

УДК 004.89
DOI <https://doi.org/10.32689/maup.it.2024.2.8>

Артур ОЛЕКСІЙ

аспірант кафедри інженерії програмного забезпечення в енергетиці,
Національний технічний університет України «Київський політехнічний інститут
імені Ігоря Сікорського», arturoleksii@gmail.com
ORCID: 0009-0006-5354-8098

Геннадій ПУХА

молодший науковий співробітник, Особливого конструкторського бюро «ШТОРМ»,
Національний технічний університет України «Київський політехнічний інститут
імені Ігоря Сікорського», ph8htos@gmail.com
ORCID: 0000-0001-5728-1577

СТВОРЕННЯ ДАТАСЕТУ АКУСТИЧНИХ СИГНАЛІВ ВОДНОГО СЕРЕДОВИЩА ДЛЯ ТРЕНУВАННЯ НЕЙРОМЕРЕЖІ ДЛЯ ПРИДУШЕННЯ ШУМІВ

Анотація. Аналіз акустичних сигналів у водному середовищі є складною задачею, яка ускладнюється невеликою кількістю доступних наборів даних, а нейронні мережі є актуальним і потужним інструментом для класифікації акустичних сигналів у водному середовищі. Враховуючи існуючі проблеми в цій галузі, доцільним є створення нейромережевого фреймворку, здатного працювати з акустичним шумом, в якому цільовий сигнал зашумлений фоновим шумом водного середовища, що відповідає реальним умовам. Для вирішення цієї задачі може бути використаний фреймворк з декількох нейронних мереж, які в результаті виконують задачі придушення шуму та подальшої класифікації. Можливість придушення фонового шуму дозволить підвищити точність класифікації за рахунок фільтрації спектральних складових, які не характерні для плавзасобів. Примутність артефактів, нехарактерних для цільового об'єкта, ускладнює процес класифікації, оскільки зайві характеристики водного середовища призводять до того, що нейронна мережа навчається за шаблонами, не характерними для водних об'єктів, і знижує точність класифікації. Для тестування нейронної мережі на придушення шуму потрібен набір даних з достатнім співвідношенням сигнал/шум, що відповідає реальним сигналам водного середовища. Крім того, для навчання таких нейронних мереж часто потрібні набори пар чистих і зашумлених зразків, де нейронна мережа буде пригнічувати шум від зашумлених зразків, і мати приклади чистих зразків як еталон для порівняння виконаної роботи. Процес отримання наборів даних про водне середовище шляхом запису реальних шумів є досить дорогим і складним процесом, який не гарантує задовільних результатів. Тому актуальною є задача створення шуму водного середовища із заданим співвідношенням сигнал/шум та наявністю специфічних шумів судна і фонових шумів у необхідному заданому співвідношенні.

Мета роботи – створення датасету для тренування нейромережі для придушення фонових шумів водного середовища.

Методологія. Програмне забезпечення для створення датасету та програмний код нейромережі розроблені застосуванням мови Python в середовищі Microsoft Visual Studio Code.

Наукова новизна. Було покращено підхід до створення датасету водного середовища з двох датасетів, запропоновано напрям подальшої роботи для отримання кращих результатів.

Висновки. Запропонований підхід до створення датасету показав нижче співвідношення сигналу та шуму в порівнянні з підходом, описаним у статті. Описані подальші плани стосовно розробки та покращення датасету.

Ключові слова: підводні акустичні сигнали, зашумлення, формування набору даних, нейронна мережа.

