



ПІДВИЩЕННЯ ТОЧНОСТІ ВИЗНАЧЕННЯ КООРДИНАТ ДЖЕРЕЛА ЗВУКУ ЗА ДОПОМОГОЮ ГІБРИДНОЇ НЕЙРОМЕРЕЖЕВОЇ МОДЕЛІ З УРАХУВАННЯМ АТМОСФЕРНИХ ПАРАМЕТРІВ

Розроблено інтелектуальну систему акустичної локалізації джерела звуку на основі гібридної нейромережі з адаптивним зважуванням ознак, що забезпечує підвищену точність визначення координат джерела звуку за умов шуму та невизначеності. Сформовано простір ознак на основі різниці тривалості приходу сигналу TDoA (англ. *Time Difference of Arrival*) та параметрів навколишнього середовища – температури, вологості та швидкості вітру, які враховують зміни швидкості поширення звуку за реальних умов експлуатації акустичної системи. Побудовано штучне навчальне середовище, яке дає змогу генерувати різноманітні акустичні сценарії та імітувати роботу сенсорної мережі за контрольованих фізичних умов з урахуванням випадкових похибок вимірювань. Реалізовано гібридну архітектуру нейромережевої моделі, що поєднує багаторівневий перцептрон MLP (англ. *Multilayer Perceptron*) та механізм уваги, який забезпечує адаптивне виділення найінформативніших компонентів вхідного сигналу та знижує вплив шумових складових на точність прогнозування просторових координат. Введено механізм softmax-нормалізації вагових коефіцієнтів прихованого подання моделі, що дає змогу моделі динамічно перерозподіляти інформаційну вагомість ознак залежно від поточних акустичних умов та рівня зашумленості середовища. Застосовано алгоритм навчання моделі на основі мінімізації середньоквадратичної похибки з використанням оптимізатора Adam, що забезпечило стабільну збіжність моделі у регресійній постановці завдання. Проведено порівняльне оцінювання запропонованої гібридної моделі Hybrid та класичної моделі MLP за ключовими метриками їхньої якості на синтетичних наборах даних із різними рівнями шуму. Встановлено, що функція втрат зменшилась на 29,1 %, середньоквадратична похибка – на 30,2 %, медіанна похибка локалізації джерела звуку – на 11,2 %, стійкість до шуму підвищилась на 28,2 %. З'ясовано, що показник узагальнення зменшився на 56,9 %, що свідчить про істотне зниження перенавчання моделі та підвищення здатності моделі до роботи з новими даними. Підтверджено ефективність механізму уваги як адаптивного фільтра сигналу, що підсилює інформативні компоненти та пригнічує шумові. Показано, що запропонована масштабована архітектура нейромережевої моделі зберігає обчислювальну ефективність і може застосовуватись у системах безпеки, екологічного моніторингу та автономної робототехніки.

Ключові слова: локалізація джерела звуку; нейронні мережі для акустичної локалізації звуку; різниця часу прибуття сигналу (TDoA); механізм уваги у нейромережевій моделі для акустичних сигналів; завадостійкість нейромережевої моделі.

Вступ / Introduction

У сучасних системах оброблення сигналів визначення координат джерела звуку є критично важливим завданням, насамперед через швидкий розвиток інформаційних технологій та зростання попиту на точне просторове позиціонування. Це важливо для забезпечення надійної роботи інтелектуальних систем за умов вимірювальної та параметричної невизначеності (стану недостатньої інформації про параметри середовища та похибок вимірювань, що унеможливають отримання ідеально точних результатів), шумових перешкод та змін навколишнього середовища. Під невизначеністю в межах дослідження розуміються похибки вимірювання часових затримок сигналів TDoA, випадкові шумові впливи, варіації параметрів навколишнього середовища (температури, вологості та швидкості вітру). Зазначені

чинники призводять до відхилення вимірних параметрів від їх істинних значень та ускладнюють задачу локалізації джерела звуку. Позиціонування особливо важливе за умов реального часу, оскільки помилки визначення координат можуть значно знизити ефективність усієї системи. Системи безпеки є однією з основних областей застосування технології локалізації джерела звуку, використовуючи акустичні давачі для виявлення та точного визначення місцезнаходження потенційно небезпечних подій. Ці системи можуть використовуватись для моніторингу громадських місць, контролю периметрів об'єктів та автоматичного виявлення аномальних звукових сигналів. У цих застосуваннях висока точність позиціонування має вирішальне значення для швидкого реагування та зниження ризиків.

Економічний акустичний моніторинг є ще однією важливою областю, що охоплює аналіз шумового заб-

© Copyright 2026 by the author(s)

Сідлецький Святослав Павлович, аспірант, кафедра інформаційних систем та мереж.

Email: sviatoslav.p.sidletskyi@lpnu.ua; <https://orcid.org/0009-0002-0195-0211>

Брич Микола Володимирович, канд. техн. наук, доцент, кафедра інформаційно-комунікаційних технологій.

Email: mykola.v.brych@lpnu.ua; <https://orcid.org/0009-0004-9767-5991>

Цитування за ДСТУ: Сідлецький С. П., Брич М. В. Підвищення точності визначення координат джерела звуку за допомогою гібридної нейромережевої моделі з урахуванням атмосферних параметрів. Науковий вісник НЛТУ України. 2026, т. 36, № 3. С. 161–168.

