

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах
рукопису

ТЕМЧУР ВЛАДИСЛАВ СЕРГІЙОВИЧ

УДК 681.518.5:621.313:004.032.26

ДИСЕРТАЦІЯ
ПРЕДИКТИВНА ДІАГНОСТИКА ОБЕРТОВОГО ЕНЕРГЕТИЧНОГО
УСТАТКУВАННЯ НА ОСНОВІ НЕЙРОМЕРЕЖЕВОГО АНАЛІЗУ
ВІБРАЦІЙНИХ СИГНАЛІВ

151 «Автоматизація та комп'ютерно-інтегровані технології»

15 Автоматизація та приладобудування

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ В. С. Темчур

Науковий керівник Баган Тарас Григорович, к.т.н., доцент

Київ – 2026

АНОТАЦІЯ

Темчур В.С. Предиктивна діагностика обертового енергетичного устаткування на основі нейромережевого аналізу вібраційних сигналів – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 151 – Автоматизація та комп'ютерно-інтегровані технології (15 – Автоматизація та приладобудування). – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, 2026.

Дисертаційна робота присвячена вирішенню актуального науково-практичного завдання підвищення ефективності предиктивної діагностики електродвигунів шляхом розроблення та дослідження методів нейромережевого аналізу вібраційних сигналів для автоматизованої класифікації технічного стану та оцінки залишкового ресурсу обладнання.

У першому розділі дисертації проведено аналітичний огляд сучасних методів діагностики обертового енергетичного устаткування. Розглянуто роль електродвигунів як критичних вузлів систем автоматизації промислових об'єктів: за статистикою IEEE, до 40% незапланованих зупинок виробничого обладнання пов'язані з відмовами електричних машин. Проаналізовано три стратегії технічного обслуговування – реактивне, планово-попереджувальне та предиктивне – та обґрунтовано перевагу предиктивного підходу, що дозволяє прогнозувати стан обладнання на основі даних реального часу. Виконано порівняльний аналіз п'яти методів моніторингу стану: термографії, аналізу мастил, аналізу струму статора, акустичної емісії та вібраційного аналізу. Обґрунтовано вибір вібраційного аналізу як основного діагностичного методу завдяки його високій чутливості до широкого спектру механічних дефектів, неінвазивності вимірювань, придатності для безперервного моніторингу та наявності добре розробленої теоретичної бази у вигляді характеристичних частот дефектів. Проведено огляд існуючих досліджень у сфері машинного та глибокого навчання для вібраційної діагностики. Встановлено, що більшість робіт зосереджені на аналізі окремих типів дефектів або

використовують обмежені вибірки з одного джерела, а систематичні порівняльні дослідження методів різної складності – від класичних статистичних показників до гібридних нейромережевих архітектур – у рамках єдиної методології практично відсутні. На основі виконаного огляду сформульовано мету та завдання дисертаційного дослідження.

У другому розділі розроблено теоретичний базис системи предиктивної діагностики. Формалізовано математичний апарат трирівневого аналізу вібраційних сигналів: у часовій області визначено шість статистичних показників (RMS, куртозис, пік-фактор, асиметрія, форм-фактор, імпульс-фактор) та метод їх нормалізації для класифікації методом k найближчих сусідів; у частотній області описано дискретне перетворення Фур'є (FFT) та спектр потужності; у часово-частотній – короткочасне перетворення Фур'є (STFT) з вікновими функціями та процедурою нормалізації спектрограм. Описано архітектуру згорткових нейронних мереж (CNN) включаючи операцію двовимірної згортки, функції активації (ReLU, Softmax, сигмоїд), шари пулінгу (MaxPooling, Global Average Pooling), Batch Normalization та Dropout, а також багатомасштабну модифікацію MSCNN з паралельними гілками ядер різного розміру та операцією конкатенації. Розроблено теоретичний опис рекурентних мереж з довгою короткостроковою пам'яттю (LSTM) та механізму уваги (Attention) для зважування часових кроків послідовності.

Обґрунтовано гібридну архітектуру MSCNN-LSTM, що поєднує багатомасштабне виділення просторових ознак зі спектрограм з аналізом часових залежностей у послідовностях вимірювань на основі принципу розділення ваг у часі. Запропоновано блок міжгілкової уваги каналів (Channel Attention), що автоматично зважує внесок паралельних гілок MSCNN залежно від характеристик вхідного зразка. Описано процедуру навчання з оптимізатором Adam та механізмом збереження найкращої моделі (ModelCheckpoint), а також метрики оцінювання якості класифікації та метод візуалізації простору ознак t-SNE. Розроблено методику побудови індексу здоров'я обладнання (HI) як ймовірності нормального стану з вектора виходу класифікатора, оцінки залишкового ресурсу

(RUL) на основі двопараметричного розподілу Вейбулла з довірчим інтервалом, адаптивного визначення порогових зон НІ через квантильний аналіз та теоретичну модель прийняття рішень щодо технічного обслуговування.

У третьому розділі представлено результати практичної реалізації та експериментального дослідження розробленої системи. Експериментальна база сформована шляхом об'єднання двох публічних датасетів вібраційних сигналів – MAFAULDA (1 951 запис, частота дискретизації 50 кГц, тривалість 5 с) та IEEE Broken Rotor Bar Database (320 записів, 7,6 кГц, 18 с) – що охоплює шість типів технічного стану електродвигунів: нормальний стан, дисбаланс ротора, зміщення осей, дефект зовнішнього кільця підшипника, дефект внутрішнього кільця підшипника та пошкодження стержнів ротора. Для узгодження даних з різних вимірювальних систем виконано передискретизацію до єдиної частоти 8 кГц та дворівневу нормалізацію амплітуд (z-score сирого сигналу та min-max нормалізація спектрограм). Вібраційні записи сегментовано на фрагменти по 4 096 відліків (0,512 с), для кожного з яких сформовано двоканальну спектрограму STFT розміром $128 \times 128 \times 2$ пікселів з вікном Блекмана-Харріса (довжина 256 відліків, перекриття 87,5%). Після балансування класів методом undersampling загальний обсяг вибірки склав 2 646 зразків (441 на клас).

Класифікацію виконано послідовно методами зростаючої складності. Метод k найближчих сусідів ($k = 5$) на ручних ознаках часової області забезпечив точність 52,8%, на частотних ознаках – 73,5%, на об'єднаному векторі – 76,2%. Згорткова нейронна мережа CNN з трьома згортковими блоками (~102 000 параметрів) досягла точності 94,1% ($F1 \text{ macro} = 94,1\%$), що на 17,9% перевищує найкращий результат класичних методів. Гібридна модель MSCNN-LSTM з механізмом уваги (~129 000 параметрів), що аналізує послідовності з п'яти спектрограм, досягла точності 98,3% ($F1 \text{ macro} = 98,3\%$). Візуалізація простору ознак методом t-SNE підтвердила формування шести чітко ізольованих кластерів гібридною моделлю.

