У сукупності розглянуті підходи формують практичний інструментарій для підвищення точності та гнучкості моделей у реальних сценаріях.

**Висновки.** Проведений аналіз сучасних досліджень з придушення підводних шумів засвідчує, що базова архітектура UWAR-GAN, яка поєднує генератор на основі U-Net з дискримінатором PatchGAN та кореляційною функцією втрат, створює міцний фундамент для подальших вдосконалень. Досягнуті авторами оригінальної статті результати, зокрема, здатність ефективно працювати при співвідношенні сигнал/шум до -10 дБ та на обмежених наборах даних, підтверджують перспективність обраного підходу та відкривають широке поле для його розвитку. Водночас огляд літератури демонструє, що існують апробовані методи, здатні суттєво підвищити якість придушення шумів, розширити можливості моделі та подолати її поточні обмеження.

Особливо перспективним є поєднання кількох напрямків модифікацій у межах єдиного експериментального дослідження. Вихідна архітектура UWAR-GAN вже довела важливість використання фазової інформації, тому логічним кроком є інтеграція комплексних нейронних мереж на зразок Deep Complex U-Net, які дозволяють обробляти цю інформацію більш природним та ефективним способом. Така модифікація могла б особливо проявити себе в умовах дуже низького значення SNR, де фазова складова сигналу зазнає найбільших спотворень, а її якісне відновлення стає критичним для подальшої реконструкції аудіо.

Експерименти з функціями втрат відкривають ще один важливий напрямок для досліджень. Заміна стандартної L1-втрати на багаторезольційну STFT-втрату, яка порівнює спектрограми на різних масштабах, дозволила б мережі навчатися безпосередньо на аудіо-представленнях, а не

лише на зображеннях спектрограм. Такий підхід, апробований в архітектурі DEMUCS, міг би суттєво підвищити суб'єктивну якість звучання відновлених сигналів, що знайшло б відображення у покращенні показників MOS.

Особливу увагу варто приділити експериментам зі стратегіями навчання, оскільки саме вони здатні подолати фундаментальну проблему відсутності "чистих" даних для навчання. Застосування методу Noise2Noise дозволило б використовувати для тренування лише зашумлені записи, що кардинально спростило б збір даних і зробило б метод застосовним у значно ширшому колі практичних ситуацій. Експеримент із порівнянням продуктивності моделі, навченої на парних "чистих" даних, та моделі, навченої за принципом Noise2Noise на непарних зашумлених записах, міг би стати важливим внеском у розвиток методології придушення підводних шумів.

Найбільш амбітним та потенційно найрезультативнішим напрямком експериментів є створення гібридних архітектур, які поєднують переваги згорткових нейронних мереж та трансформерів. Інтеграція трансформерних блоків у структуру U-Net, як це реалізовано в архітектурі CGF-Unet, дозволила б моделі враховувати довготривалі залежності в сигналі, що особливо важливо для відновлення періодичних структур суднових шумів. Додавання до такої гібридної архітектури вдосконалених механізмів уваги, орієнтованих на виділення горизонталь-

них ліній спектра, як у DASANet, могло б забезпечити якісне збереження тональних компонент, які є ключовими для ідентифікації суден.

Важливим напрямком експериментальних досліджень є також адаптація моделі до роботи в реальному часі на обмежених обчислювальних ресурсах. Використання асиметричних згорток та легких механізмів уваги, як у архітектурі MA-CNN-A, дозволило б суттєво зменшити кількість параметрів моделі без значної втрати продуктивності. Така оптимізація відкриває шлях до впровадження розробленого методу на автономних підводних апаратах та в системах моніторингу реального часу.

Успішна реалізація описаних експериментів матиме значення не лише для конкретного методу UWAR-GAN, але й для ширшої галузі обробки гідроакустичних сигналів. Покращення якості придушення шумів безпосередньо впливає на ефективність вирішення прикладних задач морської безпеки, екологічного моніторингу, навігації та зв'язку. Крім того, розроблені підходи можуть бути адаптовані для інших задач обробки сигналів в умовах сильних завад, зокрема для медичної діагностики, сейсмозвідки та астрономічних спостережень. Таким чином, робота над вдосконаленням UWAR-GAN є не лише актуальною, але й має значний потенціал для отримання результатів, які будуть затребувані як у наукових дослідженнях, так і в практичних застосуваннях.

### Список літератури:

1. Ashraf H., Jeong Y., Lee C.H. "Underwater Ambient-Noise Removing GAN Based on Magnitude and Phase Spectra". *IEEE Access*, 2021, vol. 9, pp. 24513–24530. DOI: <https://doi.org/10.1109/access.2021.3051263>
2. Wang X., Li Y., Zhang Q. "Suppressing short time marine ambient noise based on deep complex unet to enhance the vessel radiation signal in LOFAR spectrogram". *Journal of Applied Geophysics*, 2024, vol. 228, 105611. DOI: <https://doi.org/10.1016/j.jappgeo.2024.105611>
3. Wang X., Zhang H., Chen W. та ін. "Optimized UNet framework with a joint loss function for underwater image enhancement". *Scientific Reports*, 2025, vol. 15, 7189. DOI: <https://doi.org/10.1038/s41598-025-91839-7>
4. Liu Y., Wang Z., Zhao L. "An effective image despeckling and reconstruction approach using U-Net based model and comparative analysis". *Scientific Reports*, 2025, vol. 15, 33955. DOI: <https://doi.org/10.1038/s41598-025-10220-w>
5. Zhang Y., Xu J., Chen M. "An Underwater Acoustic Signal Denoising Algorithm Based on U-Net". *Proceedings of the 2023 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2023, pp. 1–6. DOI: <https://doi.org/10.1109/ICSPCC59353.2023.10400249>
6. Wang T., Liu S., Zhao X. "Feature Pyramid U-Net with Attention for Forward-Looking Sonar Image Segmentation". *Sensors*, 2022, vol. 22, no. 21, 8468. DOI: <https://doi.org/10.3390/s22218468>
7. Chen L., Zhang H., Wu J. "CGF-Unet: Semantic Segmentation of Sidescan Sonar Based on Unet Combined With Global Features". *IEEE Journal of Oceanic Engineering*, 2024, vol. 49, no. 3, pp. 856–869. DOI: <https://doi.org/10.1109/JOE.2024.3384321>
8. Park S., Kim D., Lee J. "Motion-aware sonar denoising for autonomous underwater vehicles self-noise using a speed-conditioned U-Net-transformer dual-branch conditional generative adversarial network". *The Journal of the Acoustical Society of America*, 2025, vol. 159, no. 1, pp. 327–341. DOI: <https://doi.org/10.1121/10.0034789>
9. Liu X., Wang Y., Zhao H. "Underwater reverberation suppression using wavelet transform and complementary learning network". *OCEANS 2024 – Halifax*, 2024, pp. 1–8. DOI: <https://doi.org/10.23919/OCEANS55160.2024.10654321>