Artur OLEKSII, Hennadii PUKHA. CREATING A DATASET OF ACOUSTIC SIGNALS OF THE WATER ENVIRONMENT FOR TRAINING A NEURAL NETWORK FOR NOISE SUPPRESSION

Abstract. The analysis of acoustic signals in the water environment is a complex task, complicated by the small number of available datasets and Neural networks are a relevant and powerful tool for classifying acoustic signals in the water environment. Taking into account the current problems in this area, it is advisable to create a neural network framework that can work with acoustic noise, in which the target signal is noisy with background noise of the water environment, which corresponds to real conditions. To solve this problem, a framework of several neural networks can be used, which as a result perform the tasks of noise suppression and subsequent classification. The ability to suppress background noise will improve classification accuracy by filtering out spectral components that are not typical of watercraft. The presence of artifacts uncharacteristic of the target object complicates the classification process, as unnecessary characteristics of the water environment lead to the neural network learning patterns that are not typical for watercraft and reduce classification accuracy. Testing a neural network for noise suppression requires a dataset with a sufficient signal-to-noise ratio that corresponds to real water environment signals. Also, training such neural networks often requires sets of pairs of clean and noisy samples, where the neural network will suppress noise from the noisy samples, and have examples of clean samples as a reference for comparing the work done. The process of obtaining water environment datasets by recording real noises is a rather costly and complex process that does not guarantee satisfactory results. Therefore, the task of creating water environment noise with a given signal-to-noise ratio and the presence of specific vessel noise and background noise in the required specified ratio is relevant.

The aim of this work is to create a dataset for training a neural network to suppress background noise in the aquatic environment.

Methodology. The software for creating the dataset and the neural network code were developed using Python in the Microsoft Visual Studio Code environment.

Scientific novelty. The approach to creating an aquatic environment dataset from two datasets was improved, and a direction for further work to achieve better results was proposed.

Conclusions. The proposed approach to dataset creation showed a lower signal-to-noise ratio compared to the approach described in the article. Future plans for the development and improvement of the dataset are described.

Key words: underwater acoustic signals, noise suppression, dataset formation, neural network.

Вступ. Постановка проблеми. Через складність підводного середовища, характеристик морських об'єктів та обмежень, що накладаються обладнанням, ефективність виявлення з точки зору швидкості, точності та надійності може різко погіршитися при використанні традиційних підходів. Встановлено, що глибоке навчання має значний вплив на різні сфери застосування [4], в тому числі для роботи даними морського середовища. Тому, можливим ефективним способом вирішення задачі покращення даних морського середовища є нейромережі. Задача придушення шумів є доволі складною задачею, що потребує достатньої кількості потрібних даних. До того ж, датасети для придушення шумів, у випадку застосування нейромереж GAN потребують датасетів достовірних та неправильних зразків. Такі датасети не мають широкого розповсюдження та потребують додаткових зусиль для їх створення. Отже, метою цієї роботи є огляд підходів до створення датасетів водного середовища, вибору найбільш оптимального підходу, створення датасету та перевірка результатів його застосування для тренування нейромережі.

Аналіз досліджень і публікацій. Є різні шляхи створення, обробки та застосування датасетів водного середовища. Далі зроблено огляд датасетів водного середовища, які були застосовані для класифікації та придушення шумів. Також буде оцінена доцільність застосування цього підходу для вирішення задачі придушення фонових шумів.

У статті [6] датасет був створений шляхом збору зразків у реальних умовах за допомогою гідрофона. Такий спосіб створення датасету є доволі ресурсозатратним, займає багато часу та не завжди дає бажані результати. У даній роботі вдалося отримати багато зразків, проте запропонована модель не показала високих результатів при класифікації датасету, оскільки датасет виявився незбалансованим і не всі судна вдалося ефективно розподілити на класи. Для задачі придушення шумів цей датасет навряд чи підійде, оскільки він потребує присутності пар чистих і зашумлених зразків, а для такого датасету їх навряд чи вдасться отримати.

У статті [4] датасет був створений шляхом запису у штучних умовах. Було спроектовано 6 гвинтів і створено спеціальне обладнання, яке дозволяє симуляцію кавітаційного тунелю та вимірювання згенерованих шумів. Фонові шуми були записані шляхом використання втулки замість реального гвинта корабля. Перевагою такого підходу є, по суті, ручний режим створення записів, коли записуються сигнали потрібних гвинтів кораблів. Проте, у штучних умовах майже неможливо точно відтворити усі нюанси реального водного середовища. Працюючи з реальними шумами в природних умовах, алгоритми, тренувані на датасетах, записаних у штучних умовах, навряд чи покажуть хорошу точність класифікації.