Досліджено два удосконалення гібридної моделі. Запропонований блок міжгілкової уваги каналів (Channel Attention, коефіцієнт стиснення $r = 4$, ~8 000 додаткових параметрів – менше 5% від загальної кількості) не змінив загальну

точність (Accuracy = 98,3%, F1 macro = 98,3%), проте усунув помилку змішування нормального стану зі зміщенням осей за рахунок адаптивного перерозподілу. Запропонований метод адаптивного визначення порогових зон індексу здоров'я на основі квантильного аналізу ($\alpha = 0,05$) автоматично обчислив пороги $HI_warn = 0,87$ та $HI_crit = 0,12$, що забезпечив ранішнє виявлення початку деградації порівняно з фіксованими порогоми при контрольованому рівні хибних тривог $\sim 5\%$. Оцінку залишкового ресурсу реалізовано шляхом інтеграції індексу здоров'я з двопараметричним розподілом Вейбулла: на модельному прикладі деградаційної кривої з параметрами $\beta = 3,2$ та $\eta = 185$ діб прогнозований RUL склав 38 діб з довірчим інтервалом 90%. Розроблено інтерактивний діагностичний дашборд засобами фреймворку Streamlit, що забезпечує візуалізацію вібраційних сигналів, результатів класифікації, тренду індексу здоров'я та прогнозу залишкового ресурсу в єдиному інтерфейсі.

Наукова новизна одержаних результатів полягає в наступному: вперше для задач діагностики електродвигунів розроблено гібридну нейромережеву модель, яка поєднує багатомасштабний згортковий аналіз вібраційних спектрограм з рекурентним аналізом часових послідовностей та автоматично визначає діагностичну значущість ознак різного просторового масштабу залежно від типу несправності; удосконалено методологію предиктивної діагностики електродвигунів шляхом побудови наскрізного діагностичного конвеєра, що забезпечує послідовний аналіз від класичних статистичних показників до нейромережевої класифікації з подальшою оцінкою залишкового ресурсу на єдиній експериментальній базі з шести типів технічного стану; дістали подальшого розвитку методи машинного та глибокого навчання у вібраційній діагностиці за рахунок дослідження їх ефективності на вибірках, наближених до реальних умов експлуатації обладнання; набули подальшого розвитку методи аналізу вібраційних сигналів електродвигунів шляхом інтеграції цифрової обробки сигналів та нейромережевої класифікації в єдиній системі оцінки технічного стану.

Практичне значення отриманих результатів полягає в тому, що: розроблено методику формування двоканальних спектрограм STFT вібраційних сигналів

електродвигунів для автоматизованої діагностики; розроблено та навчено нейромережеву модель MSCNN-LSTM для класифікації шести типів технічного стану з точністю 98,3%; розроблено програмний модуль оцінки залишкового ресурсу обладнання на основі індексу здоров'я та моделі Вейбулла; розроблено інтерактивний діагностичний дашборд засобами фреймворку Streamlit. Результати досліджень використовуються у діяльності ТОВ «Вентконтрол» (м. Київ), що підтверджується відповідним актом впровадження.

Ключові слова: вібраційний аналіз, електродвигун, предиктивна діагностика, модельно-прогнозуюче керування, нейронна мережа, машинне навчання, глибоке навчання, класифікація несправностей, цифрова обробка сигналів, індекс здоров'я, залишковий ресурс, діагностика, технічне обслуговування.

Список публікацій, в яких опубліковано основні наукові результати

1. Темчур В.С., Баган Т.Г. Методи глибокого навчання моделей для прогнозного обслуговування. Вчені записки Таврійського національного університету імені В.І. Вернадського. 2023. Том 34 (73), № 6, 155-162. <https://doi.org/10.32782/2663-5941/2023.6/23>, ISSN: 2663-5941 (Print), 2663-595X (Online)

Науковим керівником Баганом Т.Г. виконано постановку завдання дослідження, визначення напрямку огляду та перевірку коректності викладеного матеріалу.

Здобувачем Темчуром В.С. проведено огляд та порівняльний аналіз архітектур глибокого навчання для задач прогнозного технічного обслуговування, систематизовано етапи побудови діагностичних систем на основі даних.

2. Темчур В.С., Баган Т.Г. Предиктивна діагностика вібраційного сигналу електродвигуна на основі технології MSCNN-LSTM. Збірник наукових праць Національного університету кораблебудування імені адмірала Макарова. 2025. №1 (499), 107-112. [https://doi.org/10.15589/znp2025.1\(499\).15](https://doi.org/10.15589/znp2025.1(499).15), ISSN 2311-3405 (Print), ISSN 2313-0415 (Online)

Науковим керівником Баганом Т.Г. запропоновано концепцію поєднання багатомасштабної згорткової мережі з рекурентною мережею для аналізу

вібраційних сигналів та визначено методiku оцінки результатів.

Здобувачем Темчуром В.С. розроблено архітектуру нейромережевої моделі, виконано попередню обробку вібраційних сигналів, навчання та тестування моделі, проведено аналіз та візуалізацію результатів.

3. Bahan, T, Temchur, V, Boun, V. (2025). Determining the efficiency of vibration signal processing methods for predictive diagnostics of electric motors. Eastern-European Journal of Enterprise Technologies, 133 (1), 6-16. (Scopus) <https://doi.org/10.15587/1729-4061.2025.320425>, ISSN 1729-3774, E- ISSN 1729-4061, ISSN-L 1729-3774

Баганом Т.Г. запропоновано методологію послідовного аналізу вібраційних сигналів та визначено структуру обчислювального експерименту.

Здобувачем Темчуром В.С. виконано обробку вібраційних сигналів у часовій та частотній областях, розроблено та навчено нейромережеву модель, проведено аналіз результатів класифікації.

Співавтором Бунем В.П. виконано підготовку експериментальних даних та верифікацію результатів спектрального аналізу.

Список публікацій, які засвідчують апробацію результатів

4. Предиктивне обслуговування промислового обладнання з використанням нейромережевого підходу / Т.Г. Баган, В.С. Темчур // Сучасні проблеми наукового забезпечення енергетики : XX міжнар. наук.-прак. конф. молодих вчених та студентів, 25-28 квітня 2023 р. : мат.конф. – К.: КПІ ім. Ігоря Сікорського, 2023. – Т.2. – С. 82-83. https://iate.kpi.ua/uploads/p_21_72711255.pdf, ISBN 978-966-990-025-8(Заг.) ISBN 978-966-990-027-2(Т. 2)

Науковим керівником Баганом Т.Г. запропоновано підхід до прогнозування технічного стану промислового обладнання на основі рекурентних нейронних мереж та визначено напрям дослідження.

Здобувачем Темчуром В.С. виконано аналіз сучасних методів виявлення аномалій, досліджено архітектуру мереж LSTM для прогнозування багатовимірних часових рядів та підготовлено текст тез.

5. Методи вібраційного аналізу для прогнозного обслуговування обладнання/ Т.Г. Баган, В.С. Темчур // Сучасні проблеми наукового забезпечення енергетики : XXI міжнар. наук.-прак. конф. молодих вчених та студентів, 23-26 квітня 2024 р. : мат.конф. – К.: КПІ ім. Ігоря Сікорського, 2024. – Т.2. – С. 15-17. https://iate.kpi.ua/uploads/p_144_77938393.pdf, ISBN 978-966-990-111-8 (Заг.) ISBN 978-966-990-112-5 (Т. 1)

Науковим керівником Баганом Т.Г. запропоновано напрям дослідження вібраційного аналізу для прогнозного обслуговування промислового обладнання та визначено структуру огляду методів обробки сигналів.