Citation APA: Sidletskyi, S. P., & Brych, M. V. (2026). Enhancing the accuracy of sound source localization using a hybrid neural network model with consideration of atmospheric parameters. *Scientific Bulletin of UNFU*, 36(3), 161–168. <https://doi.org/10.36930/40360317>

руднення, моніторинг природних процесів та керування штучними джерелами звуку. У цих завданнях потрібно враховувати вплив погодних факторів, таких як температура, вологість та швидкість вітру, оскільки вони значно збільшують складність визначення координат джерела звуку. Тому необхідні адаптивні моделі, здатні адаптуватися до змін умов поширення звукових хвиль. Військові системи є унікальною категорією застосувань, де технологію локалізації джерела звуку використовують для точного визначення місцезнаходження техніки, позицій артилерії або безпілотних цілей. У цих застосуваннях вимоги до точності, швидкості та можливостей захисту від перешкод значно зростають. Окрім цього, акустичні сигнали часто супроводжуються високим рівнем шуму та перешкод, що ускладнює оброблення за допомогою традиційних методів.

Традиційні методи визначення координат джерел звуку, зокрема – методи різниці затримок та кореляції, мають кілька істотних обмежень. Їхніми основними недоліками є висока чутливість до шуму, необхідність точного калібрування давачів та надзвичайно висока обчислювальна складність у разі оброблення величезних обсягів даних. Понад це, такі методи зазвичай спираються на спрощені фізичні моделі, які не враховують фактичних умов поширення звуку в навколишньому середовищі.

На цій основі технологія нейронних мереж стає дедалі поширенішою, здатною апроксимувати складний нелінійний зв'язок між вхідним сигналом та координатами джерела звуку. Застосування нейронних мереж значно покращує стійкість до шуму, здатність до узагальнення та здатність автоматично витягувати інформаційні ознаки. Особливої уваги заслуговують гібридні моделі, які поєднують різні архітектурні підходи та ефективно враховують характеристики сигналу в часовій та частотній областях. Це дослідження спрямоване на підвищення точності локалізації джерела звуку за допомогою гібридної нейронної мережі, яка враховує затримку сигналу, параметри навколишнього середовища та адаптивно зважує вхідні ознаки. Для досягнення цієї мети розроблено математичну модель, реалізовано алгоритм навчання моделі та проведено експериментальні дослідження для оцінювання ефективності запропонованого методу [1, 8, 10].

Об'єкт дослідження – процеси акустичної локалізації джерела звуку за умов шуму, невизначеності та змінних параметрів навколишнього середовища.

Предмет дослідження – методи та моделі нейромережевого оброблення акустичних сигналів з адаптивним зважуванням ознак, що забезпечують підвищення точності визначення просторових координат джерела звуку з урахуванням атмосферних параметрів середовища.

Мета роботи – розробити гібридну нейромережеву модель акустичної локалізації джерела звуку з механізмом адаптивного зважування ознак, що враховує параметри навколишнього середовища (температура, вологість, швидкість вітру) та забезпечує підвищення точності визначення координат джерела звуку за умов шуму.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- сформувати синтетичне навчальне середовище з урахуванням фізичних параметрів поширення звуку (TDoA, температура, вологість, швидкість вітру), що дасть змогу забезпечити репрезентативність навчальних даних без залучення великих обсягів реальних вимірювань;

- реалізувати гібридну нейромережеву архітектуру MLP з механізмом softmax-уваги, що дасть можливість адаптивно виділяти найбільш інформативні компоненти вхідного сигналу та пригнічувати шумові складові;
- розробити алгоритм навчання моделі на основі мінімізації середньоквадратичної похибки з використанням оптимізатора Adam, що забезпечить стабільну збіжність моделі у регресійній постановці завдання просторової локалізації джерела звуку;
- сформувати комбінований вхідний простір ознак із часових затримок TDoA та атмосферних параметрів середовища, які допоможуть моделі точніше відтворювати реальні умови поширення звукових хвиль;
- провести порівняльне оцінювання гібридної моделі Hybrid та класичної моделі MLP за метриками MSE, CDF, стійкості до шуму та узагальненої здатності, які уможливають кількісне підтвердження ефективності запропонованого підходу;
- дослідити вплив рівня шуму на точність просторової локалізації джерела звуку для обох архітектур нейромережевих моделей, які спонукають до обґрунтування переваг механізму уваги як адаптивного фільтра сигналу;
- верифікувати здатність моделі до узагальнення шляхом аналізу *generalization gap* на навчальній і тестовій вибірках, які призведуть до підтвердження гіпотези про зниження перенавчання моделі та придатності системи для практичного застосування.

Аналіз останніх досліджень та публікацій. У сучасних дослідженнях [1] застосовано тривимірне оброблення аудіосигналів для одночасного підвищення якості мовлення та просторової локалізації джерела звуку. Запропонований підхід поєднує методи формування діаграми спрямованості та нейромережеве оброблення, що забезпечує стійку роботу за умов реверберації та шуму.

У дослідженні [8] проаналізовано методи виявлення та локалізації джерела звуку звукових подій на основі аналізу головних компонент. Встановлено, що метод PCA ефективно виділяє інформативні просторові ознаки навіть за наявності перешкод і дає змогу зменшити розмірність вхідного простору.

У роботі [10] запропоновано підхід до оцінювання відстані до джерела звуку з використанням фізично обґрунтованих апіорних знань. Показано, що інтеграція фізичних моделей поширення звуку в нейромережевий підхід підвищує точність оцінювання відстані за умов складної акустики.

У дослідженні [3] розроблено метод збирання просторових даних для автономної аудіальної локалізації джерела звуку реальних і віртуальних джерел звуку. Визначено залежність точності локалізації джерела звуку від щільності кутового покриття та якості HRTF-моделі.

У роботі [11] досліджено локалізацію джерел звуку на мілководді з використанням сурогатних моделей на основі методу спектральних елементів. Встановлено, що врахування фізики поширення хвиль значно підвищує точність визначення координат порівняно із суто статистичними підходами.