У роботі [3] були використані реальні записи кораблів, які були розділені експертами на 12 класів та зашумлені білими шумами з різним SNR. Цей датасет було застосовано для класифікації сигналів водного середовища. Подібний датасет може бути корисний для тестування роботи нейромережі, але білі шуми не відповідають реальним шумам водного середовища. Тому алгоритм, що тренувався на зразках з білими шумами, не може ефективно придушити сигнали з реальними шумами водного середовища.

У роботі [4] був застосований датасет Shipsear, де були відібрані шуми кораблів з низьким рівнем фонових шумів та зашумлені відібраними фоновими шумами, що в комбінації дають потрібне співвідношення SNR. Дійсно, порівнюючи цей підхід з іншими, такий підхід передбачає використання вже створеного та перевіреного датасету, з допомогою якого тренувались інші моделі та показали високі результати. Тому за основу для створення датасету було взято цей підхід. Для вирішення проблеми придушення шумів та класифікації було обрано датасети Shipsear та Storm.

Мета статті. Висвітлення підходу до створення датасету для тренування нейромережі для придушення шумів використовуючи датасет Shipsear та Storm і представлення результатів тренування нейромережі.

Виклад основного матеріалу. Для задачі придушення шумів застосовується нейромережа UWAR-GAN. Основною архітектурою є нейромережа GAN [5], що передбачає тренування генератора та дискримінатора в змагальній манері. При змагальному тренуванні в якості дискримінатора була застосована нейромережа PatchGAN, що вперше була представлена у статті [7], що є згортковим бінарним класифікатором. Генератор представлений нейромережею U-net [10] та відповідає за фільтрацію зашумлених зразків. Нейромережа для фільтрації представлена згортковим автоенкодером з пропусковими зв'язками. В якості вхідних даних подаються магнітудні та фазові спектрограми зашумлених семплів. Після придушення шумів результат піддається оберненому швидкому перетворенню Фур'є, результатом якого є сигнал з придушеними фоновими шумами. Архітектура нейромережі фільтрації сигналів представлена на рис. 1.

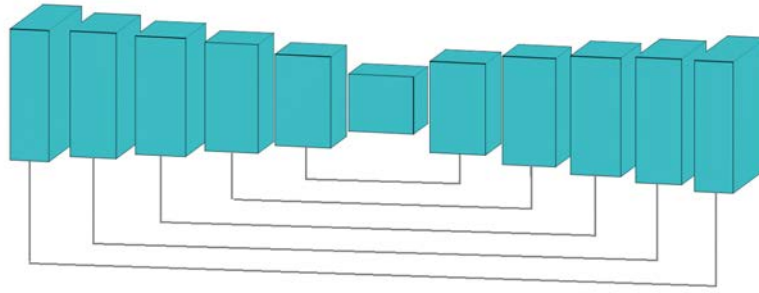


Рис. 1. Архітектура нейромережі U-net

Було застосовано два датасети водного середовища, а саме Shipsear та Storm, що був розроблений в рамках науково дослідницької роботи [1]. У датасеті Storm представлено п'ять класів кораблів, де кожен клас представлений однією годиною записаних даних. Шуми кораблів, записані в датасеті Storm, мають невелику кількість фонових шумів, а довжина файлів для кожного класу дозволяє мати значну кількість вхідних даних для тренування нейромережі. У датасеті Shipsear міститься близько 80 суден, які вже поділено на чотири класи кораблів. Разом із датасетом представлено файл, в якому детально описаний кожен файл датасету, що дозволяє мати детальну інформацію стосовно самого судна, обставин, за яких було зроблено запис, та якості самого запису. Важливою перевагою датасету Shipsear є присутність записаних фонових шумів, що зберігаються окремо від записів шумів кораблів. Всього представлено чотири класи шумів: шуми максимального потоку, шуми вітру, шуми дощу та шуми удару хвиль. Наявність окремо записаних фонових шумів дозволяє створювати датасети з присутністю різних видів шумів та комбінувати їх.