Здобувачем Темчуром В.С. виконано аналіз та систематизацію методів обробки вібраційних сигналів у часовій, частотній і частотно-часовій областях, проведено їх порівняння та підготовлено текст тез.

6. Предиктивна діагностика вібраційного сигналу електродвигуна на основі технології MSCNN-LSTM / Т.Г. Баган, В.С. Темчур // Сучасні проблеми наукового забезпечення енергетики : XXII міжнар. наук.-прак. конф. молодих вчених та студентів, 22-25 квітня 2025 р. : мат.конф. – К.: КПІ ім. Ігоря Сікорського, 2025. – Т.2. – С. 26-28. iate.kpi.ua/uploads/p_164_22642045.pdf, ISBN 978-966-990-152-1 (Заг.) ISBN 978-966-990-154-5 (Т. 2)

Науковим керівником Баганом Т.Г. запропоновано концепцію гібридної нейромережевої архітектури для діагностики несправностей електродвигунів за вібраційними сигналами та визначено методiku оцінки результатів.

Здобувачем Темчуром В.С. розроблено архітектуру нейромережевої моделі, виконано сегментацію та нормалізацію вібраційних сигналів, навчання й тестування моделі та проведено аналіз отриманих результатів.

ABSTRACT

Temchur V.S. Predictive diagnostics of rotating energy equipment based on neural network analysis of vibration signals. – Qualifying scientific work as a manuscript.

Dissertation for the degree of Doctor of Philosophy in specialty 151 – Automation and Computer-Integrated Technologies (15 – Automation and Instrumentation). – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2026.

The dissertation is devoted to solving a relevant scientific and practical problem of improving the efficiency of predictive diagnostics of electric motors by developing and investigating methods of neural network analysis of vibration signals for automated classification of technical condition and estimation of remaining useful life of equipment.

The first chapter of the dissertation presents an analytical review of modern methods for diagnostics of rotating energy equipment. The role of electric motors as critical components of industrial automation systems is examined: according to IEEE statistics, up to 40% of unplanned shutdowns of production equipment are related to failures of electric machines. Three maintenance strategies – reactive, preventive, and predictive – are analyzed, and the advantage of the predictive approach, which enables forecasting of equipment condition based on real-time data, is substantiated. A comparative analysis of five condition monitoring methods is performed: thermography, lubricant analysis, motor current signature analysis, acoustic emission, and vibration analysis. The selection of vibration analysis as the primary diagnostic method is justified due to its high sensitivity to a wide range of mechanical defects, non-invasiveness of measurements, suitability for continuous monitoring, and the availability of a well-developed theoretical framework in the form of characteristic defect frequencies. A review of existing research in the field of machine learning and deep learning for vibration-based diagnostics is conducted. It is established that most works focus on analyzing individual types of defects or use limited datasets from a single source, while systematic comparative studies of methods of varying complexity – from classical statistical indicators to hybrid neural network architectures – within a unified

methodology are practically absent. Based on the review, the objectives and tasks of the dissertation research are formulated.

In the second chapter, the theoretical framework of the predictive diagnostics system is developed. The mathematical apparatus for three-level analysis of vibration signals is formalized: in the time domain, six statistical indicators (RMS, kurtosis, crest factor, skewness, shape factor, impulse factor) and a normalization method for k-nearest neighbors classification are defined; in the frequency domain, the discrete Fourier transform (FFT) and power spectrum are described; in the time-frequency domain – the short-time Fourier transform (STFT) with window functions and spectrogram normalization procedure. The architecture of convolutional neural networks (CNN) is described, including the two-dimensional convolution operation, activation functions (ReLU, Softmax, sigmoid), pooling layers (MaxPooling, Global Average Pooling), Batch Normalization, and Dropout, as well as the multi-scale modification MSCNN with parallel branches of different kernel sizes and the concatenation operation. A theoretical description of recurrent networks with long short-term memory (LSTM) and the attention mechanism for weighting temporal steps of a sequence is developed.

The hybrid MSCNN-LSTM architecture is substantiated, combining multi-scale extraction of spatial features from spectrograms with analysis of temporal dependencies in measurement sequences based on the principle of weight sharing across time. A channel attention block is proposed that automatically weights the contribution of parallel MSCNN branches depending on the characteristics of the input sample. The training procedure with the Adam optimizer and the best model checkpoint mechanism (ModelCheckpoint) is described, along with classification quality evaluation metrics and the t-SNE feature space visualization method. A methodology is developed for constructing the equipment health index (HI) as the probability of a normal state from the classifier output vector, estimating the remaining useful life (RUL) based on the two-parameter Weibull distribution with a confidence interval, adaptive determination of HI threshold zones through quantile analysis, and a theoretical model for maintenance decision-making.

The third chapter presents the results of practical implementation and experimental

investigation of the developed system. The experimental dataset is formed by combining two public vibration signal datasets – MAFAULDA (1,951 records, sampling frequency 50 kHz, duration 5 s) and IEEE Broken Rotor Bar Database (320 records, 7.6 kHz, 18 s) – covering six types of electric motor technical condition: normal state, rotor imbalance, shaft misalignment, outer bearing race defect, inner bearing race defect, and broken rotor bars. To harmonize data from different measurement systems, resampling to a unified frequency of 8 kHz and two-level amplitude normalization (z-score of raw signal and min-max normalization of spectrograms) are performed. Vibration records are segmented into fragments of 4,096 samples (0.512 s), for each of which a two-channel STFT spectrogram of $128 \times 128 \times 2$ pixels is formed using a Blackman-Harris window (length 256 samples, overlap 87.5%). After class balancing using the undersampling method, the total dataset comprises 2,646 samples (441 per class).

Classification is performed sequentially using methods of increasing complexity. The k-nearest neighbors method ($k = 5$) on manually extracted time-domain features achieved an accuracy of 52.8%, on frequency-domain features – 73.5%, and on the combined feature vector – 76.2%. A convolutional neural network CNN with three convolutional blocks ($\sim 102,000$ parameters) achieved an accuracy of 94.1% (F1 macro = 94.1%), exceeding the best classical method result by 17.9 percentage points. The hybrid MSCNN-LSTM model with the attention mechanism ($\sim 129,000$ parameters), which analyzes sequences of five spectrograms, achieved an accuracy of 98.3% (F1 macro = 98.3%). Feature space visualization using the t-SNE method confirmed the formation of six clearly isolated clusters by the hybrid model.