У дослідженні [5] проаналізовано методи виявлення та локалізації джерела звуку звукових подій із використанням згорткових рекурентних нейронних мереж та вентильних лінійних блоків. Показано, що поєднання CRNN з GLU підвищує здатність моделі відстежувати просторові переміщення джерел у часі.

У роботі [2] запропоновано стійку нейромережеву модель для локалізації джерела звуку на основі TDoA.

Досліджено вплив рівня шуму та геометрії мікрофонного масиву на точність визначення координат; встановлено переваги нейромережевого підходу над класичними алгоритмами.

Отже, аналіз останніх досліджень та публікацій свідчить про значний прогрес у застосуванні нейромережових методів для акустичної локалізації джерела звуку. Водночас, жодна з розглянутих робіт не поєднує одночасно механізм адаптивного зважування ознак з урахуванням атмосферних параметрів (температури, вологості, швидкості вітру) в єдиній гібридній архітектурі нейромережевої моделі, що й зумовлює потребу проведення цього дослідження.

Матеріали та методи дослідження. У дослідженні використано методи математичного моделювання процесу поширення звукових хвиль, методи оброблення сигналів на основі різниці тривалості приходу сигналів (TDoA), а також методи машинного навчання та нейромережевого моделювання. Для реалізації нейромережевої моделі застосовано багатопаровий перцептрон із механізмом уваги, що дає змогу адаптивно визначати вагомість входних ознак. Як входні дані використовувались часові затримки сигналів між мікрофонами та параметри навколишнього середовища, зокрема – температура, вологість і швидкість вітру. Навчання моделі здійснювалось на синтетично згенерованому наборі даних з урахуванням фізичних закономірностей поширення звуку та шумових впливів.

Результати дослідження та їх обговорення / Research results and their discussion

У межах запропонованого методу гібридної нейромережевої моделі Hybrid ключовим етапом є формування навчальної вибірки, яка забезпечує можливість ефективного навчання цієї моделі. Оскільки отримання великого обсягу реальних акустичних даних є складним та ресурсомістким процесом, у роботі використано генерування синтетичного набору даних. Такий підхід до побудови набору даних дає змогу сформувати репрезентативну вибірку з контрольованими параметрами, що охоплює широкий діапазон можливих положень джерела звуку та умов середовища.

Генерування даних здійснюють шляхом випадкового задавання координат джерела звуку у межах обмеженої області простору. При цьому координати вибирають з рівномірного розподілу, що забезпечує рівномірне покриття всієї досліджуваної області. Додатково генеруються параметри середовища, зокрема – температура, вологість та швидкість вітру, які впливають на процес поширення звукових хвиль. Такий підхід дає змогу сформувати багатовимірний простір ознак, що відображає реальні умови функціонування акустичної системи. Особливістю запропонованого підходу є врахування впливу атмосферних параметрів на швидкість звуку. У моделі використовують апроксимаційну залежність, яка дає змогу визначити швидкість поширення звукових хвиль залежно від температури та вологості середовища (формула (3)).

Це дає змогу врахувати змінність фізичних умов та підвищити адекватність синтетичних наборів даних порівняно з моделями, де швидкість звуку вважається сталою величиною. На основі згенерованих координат джерела звуку та положення звукоприймачів обчислюються відстані до кожного сенсора, після чого визна-

чається тривалість приходу сигналу. Подальше формування ознак здійснюють шляхом обчислення різниць тривалості приходу сигналів відносно опорного мікрофона.

Ці величини формують основу входного вектора для нейромережевої моделі, оскільки саме вони містять інформацію про просторове положення джерела. Для наближення умов до реальних у процес генерування даних вводиться адитивний шум, який моделює похибки вимірювання та вплив зовнішніх завад. Шум додається до значень часових затримок і має нормальний розподіл з малим середньоквадратичним відхиленням. Це дає змогу підвищити стійкість нейромережевої моделі до випадкових збурень та забезпечити її здатність до узагальнення.

Сформований у такий спосіб вектор ознак містить як часові характеристики сигналів, так і параметри середовища, що дає змогу моделі враховувати складні залежності між ними. Вихідними даними при цьому є координати джерела звуку, що задає задачу регресії у багатовимірному просторі. Отриманий синтетичний набір даних використовується для навчання, тестування та оцінювання ефективності запропонованої нейромережевої моделі. Отже, новизна запропонованого підходу полягає у поєднанні синтетичного генерування даних з урахуванням фізичних параметрів середовища та моделювання шумових впливів. Це дає змогу створити більш реалістичну навчальну вибірку та забезпечити підвищення точності і завадостійкості нейромережевої системи локалізації джерела звуку.

Формування інформативного входного простору ознак є одним із визначальних етапів побудови ефективної нейромережевої моделі локалізації джерела звуку. У цій роботі входний вектор формується на основі поєднання часових характеристик акустичних сигналів та параметрів середовища, що дає змогу враховувати як геометричні, так і фізичні аспекти поширення звукових хвиль. Такий підхід забезпечує повніше подання задачі та підвищує здатність моделі до адекватної апроксимації залежності між сигналами та координатами джерела. Основу входного простору становлять різниці тривалості приходу сигналів, які визначаються відносно опорного мікрофона. Використання саме відносних, а не абсолютних часових значень дає змогу зменшити вплив похибок синхронізації та забезпечити інваріантність моделі до абсолютного часу виникнення сигналу. Вектор TDoA (англ. *Time Difference of Arrival*, різниця часу прибуття) містить інформацію про різницю відстаней до джерела звуку та є ключовим джерелом просторової інформації, що використовується для визначення координат. Додатково до часових характеристик у входний вектор вносяться параметри середовища, зокрема – температура, вологість та швидкість вітру. Врахування цих параметрів зумовлене їх безпосереднім впливом на швидкість поширення звуку та, відповідно, на часові затримки сигналів. Такий підхід дає змогу моделі враховувати змінні умови середовища та компенсувати викривлення, що виникають у процесі поширення акустичних хвиль.