При відборі шумів з низьким рівнем шуму важливо обрати достатньо потужні зразки, в яких характерні для кораблів дискрети мали значно більшу потужність на фоні шумів, але в комбінації з фоновими шумами дали б достатньо низьке значення SNR. Всього вдалося обрати близько 2000 комбінованих зразків, чий середній SNR сягав значення -10. Але не всі зразки мали достатнє значення амплітуди, що свідчить про слабкість окремих сигналів, тому можуть мати низький SNR на фоні шуму. Тому серед сигналів, що мали низьке співвідношення сигналу та шуму, бралися сигнали з більшою амплітудою, максимальне значення якої перевищувало 0,03. Було згенеровано представлення семплів акустичних сигналів водного середовища, що містило наступні дані: спектрограму, нормалізовану потужність спектру, логарифмічне представлення спектру, амплітуду сигналу (рис. 2).

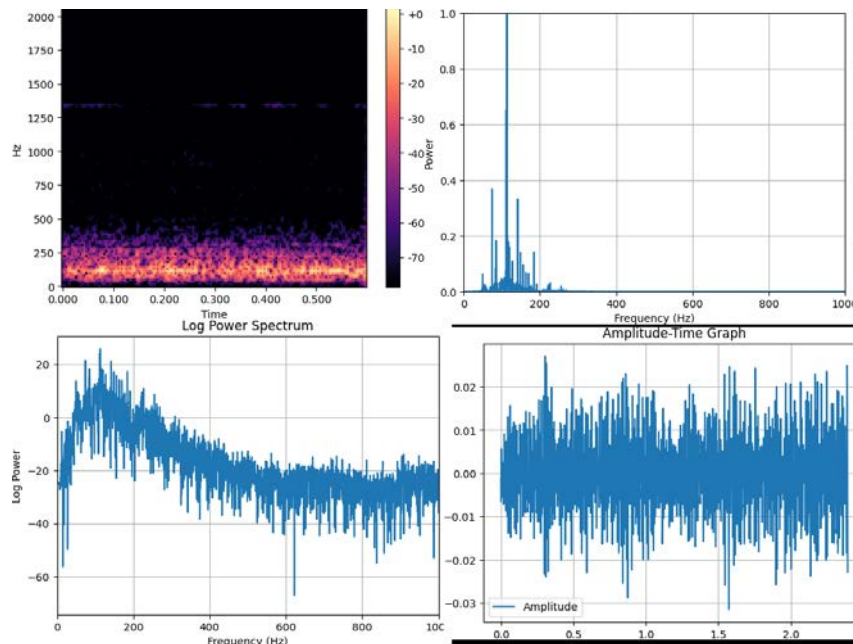


Рис. 2. Приклад представлення семпла

Для створення семплів шумів водного середовища був використаний датасет Shipsear. В ньому представлені як шуми кораблів, так і фонові шуми водного середовища. Для відповідності семплам шумів кораблів кожен аудіо-файл фонових шумів було сегментовано на семпли довжиною 2 секунди. Всього представлено 4 типи фонових шумів, а саме: шуми вітру, шуми хвиль, шуми дощу та шуми максимального потоку. Оскільки такої кількості шумів виявилось недостатньо для отримання достатньої кількості семплів, було вирішено скомбінувати шуми між собою, що дозволило отримати більшу варіацію шумів та більшу кількість комбінацій шумів кораблів та фонових шумів, що мали низький рівень SNR. Для скомбінованих зразків було досягнуто наступні значення: SNR -10.349, RMSE 0.161 та SSIM 0.158 (рис. 3).