Two enhancements to the hybrid model are investigated. The proposed channel attention block (compression ratio $r = 4$, $\sim 8,000$ additional parameters – less than 5% of the total) did not change the overall accuracy (Accuracy = 98.3%, F1 macro = 98.3%), but eliminated the misclassification between the normal state and shaft misalignment through adaptive redistribution of channel significance. The proposed method for adaptive determination of health index threshold zones based on quantile analysis ($\alpha = 0.05$) automatically computed thresholds $HI_warn = 0.87$ and $HI_crit = 0.12$, providing earlier detection of the onset of degradation compared to fixed thresholds with a

controlled false alarm rate of approximately 5%. Remaining useful life estimation is implemented by integrating the health index with the two-parameter Weibull distribution: on a model example of a degradation curve with parameters $\beta = 3.2$ and $\eta = 185$ days, the predicted RUL was 38 days with a 90% confidence interval. An interactive diagnostic dashboard is developed using the Streamlit framework, providing visualization of vibration signals, classification results, health index trends, and remaining useful life predictions in a unified interface.

The scientific novelty of the obtained results is as follows: for the first time in electric motor diagnostics, a hybrid neural network model is developed that combines multi-scale convolutional analysis of vibration spectrograms with recurrent analysis of temporal sequences and automatically determines the diagnostic significance of features at different spatial scales depending on the fault type; the methodology of predictive diagnostics of electric motors is improved by constructing an end-to-end diagnostic pipeline that provides sequential analysis from classical statistical indicators to neural network classification with subsequent remaining useful life estimation on a unified experimental dataset of six technical condition types; methods of machine learning and deep learning in vibration-based diagnostics are further developed through investigating their effectiveness on datasets approximating real-world operating conditions of equipment; methods of vibration signal analysis of electric motors are further developed through integration of digital signal processing and neural network classification in a unified technical condition assessment system.

The practical significance of the obtained results is as follows: a methodology for forming two-channel STFT spectrograms of electric motor vibration signals for automated diagnostics is developed; the MSCNN-LSTM neural network model is developed and trained for classification of six types of technical condition with an accuracy of 98.3%; a software module for equipment remaining useful life estimation based on the health index and Weibull model is developed; an interactive diagnostic dashboard using the Streamlit framework is developed. The research results are implemented in the operations of LLC "Ventcontrol" (Kyiv), which is confirmed by the corresponding implementation certificate.

Keywords: vibration analysis, electric motor, predictive diagnostics, predictive model control, neural network, machine learning, deep learning, fault classification, digital signal processing, health index, remaining useful life, diagnostics, maintenance.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ	18
ВСТУП	19
Розділ 1. Аналітичний огляд методів діагностики обертового енергетичного устаткування	26
1.1.Сучасні тенденції у сфері автоматизації енергетичного устаткування	26
1.2.Методи моніторингу стану обертового устаткування	33
1.2.1. Термографічний контроль.....	34
1.2.2. Аналіз мастильних матеріалів	35
1.2.3. Аналіз струму статора	36
1.2.4. Акустична емісія	36
1.2.5. Вібраційний аналіз	37
1.3.Вібраційний аналіз як основний метод діагностики	39
1.3.1. Характеристичні частоти дефектів підшипників.....	41
1.3.2. Характеристичні ознаки інших типів несправностей	42
1.3.3. Аналіз вібраційних сигналів у часовій області	45
1.3.4. Аналіз вібраційних сигналів у частотній області	46
1.4.Огляд існуючих досліджень та публікацій	47
1.4.1. Класичні методи обробки вібраційних сигналів у задачах діагностики	47
1.4.2. Методи машинного навчання у задачах діагностики обертового устаткування.....	49
1.4.3. Методи глибокого навчання та нейромережеві архітектури.....	51
1.5.Висновки до розділу	53
1.5.1. Сучасний стан та тенденції у діагностиці обертового устаткування	54
1.5.2. Методи обробки вібраційних сигналів	54

1.5.3. Нейромережеві підходи до діагностики	55
1.5.4. Подальші напрями досліджень	55
Розділ 2. Методи обробки вібраційних сигналів та нейромережеві архітектури	57
2.1. Математичні основи аналізу вібраційних сигналів.....	57
2.1.1. Часова область. Статистичні показники, формування вектора ознак та базовий класифікатор	57
2.1.2. Дискретна реалізація FFT та практичні аспекти спектрального аналізу	63
2.1.3. Розширений математичний апарат STFT для формування спектрограм	67
2.1.4. Обґрунтування та вибір параметрів STFT	71
2.2. Архітектура згорткових нейронних мереж	73
2.2.1. Операція згортки та принципи побудови CNN.....	74
2.2.2. Шари пулінгу, нормалізації та регуляризації.....	77
2.2.3. Багатомасштабна CNN з паралельними гілками (MSCNN)	83
2.3. Рекурентні нейронні мережі та механізм уваги.....	87
2.3.1. Рекурентні нейронні мережі та проблема зникнення градієнта.....	87
2.3.2. Архітектура LSTM та вентильний механізм	90
2.3.3. Обробка часових послідовностей у LSTM	93
2.3.4. Механізм уваги (Attention Mechanism)	95
2.4. Гібридна архітектура MSCNN-LSTM.....	100
2.4.1. Обґрунтування гібридної архітектури	100
2.4.2. Принцип з'єднання MSCNN і LSTM	102
2.4.3. Запропонований блок міжгілкової уваги каналів	106
2.4.4. Навчання та оптимізація гіперпараметрів	109
2.4.5. Метрики оцінювання якості класифікації	111

2.5.Методи оцінки залишкового ресурсу та прийняття діагностичних рішень ...	114
2.5.1. Інтерпретація виходу нейромережі	115
2.5.2. Індекс здоров'я обладнання.....	116
2.5.3. Оцінка залишкового ресурсу на основі моделі Вейбулла	118
2.5.4. Порогові зони індексу здоров'я та адаптивне визначення порогів .	122
2.5.5. Теоретична модель прийняття рішень	125
2.6.Висновки до розділу	128
Розділ 3. Розробка та дослідження системи предиктивної діагностики	130
3.1.Загальна архітектура розробленої системи	130
3.1.1. Програмне середовище та інструментальні засоби	134
3.2.Експериментальна база дослідження.....	136
3.2.1. Датасет MAFAULDA.....	137
3.2.2. Датасет IEEE Broken Rotor Bar Database	138
3.2.3. Вибір вимірювальних осей та обґрунтування.....	139
3.2.4. Узгодження даних із різних джерел.....	140
3.2.5. Зведена характеристика експериментальної бази	142
3.3.Підготовка та попередня обробка даних	143
3.3.1. Сегментація сигналів та формування вибірки	144
3.3.2. Формування двоканальних спектрограм STFT.....	146
3.3.3. Розподіл даних та балансування класів	150
3.4.Результати аналізу вібраційних сигналів класичними методами	152
3.4.1. Аналіз у часовій області	152
3.4.2. Аналіз у частотній області (FFT).....	159
3.5.Розробка та навчання моделі CNN.....	164
3.5.1. Архітектура мережі.....	164