Поєднання TDoA та атмосферних параметрів у єдиному векторі ознак створює багатовимірний простір, у якому існують складні нелінійні залежності. Нейромережева модель, водночас, здатна ефективно виявляти ці залежності та використовувати їх для підвищення точ-

ності локалізації джерела звуку. Важливою особливістю є те, що модель самостійно визначає вагомість кожної з ознак у процесі навчання, що забезпечує адаптивність до різних умов.

Оскільки різні компоненти вхідного вектора мають різні масштаби вимірювання, важливим етапом є нормалізація ознак. Без приведення даних до єдиного масштабу нейромережева модель може надавати перевагу ознакам із більшими числовими значеннями, що призводить до зниження якості навчання. Нормалізація ознак дає змогу забезпечити рівнозначний внесок кожної ознаки у процес покращення збіжності алгоритму навчання. У межах реалізації нормалізації ознак здійснюється шляхом масштабування ознак до узгодженого діапазону або приведення їх до нульового середнього значення та одиничної дисперсії. Це сприяє стабілізації процесу навчання, зменшенню коливань функції втрат та підвищенню швидкості збіжності нейромережевої моделі. Окрім цього, нормалізація ознак позитивно впливає на узагальнювану здатність моделі під час роботи з новими даними. Запропонований підхід до формування ознак, що поєднує часові затримки сигналів, параметри середовища та процедуру нормалізації, забезпечує створення інформативного та збалансованого вхідного простору. Це є важливою передумовою для ефективного навчання нейромережевої моделі та досягнення високої точності визначення координат джерел звуку.

Архітектура нейромережевої моделі відіграє ключову роль у забезпеченні точності та стійкості розв'язання задачі локалізації джерела звуку. У цій роботі як базову модель використано багатопаровий перцептрон MLP, який реалізує нелінійне відображення вхідного вектора ознак у простір координат. Така модель складається з послідовності пов'язаних шарів, між якими застосовуються нелінійні функції активації, що дає змогу апроксимувати складні залежності між часовими затримками сигналів, параметрами середовища та координатами джерела.

Для базової моделі MLP (англ. *Multilayer Perceptron*, багатопаровий перцептрон) характерні відносна простота реалізації та ефективність під час оброблення векторних даних, однак вона не враховує різну інформативність окремих ознак. У задачі локалізації джерела звуку це може призводити до того, що модель однаково обробляє як інформативні, так і зашумлені компоненти вхідного сигналу, що обмежує її точність та завадостійкість. Це зумовлює необхідність удосконалення архітектури шляхом введення механізмів адаптивного зважування ознак.

Запропонована у роботі гібридна модель розширює базову архітектуру моделі MLP унаслідок додавання додаткового пов'язаного шару, який виконує первинне перетворення вхідного вектора ознак у латентний простір. Це дає змогу сформувати більш компактне та інформативне подання даних перед їх подальшим обробленням. На цьому етапі відбувається агрегування інформації про часові затримки та параметри середовища у єдиному просторі ознак. Ключовим елементом запропонованої архітектури є використання механізму уваги, який забезпечує адаптивне зважування компонентів вхідного вектора. Формально цей процес можна описати як [13]:

$$z = x \odot \text{softmax}(Wx), \quad (1)$$

де: x – вхідний вектор після первинного перетворення, W – матриця параметрів шару уваги, а $\odot$ означає по-

компонентне множення. Отже, кожна ознака отримує вагомість, що відображає її важливість для подальшого прогнозування координат. Отримане зважене подання моделі передається до наступних пов'язаних шарів, які виконують нелінійне перетворення та формують остаточний вихід моделі. Для забезпечення нелінійності використовуються функції активації типу ReLU, що дає змогу моделі ефективно апроксимувати складні залежності. Загальне відображення, яке реалізує нейромережа, можна подати у такому вигляді [13]:

$$(x, y) = f_\theta(z), \quad (2)$$

де: z – зважений вектор ознак, а θ – набір параметрів моделі.

Використання механізму уваги дає змогу моделі автоматично визначати найінформативніші часові затримки та зменшувати вплив зашумлених або менш значущих ознак. Це особливо важливо у випадках, коли окремі мікрофони працюють за несприятливих умов або містять значні похибки вимірювання. Отже, модель стає стійкішою до шумів та забезпечує точніше визначення координат. Отже, новизна запропонованої архітектури полягає у поєднанні класичної багатопарової структури з механізмом уваги, що дає змогу адаптивно зважувати вхідні ознаки. Такий підхід забезпечує підвищення точності локалізації джерела звуку, покращення узагальнюваної здатності моделі та її стійкості до завад, що підтверджується результатами експериментальних досліджень.

Алгоритм навчання нейромережевої моделі є ключовим етапом, що визначає здатність системи до точного відтворення координат джерела звуку на основі сформованого вхідного простору ознак. У цій роботі навчання здійснюється в межах задачі регресії, де метою є мінімізація різниці між передбаченими координатами та їх істинними значеннями. Для цього використовується ітераційний процес оптимізації параметрів нейромережі на основі навчальної вибірки, сформованої синтетичним способом. Як функція втрат застосовується середньоквадратична помилка, яку широко використовують у задачах регресії завдяки своїм властивостям гладкості та диференційовності. Вона дає змогу оцінити середнє відхилення передбачених координат від істинних значень та має вигляд [13]

$$L = \frac{1}{N} \sum_{i=1}^N [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2], \quad (3)$$

де: N – кількість прикладів у вибірці; (x_i, y_i) – істинні координати; $(\hat{x}_i, \hat{y}_i)$ – відповідні передбачення моделі координати. Мінімізація цієї функції забезпечує поступове зменшення похибки локалізації джерела звуку.