Для поєднання чистих та зашумлених зразків, а також зашумлених зразків між собою, була застосована бібліотека AudioSegment, що є частиною модуля PyDub [9]. Ця бібліотека має багато можливостей, що дають більший контроль над модифікацією зразків. Наприклад, для комбінації зразків був застосований метод overlay. Застосування цього методу дозволяє досягти синхронізації аудіо за рахунок одночасного накладання двох аудіосигналів. Комбінація за допомогою цього методу забезпечує багатшаровість, що дозволяє створити ефект, якого важко досягти за допомогою простої конкатенації (рис. 4).

Створений датасет було оброблено та отримано набір .pt файлів з магнітудною та фазовою спектрограмами та подано на тренування нейромережі. Тренування нейромережі тривало 10000 епох. Натренована нейромережа була перевірена на семплах, що не брали участь у тренуванні. Результати були порівняні з зашумленими та чистими семплами кораблів. Результати представлення зашумлених, чистих семплів та семплів з придушеними фоновими шумами показано на рис. 5.

Значення SNR вдалось зменшити до -0.158, SSIM до 0.067 а RMSE збільшилось до 0.182. Хоча співвідношення сигнал шум і вдалось зменшити у порівнянні зі статтею з оригінальним підходом, але не усі скомбіновані зразки були очищені від фонових шумів в повній мірі, а значення середньоквадратичної похибки зросло. Причиною може бути недостатня збалансованість датасету та необхідністю продовження підбору гіперпараметрів нейромережі.

Висновки. Отже, в даній статті було обґрунтовано складність задачі придушення шумів водного середовища та необхідність наявності достатньої кількості даних для роботи з ними. Також обґрунтовано актуальність формування датасетів для тренування нейромережі з придушення шумів. Було виконано огляд літератури, де розглянуто різні підходи до створення датасетів водного середовища. В огляді оцінено переваги та недоліки кожного з запропонованих підходів та обрано оптимальний. Описано основну ідею та архітектуру нейромережі для придушення шумів. Обрано та описано датасети Shipsear та Storm. Також описано процес відбору чистих семплів водного середовища,