3.5.2.	Підбір гіперпараметрів	167
3.5.3.	Процес навчання.....	171
3.5.4.	Результати класифікації.....	172
3.6.	Розробка та навчання гібридної моделі MSCNN-LSTM	174
3.6.1.	Формування послідовностей.....	175
3.6.2.	Архітектура мережі.....	177
3.6.3.	Підбір гіперпараметрів та навчання.....	180
3.6.4.	Результати класифікації.....	182
3.6.5.	Візуалізація простору ознак (t-SNE)	185
3.7.	Удосконалення моделі.....	187
3.7.1.	Блок міжгілкової уваги каналів (Channel Attention).....	187
3.7.2.	Адаптивне визначення порогових зон індексу здоров'я	191
3.8.	Порівняльний аналіз методів.....	194
3.9.	Оцінка залишкового ресурсу обладнання	198
3.9.1.	Побудова деградаційної кривої	198
3.9.2.	Розрахунок залишкового ресурсу.....	200
3.10.	Розробка інтерфейсу системи діагностики	202
3.10.1.	Архітектура та функціональні блоки дашборду	203
3.10.2.	Блок поточного стану та прогнозу ресурсу.....	203
3.10.3.	Блок аналізу вібраційного сигналу	204
3.10.4.	Блок трендів, журналу та класифікації.....	205
3.11.	Висновки до розділу	207
	ВИСНОВКИ.....	209
	СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	212

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ

- PdM (Predictive Maintenance) – предиктивне технічне обслуговування;
- IIoT (Industrial Internet of Things) – промисловий Інтернет речей;
- RNN (Recurrent Neural Network) – рекурентна нейронна мережа;
- CNN (Convolutional Neural Network) – згорткова нейронна мережа;
- LSTM (Long Short-Term Memory) – довга короткострокова пам'ять;
- MSCNN (Multi-Scale Convolutional Neural Network) – багатомасштабна згорткова нейронна мережа;
- FFT (Fast Fourier Transform) – швидке перетворення Фур'є;
- STFT (Short-Time Fourier Transform) – короткочасне перетворення Фур'є;
- k-NN (k-Nearest Neighbors) – метод k найближчих сусідів;
- RUL (Remaining Useful Life) – залишковий ресурс;
- HI (Health Index) – індекс здоров'я обладнання;
- t-SNE (t-distributed Stochastic Neighbor Embedding) – метод стохастичного вкладення сусідів з t-розподілом.
- BPFI (Ball Pass Frequency Inner) – частота проходження тіл кочення через дефект внутрішнього кільця підшипника;
- BPFO (Ball Pass Frequency Outer) – частота проходження тіл кочення через дефект зовнішнього кільця підшипника;
- CA (Channel Attention) – увага каналів.

ВСТУП

Обґрунтування вибору теми дослідження

У сучасних умовах розвитку промисловості особливого значення набувають питання надійності, енергоефективності та безперервності роботи технологічного обладнання. Електродвигуни є одним із ключових елементів більшості промислових систем, зокрема у насосних установках, компресорах, вентиляційних системах, конвеєрних лініях та інших об'єктах безперервного виробництва. Їх відмова або некоректна робота призводить до значних економічних втрат, простоїв обладнання, зниження якості продукції, а в окремих випадках – до аварійних ситуацій та загрози безпеці персоналу [1]. За статистикою незаплановані простої, спричинені відмовами електродвигунів, становлять від 30 до 40% від загального обсягу виробничих втрат на промислових підприємствах [2].

Традиційні підходи до технічного обслуговування електродвигунів, як правило, базуються на регламентних або реактивних стратегіях, коли ремонтні роботи виконуються за наперед визначеним графіком або вже після виникнення несправності. Такий підхід не враховує реальний технічний стан обладнання, що зумовлює або передчасне обслуговування справних агрегатів, або запізніле виявлення дефектів, які вже перейшли у критичну стадію [3]. У зв'язку з цим актуальним є перехід до концепції предиктивної діагностики, що передбачає прогнозування технічного стану електродвигунів на основі аналізу експлуатаційних даних у реальному часі [4].

Серед різних типів діагностичних сигналів особливе місце займають вібраційні сигнали, оскільки вони безпосередньо відображають динамічні процеси, пов'язані з роботою обертових елементів електродвигуна. Вібраційний аналіз дозволяє виявляти широкий спектр механічних та електромеханічних дефектів, зокрема дисбаланс, зміщення, вигин вала, ослаблення кріплень, дефекти підшипників та пошкодження ротора [5]. Разом з тим, практичне застосування класичних методів вібраційної діагностики ускладнюється високою зашумленістю сигналів, залежністю результатів від режимів роботи, а також необхідністю залучення висококваліфікованих фахівців для інтерпретації спектральних

характеристик [6].

Розвиток методів машинного навчання та глибоких нейронних мереж відкрив нові можливості для автоматизації процесів діагностики технічного стану електродвигунів. Застосування згорткових нейронних мереж у поєднанні з методами частотно-часового аналізу, зокрема короточасного перетворення Фур'є, дозволило досягти високої точності класифікації дефектів за рахунок використання спектрограм як інформативного представлення вібраційних сигналів [7]. Однак такі підходи мають низку обмежень, пов'язаних із залежністю від параметрів спектрального вікна та недостатнім урахуванням часової динаміки розвитку дефекту [8].

У зв'язку з цим особливий інтерес становлять методи, що безпосередньо працюють із сирими вібраційними сигналами та здатні враховувати їхню часову динаміку. Поєднання багатомасштабних згорткових нейронних мереж із рекурентними архітектурами, зокрема мережами довгої короточасної пам'яті, створює передумови для більш повного вилучення діагностично значущих ознак і підвищення стійкості системи до змін режимів роботи [9]. Такий підхід є особливо перспективним у контексті створення систем предиктивної діагностики, орієнтованих на реальні промислові умови [10].

Більшість існуючих досліджень у цій галузі зосереджені на аналізі окремих типів дефектів або використовують обмежені експериментальні вибірки, що ускладнює узагальнення отриманих результатів [11]. Крім того, часто недостатньо уваги приділяється порівнянню різних підходів обробки вібраційних сигналів та обґрунтуванню вибору архітектури нейронної мережі з урахуванням можливості подальшого прогнозування технічного стану обладнання [12].

Таким чином, вибір теми дослідження, присвяченої розробці та аналізу методів предиктивної діагностики електродвигунів на основі вібраційних сигналів із використанням сучасних методів глибокого навчання, зумовлений такими факторами:

- Зростаючою потребою у підвищенні надійності та безпеки промислового обладнання.

- Широким застосуванням електродвигунів у критично важливих технологічних процесах.
- Обмеженими можливостями традиційних методів діагностики в умовах змінних режимів роботи та високої зашумленості сигналів.
- Перспективністю використання нейронних мереж для автоматизованого аналізу складних вібраційних даних.
- Необхідністю створення гнучких і масштабованих діагностичних систем, орієнтованих на практичне впровадження.

Отже, обрана тема дослідження є актуальною з наукового та практичного погляду й спрямована на розв'язання важливого завдання підвищення ефективності та надійності експлуатації електродвигунів шляхом впровадження інтелектуальних методів предиктивної діагностики.

Мета дослідження

Підвищення ефективності діагностики електродвигунів шляхом розроблення методів нейромережевого аналізу вібраційних сигналів для автоматизованої класифікації технічного стану та оцінки залишкового ресурсу обладнання.