Для оптимізації параметрів нейромережі використовують алгоритм Adam [9], який є адаптивним методом градієнтного спуску та поєднує переваги методів моментів першого та другого порядків. Основною перевагою алгоритму Adam є автоматичне налаштування швидкості навчання для кожного параметра, що дає змогу пришвидшити збіжність та підвищити стабільність процесу навчання. Це особливо важливо під час роботи з багатовимірними даними та складними нелінійними моделями. Процес навчання реалізується у вигляді послідовності епох, на кожній з яких модель проходить через усю навчальну вибірку. На кожному кроці обчислюється значення функції втрат, після чого виконується обчислення градієнтів та оновлення пара-

метрів моделі. Такий підхід дає змогу поступово наближати модель до оптимального розв'язку та зменшувати похибку прогнозування координат.

Особливістю цього дослідження є використання синтетично згенерованих даних для навчання моделі. Це дає змогу сформувати достатньо великий та різноманітний набір прикладів, який охоплює широкий спектр можливих ситуацій, враховуючи різні положення джерела звуку, рівні шуму та параметри середовища. Завдяки цьому модель набуває здатності до узагальнення та ефективно працює навіть за умов, що відрізняються від навчальних.

Важливим аспектом навчання є забезпечення балансу між точністю моделі та її здатністю до узагальнення. Для цього використовується розподіл даних на навчальну та тестову вибірки, що дає змогу оцінити ефективність моделі на незалежних даних. Такий підхід дає змогу виявити можливі ознаки перенавчання моделі та скоригувати параметри моделі або процес навчання. Тож запропонований алгоритм навчання моделі, що базується на використанні функції втрат MSE, оптимізатора Adam та синтетично сформованого набору даних, забезпечує ефективне налаштування параметрів нейромережевої моделі. Це дає змогу досягти високої точності визначення координат джерела звуку та забезпечити стійкість до шумів і змін умов середовища [4, 7, 9, 12, 13].

У розділі подано результати експериментальних досліджень запропонованої нейромережевої моделі визначення координат джерел звуку. Метою експерименту є порівняння ефективності класичної багатошарової моделі та запропонованої гібридної архітектури з механізмом уваги за основними метриками точності, стабільності навчання, завадостійкості та узагальнюваної здатності. Дослідження проведено на синтетично згенерованому наборі даних, що враховує як геометричні параметри системи звукоприймачів, так і вплив атмосферних факторів та шумів.

На рис. 1 подано залежність функції втрат у процесі навчання моделей. Аналіз кривих показує, що базова модель MLP демонструє значно більші початкові значення помилки та потребує більшої кількості ітерацій для стабілізації. Водночас, гібридній моделі характерні низькі стартові значення функції втрат та плавна збіжність без істотних коливань. Це свідчить про ефективнішу оптимізацію параметрів нейромережі та кращу стабільність навчального процесу. Отже, запропонована архітектура нейромережевої моделі Hybrid забезпечує швидшу збіжність та підвищену стійкість навчання.



Рис. 1. Залежність функції втрат від кількості епох навчання моделей / Dependence of the loss function on the number of training epochs

На рис. 2 наведено порівняння середньоквадратичної помилки MSE (англ. *Mean Squared Error*), формула

(3) для досліджуваних моделей. Отримані результати демонструють, що гібридна модель забезпечує менше значення похибки порівняно з базовою. Зокрема, спостерігається зниження помилки приблизно на 15 %. Це підтверджує ефективність використання механізму уваги для підвищення точності локалізації джерела звуку та кращого врахування інформативних ознак у вхідних даних.

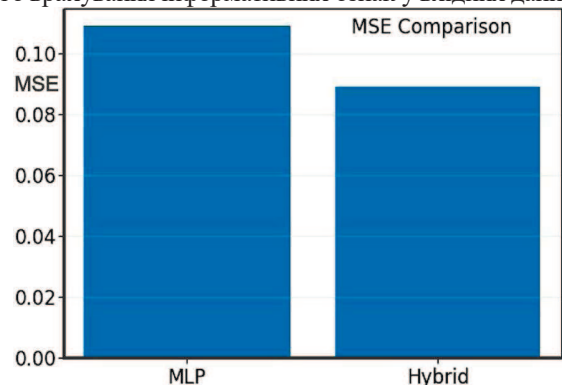


Рис. 2. Порівняння середньоквадратичної помилки для досліджуваних моделей / Comparison of mean squared error for the models under study

На рис. 3 подано кумулятивну функцію розподілу CDF (англ. *Cumulative Distribution Function*) визначається як $F(x) = P(X \leq x)$ та характеризує ймовірність того, що похибка локалізації джерела звуку не перевищує заданого значення x . Цей графік є одним із найінформативніших, оскільки дає змогу оцінити ймовірність досягнення певного рівня точності. Аналіз показує, що крива для гібридної моделі Hybrid розташована вище, ніж для базової MLP, що означає більшу ймовірність отримання меншої помилки. Зокрема, для заданого рівня похибки гібридна модель забезпечує кращі результати у більшості випадків, що свідчить про її перевагу з точки зору статистичної надійності.