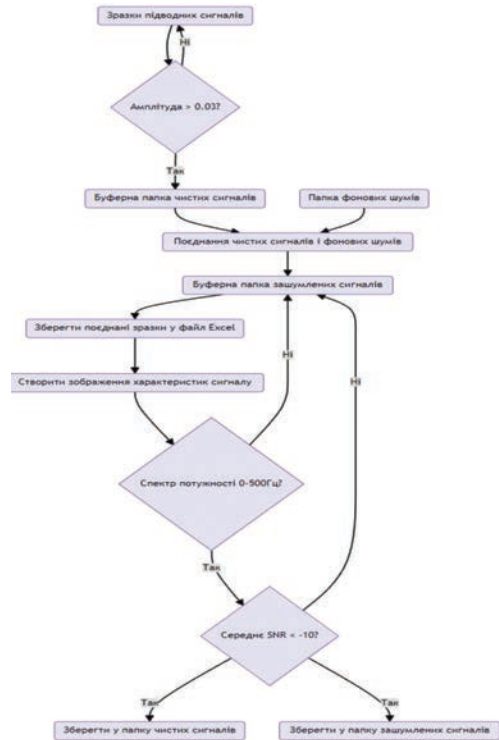


Рис. 3. Представлення роботи програмного забезпечення для створення датасету

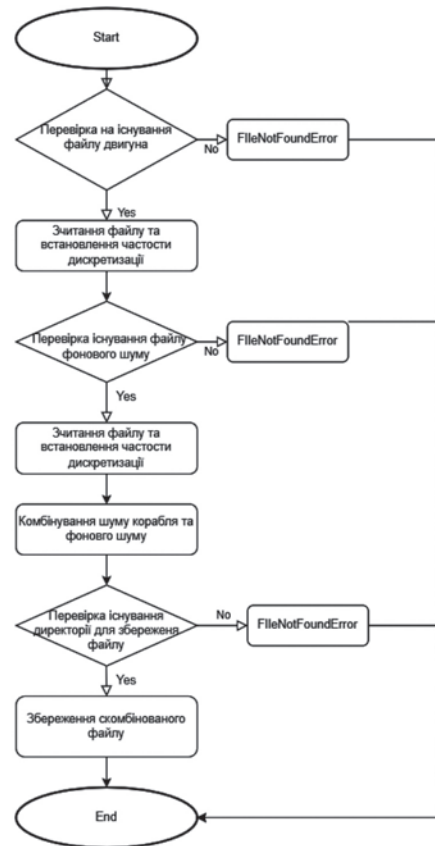


Рис. 4. Блок-схема алгоритму

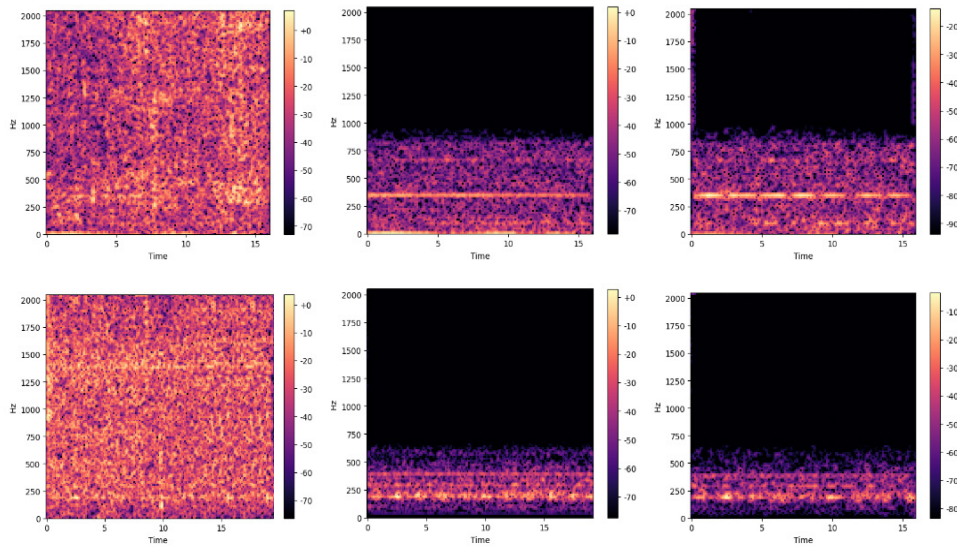


Рис. 5. Представлення зашумлених, чистих та очищених зразків

процес комбінації семплів, вказано отримані семпли, критерії оцінювання та результати застосування отриманого датасету для тренування нейромережі з придушення шумів.

Список використаних джерел:

1. Вимірювальні системи та програмне забезпечення для морських охоронних систем і дослідницьких полігонів: звіт про НДР (заключ.) НТУУ «КПІ»; кер. роб. Є. Мачуський. – К., 2012. – 104 л. + відеосюжет + CD-ROM. – Д/б №2429-п.
2. Ashraf H., Jeong Y., Lee C. H. Underwater ambient-noise removing GAN based on magnitude and phase spectra. *IEEE Access*, 2021. 9, pp.24513–24530.
3. Doan V. S., Huynh-The T., Kim D. S. Underwater acoustic target classification based on dense convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 2020. 19, pp.1–5.
4. Er M. J., Chen J., Zhang Y., Gao W. Research challenges, recent advances, and popular datasets in deep learning-based underwater marine object detection: A review. *Sensors*, 2023. 23(4), p.1990.
5. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., Generative adversarial nets. *Advances in neural information processing systems*, 27. 2014.
6. Irfan M., Jiangbin Z.H.E.N.G., 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*, 2021. 183, p.115270.
7. Isola P., Zhu J. Y., Zhou T., Efros A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017. pp. 1125–1134.
8. Khishe M., Mohammadi H. Passive sonar target classification using multi-layer perceptron trained by salp swarm algorithm. *Ocean Engineering*, 2019. 181, pp.98–108.
9. Robert J., Webbie M., others. Pydub. GitHub. 2018. Retrieved from <http://pydub.com/>.
10. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, 2015. part III 18* (pp. 234–241). Springer International Publishing.