Завдання дослідження

Для досягнення вищезазначеної мети необхідно вирішити наступні завдання дослідження:

1. Аналіз сучасних методів діагностики технічного стану електродвигунів та визначення їхніх обмежень в умовах змінних режимів роботи.
2. Дослідження інформативності вібраційних сигналів для виявлення різних типів дефектів електродвигунів у часовій та частотно-часовій областях.
3. Розроблення та обґрунтування методів попередньої обробки вібраційних сигналів для підвищення якості автоматизованої діагностики.
4. Розроблення та дослідження нейромережевих моделей для автоматизованої діагностики електродвигунів на основі аналізу вібраційних сигналів.
5. Порівняльний аналіз ефективності запропонованих підходів до обробки вібраційних сигналів та нейромережевої діагностики.

6. Розроблення системи оцінки технічного стану та прогнозування залишкового ресурсу обладнання з візуалізацією результатів діагностики.

Об'єкт дослідження

Вібраційні сигнали електродвигунів та їх зміни, зумовлені технічним станом обладнання.

Предмет дослідження

Методи неймережевого аналізу вібраційних сигналів електродвигунів для автоматизованої діагностики та оцінки залишкового ресурсу.

Методи дослідження

У дисертаційній роботі для дослідження процесів діагностики технічного стану електродвигунів за вібраційними сигналами застосовано методи системного аналізу, що дозволили визначити структуру досліджуваного процесу, основні етапи обробки даних та взаємозв'язки між складовими системи моніторингу і діагностики.

Аналіз вібраційних сигналів електродвигунів виконано з використанням методів цифрової обробки сигналів. Попередня обробка даних включала фільтрацію, нормалізацію та перетворення часових послідовностей з метою підвищення інформативності сигналів і зменшення впливу шумів. Для виявлення характерних особливостей сигналів застосовано аналіз часових, спектральних та спектрально-часових характеристик.

Для автоматизованої діагностики технічного стану електродвигунів використано методи машинного та глибокого навчання, орієнтовані на аналіз часових рядів вібраційних даних. На основі отриманих сигналів побудовано неймережеві моделі, що дозволяють здійснювати виявлення та класифікацію різних типів несправностей електродвигунів.

Оцінювання ефективності запропонованих методів здійснювалося шляхом проведення обчислювальних експериментів на вибірках вібраційних сигналів, що відповідають різним технічним станам електродвигунів. Для аналізу результатів застосовано методи статистичного аналізу та візуалізації багатовимірних даних, що дало змогу оцінити можливості використання розроблених підходів у задачах

предиктивної діагностики.

Наукова новизна отриманих результатів:

1. Вперше розроблено гібридну нейромережеву модель діагностики електродвигунів шляхом поєднання багатомасштабного згорткового та рекурентного аналізу вібраційних сигналів із механізмом адаптивного зважування діагностичних ознак різного масштабу.

2. Удосконалено цілісну методологію предиктивної діагностики електродвигунів, що охоплює повний цикл від обробки вібраційних сигналів через послідовне порівняння методів класифікації різної складності до оцінки залишкового ресурсу обладнання на єдиній експериментальній базі з шести типів технічного стану.

3. Дістали подальшого розвитку методи застосування машинного та глибокого навчання у задачах вібраційної діагностики електродвигунів за рахунок їх використання для аналізу вибірок вібраційних даних, наближених до реальних умов експлуатації електромеханічного обладнання.

4. Набули подальшого розвитку методи використання вібраційних сигналів у задачах технічної діагностики електродвигунів за рахунок поєднання методів цифрової обробки сигналів та машинного навчання в межах єдиного підходу до аналізу технічного стану.

Особистий внесок здобувача

Здобувач самостійно проаналізував особливості застосування методів вібраційного аналізу для технічної та предиктивної діагностики електродвигунів, а також дослідив можливості використання методів машинного та глибокого навчання для автоматизованого виявлення несправностей електродвигунів в умовах змінних режимів роботи.

Здобувач сформулював мету та завдання дослідження.

В опублікованих наукових працях зі співавторами особисто здобувачу належать такі внески:

- Розроблення підходів до збору, попередньої обробки та аналізу вібраційних сигналів електродвигунів, орієнтованих на підвищення

інформативності даних у задачах автоматизованої технічної діагностики [1-3].

- Побудова та дослідження нейромережових моделей для аналізу часових послідовностей вібраційних даних з метою виявлення та класифікації несправностей електродвигунів.
- Проведення експериментальних досліджень розроблених підходів на вибірках вібраційних сигналів, що відповідають різним технічним станам електродвигунів, та аналіз їхньої діагностичної здатності.
- Інтерпретація отриманих результатів і оцінювання можливостей практичного застосування запропонованих методів у задачах предиктивної діагностики електродвигунів.

Апробація матеріалів дисертації

Результати досліджень дисертації висвітлювалися на наступних конференціях, конгресах, симпозіумах, семінарах і школах:

- XX Міжнародній науково-практичній конференції молодих вчених та студентів «Сучасні проблеми наукового забезпечення енергетики» (25-28 квітня 2023 р., м. Київ, КПІ ім. Ігоря Сікорського), том 2, с. 82-83.
- XXI Міжнародній науково-практичній конференції молодих вчених та студентів «Сучасні проблеми наукового забезпечення енергетики» (23-26 квітня 2024 р., м. Київ, КПІ ім. Ігоря Сікорського), том 2, с. 15-17.
- XXII Міжнародній науково-практичній конференції молодих вчених та студентів «Сучасні проблеми наукового забезпечення енергетики» (22-25 квітня 2025 р., м. Київ, КПІ ім. Ігоря Сікорського), том 2, с. 26-27.

Структура дисертації

Дисертація складається із вступу, 3 розділів, висновку, списку використаних джерел та 1 додатка.

Обсяг дисертації

Загальний обсяг дисертації складає 218 сторінок, з них основний текст роботи викладено на 186 сторінках друкованого тексту, у тому числі 37 рисунків, 28 таблиць та 97 формул. У списку використаних джерел представлено 65 джерел.

Зв'язок роботи з науковими програмами, планами, темами, грантами

Дисертаційна робота виконана на кафедрі автоматизації енергетичних процесів КПІ ім. Ігоря Сікорського у відповідності до: Розпорядження Кабінету Міністрів України від 31 грудня 2024 року №1351-р «Про схвалення Стратегії цифрового розвитку інноваційної діяльності України на період до 2030 року та затвердження операційного плану заходів з її реалізації у 2025-2027 роках»; Закону України № 3534-IX, від 13.01.2024 «Про пріоритетні напрями розвитку науки і техніки».

Практичне значення отриманих результатів

Розроблено методику формування двоканальних спектрограм вібраційних сигналів електродвигунів на основі STFT, що включає процедури узгодження даних з різних вимірювальних систем та може використовуватися як стандартизований етап підготовки вхідних даних для систем автоматизованої діагностики.

Розроблено та навчено нейромережеву модель на основі гібридної архітектури MSCNN-LSTM, що забезпечує автоматичну класифікацію шести типів технічного стану електродвигунів з точністю 98,3% без необхідності ручного формування діагностичних ознак.