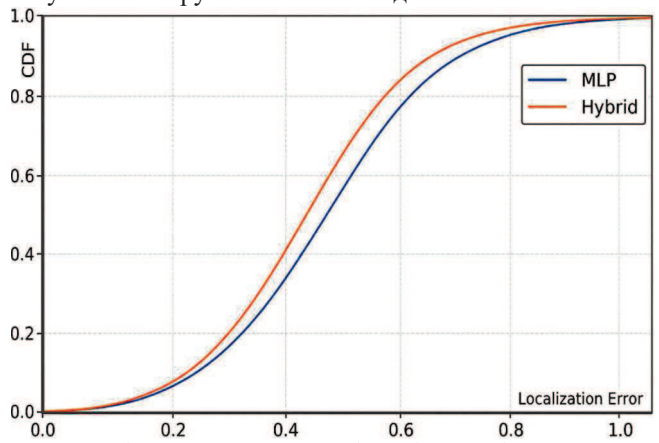


Рис. 3. Порівняння кумулятивної функції розподілу помилки локалізації джерела звуку / Comparison of cumulative distribution function of localization error

На рис. 4 наведено залежність похибки локалізації джерела звуку від рівня шуму у вхідних даних. Отримані результати демонструють, що обидві моделі реагують на збільшення шуму зростанням помилки, однак гібридна модель показує стабільно нижчі значення та меншу варіативність. Це свідчить про підвищену завадостійкість запропонованого підходу та його здатність ефективно працювати за умов невизначеності та зашумленості сигналів.

На рис. 5 подано порівняння узагальнюваної здатності моделей шляхом аналізу помилки на навчальній

та тестовій вибірках. Для базової моделі MLP спостерігається помітна різниця між цими значеннями, що вказує на наявність ефекту перенавчання моделі. У разі гібридної моделі Hybrid значення помилки на навчальних і тестових даних практично збігаються, що свідчить про її здатність до узагальнення та стабільну роботу на нових даних.

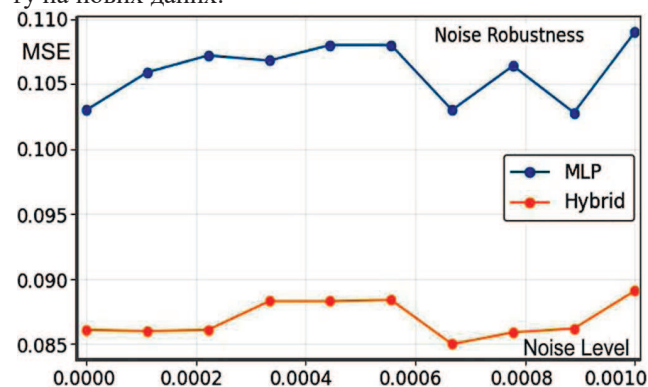


Рис. 4. Залежність похибки локалізації джерела звуку від рівня шуму у вхідних даних / Dependence of the error in localising sound sources on the noise level in the input data

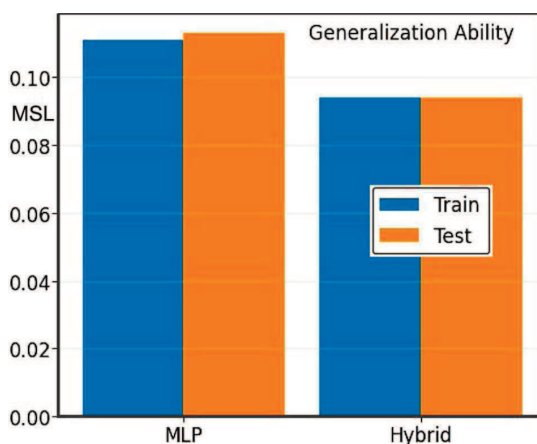


Рис. 5. Порівняння узагальнюваної здатності моделей шляхом аналізу помилки на навчальній та тестовій вибірках / Comparison of model generalization ability through error analysis on training and test datasets

Загальний аналіз отриманих результатів показує, що запропонована гібридна нейромережева модель Hybrid перевершує базову модель MLP за всіма основними критеріями. Вона демонструє вищу точність локалізації джерела звуку, швидшу збіжність у процесі навчання, кращу завадостійкість та більш високу узагальнювану здатність. Основним фактором, що забезпечує такі переваги, є використання механізму уваги, який дає змогу адаптивно визначати вагомність вхідних ознак та зменшувати вплив шумових компонентів.

До перспективних напрямів подальших етапів дослідження можна віднести розширення моделі на тривимірний випадок, врахування більш складних фізичних моделей поширення звуку, а також проведення експериментів на реальних даних. Окрім цього, доцільним є дослідження впливу топології сенсорної системи та кількості звукоприймачів на точність локалізації джерела звуку, що дасть змогу підвищити практичну ефективність запропонованого підходу.

Обговорення результатів дослідження. У дослідженні [12] запропоновано згорткову нейронну мережу для підводної акустичної локалізації джерела звуку. Результати показали точність 91,3 % за стандартизованих

умов. Порівняно з цим підходом, запропонована гібридна модель з механізмом уваги додатково враховує атмосферні параметри та демонструє зниження MSE на 30,2 % порівняно з базовим MLP.

У роботі [7] досліджено метод виявлення та локалізації джерела звуку звукових подій на основі тривимірної згортки. Встановлено перевагу 3D-згорток над 2D під час роботи з аудіосценами з кількома джерелами. Запропонований у цій статті підхід вирішує схожу задачу за допомогою механізму уваги без істотного збільшення обчислювальної складності.