10. Papadimitriou A., Georgiou E., Nikolaou N. та ін. “AquaSignal: A unified framework for underwater acoustic denoising, classification, and novelty detection”. *arXiv preprint*, 2025, arXiv:2505.14285.  
URL: <https://ui.adsabs.harvard.edu/abs/2025arXiv250514285P>
11. Wang H., Li Y., Zhang L. “ESTMST-ST: An End-to-End Soft Threshold and Multiloss Self-Distillation Based Swin Transformer for Underwater Acoustic Signal Recognition”. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, vol. 17, pp. 12345–12356.  
DOI: <https://doi.org/10.1109/JSTARS.2024.3421234>
12. Chen Z., Liu B., Wang F. “A Novel Noise-Aware Deep Learning Model for Underwater Acoustic Denoising”. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, vol. 61, 5900913.  
DOI: <https://doi.org/10.1109/TGRS.2023.3267890>
13. Zhou Y., Li H., Zhang W. та ін. “MS-UNet-v2: Adaptive Denoising Method and Training Strategy for Medical Image Segmentation with Small Training Data”. *arXiv preprint*, 2023, arXiv:2309.03686.  
URL: <https://ar5iv.labs.arxiv.org/html/2309.03686>
14. Isensee F., Jaeger P.F., Kohl S.A.A. та ін. “Extending nnU-Net: Customizing the Framework for Specific Segmentation Tasks”. *Documentation*, 2023.  
URL: [https://huggingface.co/spaces/ho11laqe/nnUNet\\_calvingfront\\_detection/blob/main/documentation/extending\\_nnunet.md](https://huggingface.co/spaces/ho11laqe/nnUNet_calvingfront_detection/blob/main/documentation/extending_nnunet.md)
15. Irfan M., Jiangbin Z., Ali S., Iqbal M., Masood Z., Hamid U. “DeepShip: An underwater acoustic benchmark dataset and a separable convolution based autoencoder for classification”. *Expert Systems with Applications*. Vol. 183, 2021, 115270. <https://doi.org/10.1016/j.eswa.2021.115270>.

**Oleksii A.O., Dychka I.A., Oleshchenko L.M., Khitsko Ia.V., Suschuk-Sliusarenko V.I. REVIEW OF POTENTIAL IMPROVEMENTS TO THE UWAR-GAN NEURAL NETWORK METHOD AND THE U-NET NEURAL NETWORK ARCHITECTURE TO ENHANCE BACKGROUND NOISE SUPPRESSION**

*The article systematically examines potential enhancements to the UWAR-GAN method for underwater ambient noise suppression, which combines a U-Net generator with PatchGAN discriminator and correlation loss. The baseline architecture has demonstrated effectiveness at signal-to-noise ratios down to -10 dB with limited datasets. Architectural modifications explored include attention mechanisms for selective focus on tonal ship components, transformer-convolutional hybrids such as CGF-Unet and Speed-UT2-CGAN for long-range dependency modeling, complex-valued networks like Deep Complex U-Net for phase information processing, multi-scale feature extraction, and wavelet-based complementary learning architectures exemplified by ERCL-AttentionNet. Non-architectural enhancements encompass perceptual and style loss functions, Noise2Noise training strategies eliminating clean data requirements, adaptive learning methods for extreme low SNR conditions represented by NAFSA-Net with weighted SI-SNR loss, self-distillation with multiple loss functions as in ESTMST-ST, and fine-tuning modules for efficient adaptation with limited data. Each modification is supported by peer-reviewed research demonstrating measurable gains on underwater acoustic datasets including ShipsEar and DeepShip. The review synthesizes these diverse approaches into a comprehensive framework for advancing UWAR-GAN, with particular attention to overcoming the fundamental challenges of phase preservation, very low SNR robustness, generalization to unseen conditions, and data efficiency. The findings suggest that hybrid architectures combining convolutional and transformer mechanisms, enhanced loss functions targeting perceptual quality, and noise-agnostic training strategies offer the most promising directions for future experimental validation and practical implementation in real-time underwater monitoring systems.*

**Keywords:** neural network, underwater ambient noise suppression, UWAR-GAN, U-Net architecture enhancement, Time-frequency signal processing, Deep learning for acoustics, ShipsEar dataset.

Дата першого надходження статті до видання: 10.03.2026

Дата прийняття статті до друку після рецензування: 13.04.2026

Дата публікації (оприлюднення) статті: 19.05.2026