Розроблено програмний модуль оцінки залишкового ресурсу обладнання на основі індексу здоров'я та статистичної моделі надійності, що дозволяє прогнозувати час до критичного стану та планувати технічне обслуговування.

Розроблено інтерактивний діагностичний дашборд засобами фреймворку Streamlit, що забезпечує візуалізацію вібраційних сигналів, результатів класифікації, тренду індексу здоров'я та прогнозу залишкового ресурсу у єдиному інтерфейсі.

Результати досліджень використовуються в ТОВ "Вентконтрол" (м.Київ) що підтверджується відповідним Актом впровадження.

Розділ 1. Аналітичний огляд методів діагностики обертового енергетичного устаткування

1.1. Сучасні тенденції у сфері автоматизації енергетичного устаткування

Сучасна промислова автоматизація переживає якісний стрибок, пов'язаний із поширенням концепції Industry 4.0 – четвертої промислової революції. На відміну від попередніх етапів індустриального розвитку, що були зосереджені на механізації та електрифікації виробничих процесів, Industry 4.0 передбачає глибоку інтеграцію кіберфізичних систем, промислового Інтернету речей (IIoT), хмарних обчислень та методів штучного інтелекту безпосередньо у виробничі та інфраструктурні процеси [3]. Ця трансформація кардинально змінює вимоги до надійності та ефективності технологічного обладнання, оскільки в умовах цифровізації та глибокої взаємозалежності компонентів збій одного вузла може миттєво призвести до порушення роботи значно більшої частини виробничого ланцюга.

У центрі концепції Industry 4.0 знаходиться інтелектуальне управління виробничими активами – підхід, відомий як Asset Performance Management (APM). Його суть полягає в переході від пасивного реагування на відмови обладнання до активного прогнозування та попередження цих відмов на основі даних, що безперервно збираються з виробничих активів. Це вимагає принципово нових підходів до моніторингу стану обладнання, обробки великих масивів технологічних даних та побудови прогностичних моделей, здатних завчасно виявляти ознаки деградації устаткування [2].

Особливого значення ці тенденції набувають у контексті систем енергетичної інфраструктури – теплопостачання, водопостачання, вентиляції та кондиціонування. Ці системи функціонують у безперервному режимі та забезпечують базові потреби як промислових підприємств, так і об'єктів комунального господарства. Порушення їх роботи навіть на короткий проміжок часу може мати серйозні наслідки – від зупинки виробничих процесів до загрози комфорту і безпеці людей, особливо в осінньо-зимовий період.

Одним з найрозповсюдженіших об'єктів автоматизації в системах

енергетичної інфраструктури є індивідуальний тепловий пункт (ІТП). Такий об'єкт являє собою комплекс інженерного обладнання, в якому реалізовано декілька взаємопов'язаних контурів автоматичного регулювання. Розглянемо декілька основних контурів:

- контур регулювання температури теплоносія, що подається до системи опалення, відповідно до температурного графіку, залежно від поточної температури зовнішнього повітря;
- контур підтримання заданого тиску і витрати теплоносія у розподільній мережі;
- контур забезпечення приготування гарячої води для господарсько-побутових потреб із підтриманням необхідної температури;
- контур керування роботою вентиляційних установок, систем припливно-витяжної вентиляції та, у ряді випадків, систем кондиціонування [13].

Кожен із зазначених контурів автоматичного регулювання включає виконавчі органи – приводні механізми, що безпосередньо здійснюють керуючий вплив на технологічний процес. До таких механізмів відносяться циркуляційні насоси контуру опалення, насоси підживлення мережі, насоси гарячого водопостачання, вентилятори припливних і витяжних установок та інші. У кожному з цих механізмів саме електродвигун є первинним перетворювачем електричної енергії на механічну і визначає функціональну спроможність усього контуру. Відмова, наприклад, циркуляційного насоса опалення призводить до повного припинення циркуляції теплоносія у відповідному контурі та виходу з ладу всієї системи теплопостачання будівлі [1]. Структурну схему ІТП з виділеними контурами регулювання та позначенням приводних механізмів наведено на рисунку 1.1.

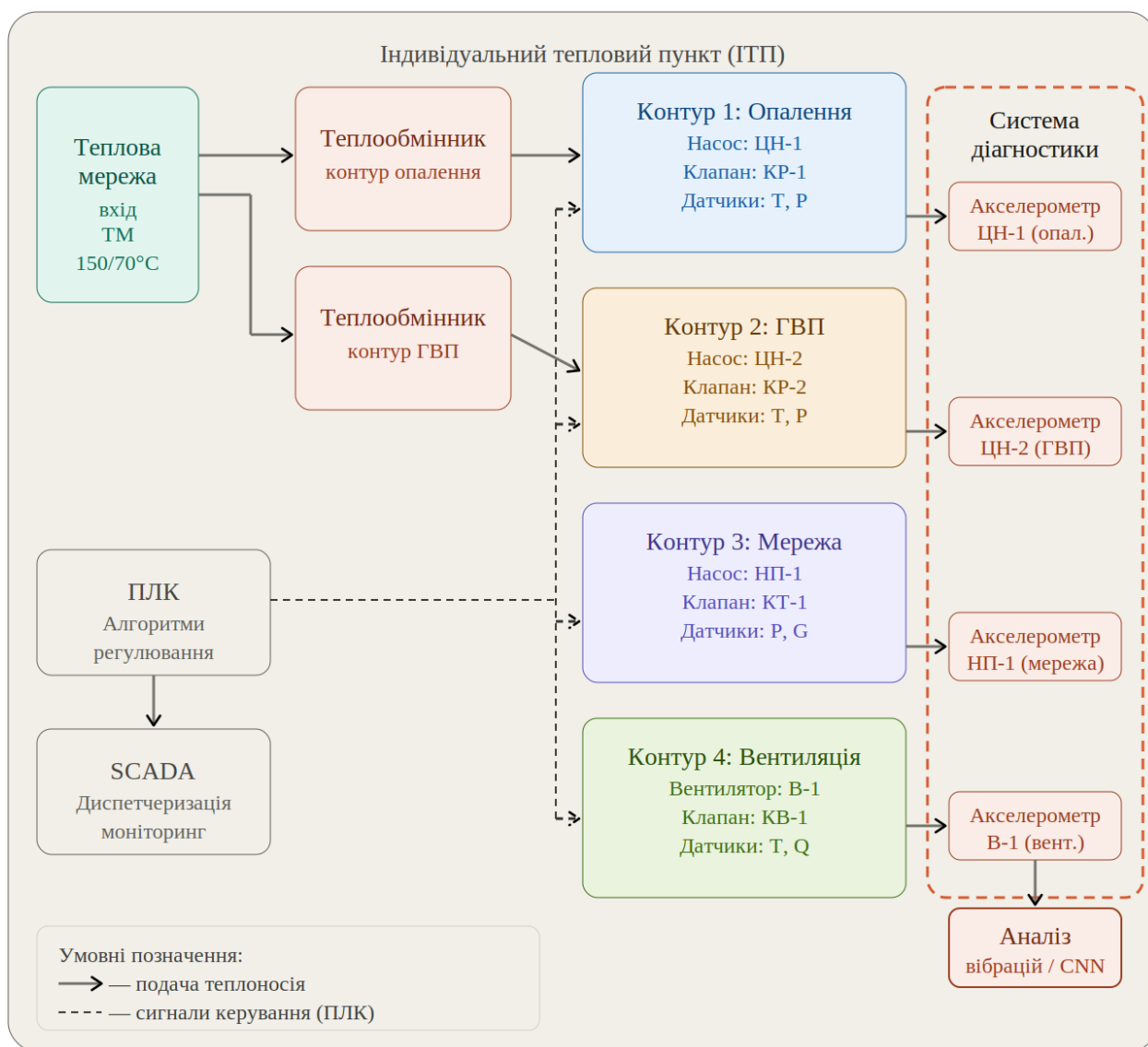


Рисунок 1. 1 Структурна схема індивідуального теплового пункту з виділеними контурами регулювання