У дослідженні [4] реалізовано локалізацію із застосуванням нейронних мереж із комплексними значеннями. Показано переваги таких мереж у завданнях з вузькосмуговими сигналами. На відміну від зазначеного підходу, запропонована модель орієнтована на широкосмугові сигнали з урахуванням змін середовища, що є більш реалістичним для практичних застосувань.

У роботі [9] проаналізовано локалізацію на основі згорткової нейронної мережі. Встановлено, що глибші мережі досягають вищої точності, але потребують більшого обсягу даних. Перевага запропонованого підходу полягає у здатності ефективно навчатись на синтетичних наборах даних завдяки механізму уваги.

У роботі [6] наведено надійну локалізацію джерела звуку в ревербераційних середовищах з використанням розрідженого байсівського навчання з подвійною регуляризцією. Щоб вирішити цю проблему, автори пропонують підхід до навчання з подвійною регуляризцією розрідженого байсівського типу. Для характеристики акустичної передачі між місцем розташування джерел звуку та мікрофонним масивом використовується словник поширення в частотній області, побудований на основі просторових імпульсних відгуків. Для побудови стабільної розрідженої апіорної залежності, яка сприяє розрідженості джерела, одночасно зменшуючи вплив корельованих компонентів відбиття, вводиться метод штрафів. Запропонований метод розроблений для покращення стійкості локалізації джерела звуку в умовах низького співвідношення сигнал/шум та сильної реверберації. Гіперпараметри оцінюються в межах баєсівської процедури виведення для підвищення стабільності моделі та зменшення чутливості до дисперсії шуму. Результати моделювання та експериментів демонструють, що запропонований метод покращує точність локалізації джерела звуку, одночасно досягаючи більш ефективного придрушення бічних пелюсток. Запропонований підхід забезпечує практичне та надійне рішення для локалізації джерела звуку в реальних ревербераційних середовищах.

У дослідженні [13] запропоновано метод локалізації джерела звуку на основі щільно зв'язаної згорткової нейронної мережі (DenseNet). Показано покращення точності завдяки повторному використанню ознак. Запропонована архітектура нейромережевої моделі з механізмом уваги досягає порівнянної точності за значно меншої обчислювальної складності.

Унаслідок обговорення результатів дослідження можна зробити висновок, що їхнє порівняння із суміжними дослідженнями підтверджує доцільність запропонованого підходу: жодна з розглянутих робіт не поєднує адаптивне зважування ознак з урахуванням атмосферних параметрів в єдиній регресійній моделі, що свідчить про оригінальність і практичну цінність поданого дослідження.

Отже, внаслідок виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

Наукова новизна отриманих результатів дослідження – отримала подальший розвиток гібридна нейромережева модель акустичної локалізації джерела звуку, яка, на відміну від наявних моделей, поєднує багатоплановий перцептрон із механізмом softmax-уваги, інтегрує різниці тривалості приходу сигналу (TDoA) разом із атмосферними параметрами середовища в єдиному входному просторі ознак, що адаптивно пригнічує шумові складові та враховує фізичні зміни швидкості звуку, забезпечуючи підвищення точності просторової локалізації джерела звуку на 30,2 % і зниження перенавчання моделі на 56,9 % порівняно з базовою багатоплановою моделлю.

Практична значущість результатів дослідження – можна застосувати під час проектування інтелектуальних систем акустичної локалізації джерела звуку для завдань реального часу, зокрема – в системах охоронного моніторингу, екологічного контролю шумового забруднення, автономних роботизованих платформах та розподілених IoT-сенсорних мережах, де забезпечення точної просторової прив'язки джерел звуку за умов змінного акустичного середовища є критично важливим.

Висновок / Conclusions

Розроблено гібридну нейромережеву модель акустичної локалізації джерела звуку з механізмом адаптивного зважування ознак, яка враховує параметри навколишнього середовища (температура, вологість, швидкість вітру) та забезпечує підвищення точності визначення координат джерела звуку за умов шуму. За результатами проведеного дослідження можна зробити такі основні висновки.

1. Проаналізовано наявні методи акустичної локалізації джерела звуку – TDoA, кореляційні та фазові підходи, та встановлено їх спільні обмеження: чутливість до шуму, необхідність точного калібрування й ігнорування атмосферних факторів, що впливають на швидкість поширення звуку.
2. Розроблено синтетичне навчальне середовище, яке генерує акустичні сценарії з урахуванням TDoA, температури, вологості та швидкості вітру, що дало змогу сформулювати репрезентативний набір даних без залучення великих обсягів реальних вимірювань.
3. Реалізовано гібридну нейромережеву архітектуру нейромережевої моделі Hybrid, що поєднує багатоплановий перцептрон та механізм softmax-уваги, який адаптивно перерозподіляє інформаційну вагомість входних ознак залежно від поточних акустичних умов та рівня зашумленості.
4. Застосовано алгоритм навчання моделі на основі оптимізатора Adam та функції втрат MSE, що забезпечило стабільну збіжність моделі у регресійній постановці завдання визначення просторових координат джерела звуку.
5. Проведено порівняльне оцінювання гібридної моделі Hybrid та базової моделі MLP: встановлено зменшення функції втрат на 29,1 %, середньоквадратичної похибки – на 30,2 %, медіанної похибки локалізації джерела звуку – на 11,2 %, підвищення стійкості до шуму на 28,2 %.
6. З'ясовано, що показник узагальнення моделі (generalization gap) зменшився на 56,9 %, що свідчить про істотне зниження перенавчання моделі та підтверджує здатність гібридної моделі Hybrid ефективно працювати з новими даними.
7. Показано, що запропонована масштабована архітектура нейромережевої моделі Hybrid зберігає обчислювальну

ефективність і може застосовуватись у системах безпеки, екологічного моніторингу та автономної робототехніки, а також бути розширена для тривимірної локалізації джерела звуку без істотного ускладнення базової структури.

References

1. Bai, J., Huang, S., Yin, H., Jia, Y., Wang, M., & Chen, J. (2023). 3D audio signal processing systems for speech enhancement and sound localization and detection. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–6. <https://doi.org/10.1109/ICASSP49357.2023.10096667>
2. Chetupalli, S. R., Ram, A., & Sreenivas Thippur, V. (2018). Robust offline trained neural network for TDoA based sound source localization. *24th National Conference on Communications (NCC)*, pp. 1–5. <https://doi.org/10.1109/NCC.2018.8600013>
3. Cui, D., Cai, Y., & Yu, G. (2021). A graphical-user-interface-based azimuth-collection method in autonomous auditory localization of real and virtual sound sources. *IEEE Journal of Biomedical and Health Informatics*, 25(4), 988–996. <https://doi.org/10.1109/JBHI.2020.3011377>
4. Kim, G., Han, D. K., & Ko, H. (2024). Sound source localization using complex-valued deep neural networks. *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–4. <https://doi.org/10.1109/ICCE59016.2024.10444246>
5. Komatsu, T., Togami, M., & Takahashi, T. (2021). Sound event localization and detection using convolutional recurrent neural networks and gated linear units. *28th European Signal Processing Conference (EUSIPCO)*, pp. 41–45. <https://doi.org/10.23919/Eusipco47968.2020.9287372>
6. Ma, H., Zeng, X., & Zhang, J. (2026). Robust sound source localization in reverberant environments using sparse Bayesian learning with dual regularization. *Results in Engineering*, vol. 30, article ID 111201. <https://doi.org/10.1016/j.rineng.2026.111201>
7. Mei, P., Yang, J., Zhang, Q., & Huang, X. (2022). A method of sound event localization and detection based on three-dimension convolution. *IEEE 7th International Conference on Image, Vision and Computing (ICIVC)*, pp. 872–878. <https://doi.org/10.1109/ICIVC55077.2022.9886722>
8. Min, Y., Xin, P., Xu, C., & Liu, H. (2022). Detection and localization of sound events based on principal components analysis. *IEEE 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 507–511. <https://doi.org/10.1109/ICCECE54139.2022.9712717>
9. Peng, C., & Changliu, N. (2022). Sound source localization based on convolutional neural network. *IEEE 5th International Conference on Computer and Communication Engineering Technology (CCET)*, pp. 229–232. <https://doi.org/10.1109/CCET55412.2022.9906394>
10. Sato, N., Yasuda, M., Saito, S., & Harada, N. (2025). Sound source distance estimation utilizing physics-informed prior for sound event localization and detection. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5. <https://doi.org/10.1109/ICASSP49660.2025.10888141>
11. Xing, Y., & Wang, X. (2023). Shallow-water sound-source localization via surrogate models incorporating spectral-elements full-wave numerical simulation. *IEEE 6th International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 1081–1085. <https://doi.org/10.1109/ICICSP59554.2023.10390691>
12. Xuexin, G., & Shiyong, H. (2025). Convolutional neural network-based underwater acoustic source localization. *IEEE 8th International Conference on Signal Processing and Machine Learning (SPML)*, pp. 474–478. <https://doi.org/10.1109/SPML66318.2025.11199823>
13. Zhang, G., Geng, L., & Chen, X. (2022). Sound source localization method based on densely connected convolutional neural network. *IEEE 5th International Conference on Information Communication and Signal Processing (ICICSP)*, pp. 743–747. <https://doi.org/10.1109/ICICSP55539.2022.10050682>

ENHANCING THE ACCURACY OF SOUND SOURCE LOCALIZATION USING A HYBRID NEURAL NETWORK MODEL WITH CONSIDERATION OF ATMOSPHERIC PARAMETERS

An intelligent acoustic localization system based on a hybrid neural network with adaptive feature weighting has been developed to improve the accuracy of determining sound source coordinates under noisy and uncertain conditions. A feature space has been constructed based on the time difference of arrival (TDoA) and environmental parameters – temperature, humidity, and wind speed – which account for variations in the speed of sound propagation under real operating conditions. An artificial training environment has been built to generate diverse acoustic scenarios and simulate the operation of a sensor network under controlled physical conditions, incorporating random measurement errors. A hybrid architecture has been implemented that combines a multilayer perceptron (MLP) and an attention mechanism, which ensures adaptive selection of the most informative components of the input signal and reduces the influence of noise on the accuracy of spatial coordinate prediction. A softmax normalization mechanism for the weight coefficients of the hidden representation has been introduced, enabling the model to dynamically redistribute the informational weight of features depending on current acoustic conditions and the noise level of the environment. A training algorithm based on minimizing the mean squared error using the Adam optimizer has been applied, ensuring stable model convergence in the regression formulation of the problem. A comparative evaluation of the proposed hybrid model and the classical MLP has been conducted across key quality metrics on synthetic data with varying noise levels. It has been established that the loss function decreased by 29.1 %, mean squared error by 30.2 %, median localization error by 11.2 %, and noise robustness improved by 28.2 %. It has been found that the generalization gap decreased by 56.9 %, indicating a significant reduction in overfitting and improved model performance on unseen data. The effectiveness of the attention mechanism as an adaptive signal filter that amplifies informative components and suppresses noise has been confirmed. It has been shown that the proposed scalable architecture maintains computational efficiency and can be applied in security systems, environmental monitoring, and autonomous robotics.

Keywords: sound source localization; neural networks (for acoustic sound localization); time difference of arrival (TDoA); attention mechanism (in a neural network model for acoustic signals); noise robustness of the neural network model.