З точки зору архітектури системи автоматичного керування, сучасний ІТП являє собою трирівневу ієрархічну систему. На нижньому (польовому) рівні розташовані датчики технологічних параметрів (температури, тиску, витрати), виконавчі механізми (клапани, насоси, вентилятори) та їх приводи. На середньому рівні функціонують програмовані логічні контролери (ПЛК), що реалізують алгоритми автоматичного регулювання. На верхньому рівні знаходиться система диспетчеризації та моніторингу (SCADA), що забезпечує збір, збереження та відображення технологічних даних, а також можливість дистанційного керування

об'єктом [3].

У цій ієрархії електродвигуни приводних механізмів є критичними вузлами нижнього рівня, стан яких безпосередньо відображається на якості виконання алгоритмів регулювання верхніх рівнів. Несправність двигуна еквівалентна втраті виконавчого органу в контурі регулювання, що робить неможливим подальше виконання керуючого алгоритму. Саме тому надійна і своєчасна діагностика електродвигунів є невід'ємною умовою коректного функціонування всієї автоматизованої системи.

Промислова статистика надає переконливі дані щодо масштабу проблеми відмов електродвигунів. За даними досліджень ІЕЕЕ, електродвигуни є одним із найбільш відмовонебезпечних класів обладнання в автоматизованих системах – близько 40% усіх незапланованих зупинок промислового обладнання безпосередньо пов'язані з несправностями електродвигунів або суміжних механічних вузлів [2]. При цьому загальна структура несправностей, за даними тієї ж статистики, виглядає наступним чином: дефекти підшипників становлять від 40 до 50% від загальної кількості відмов, несправності обмоток статора – 30–40%, дефекти ротора – 5–10%, а інші механічні несправності – решту 10–15% [1,8,14] (Рис. 1.2.).

Враховуючи, що типовий ІТП містить від 10 до 30 і більше приводних механізмів різного призначення, статистична ймовірність виникнення несправності хоча б в одному з них протягом опалювального сезону є досить значною. Саме тому питання ефективного контролю технічного стану електродвигунів набуває для таких об'єктів стратегічного значення.

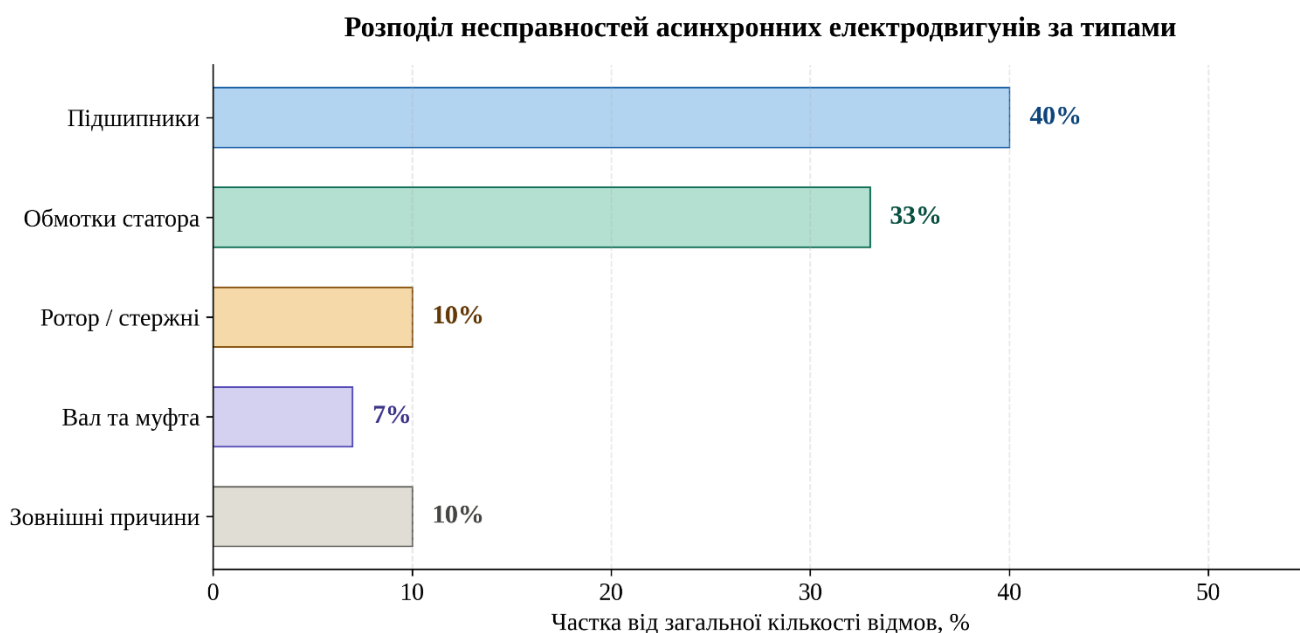


Рисунок 1. 2 Статистичний розподіл несправностей електродвигунів за типами відмов

Традиційні стратегії технічного обслуговування електродвигунів у системах автоматизації охоплюють три основні підходи. Реактивне обслуговування означає відмову від будь-яких планових заходів та виконання ремонтних робіт виключно після настання відмови. Ця стратегія є найпростішою в організаційному плані, проте призводить до найбільших економічних втрат через аварійні зупинки виробничого процесу, необхідність термінового постачання запасних частин та ризик каскадних пошкоджень суміжного обладнання внаслідок раптового відключення певних вузлів [13].

Планово-попереджувальне обслуговування передбачає виконання регламентних робіт через фіксовані часові інтервали або після досягнення встановленого напруження, незалежно від фактичного стану обладнання. Ця стратегія усуває ризик несподіваної аварійної відмови та дозволяє планувати зупинки обладнання зручним для виробничого процесу чином. Проте вона має суттєвий недолік – нераціональне використання ресурсу обладнання. За різними оцінками, від 30 до 50% планових заміन деталей і агрегатів виконуються на справному обладнанні, ресурс якого ще далеко не вичерпаний, тоді як частина обладнання може вийти з ладу між плановими обслуговуваннями [4].

Обслуговування за станом (Condition-Based Maintenance, CBM) та його найбільш досконала форма – предиктивне технічне обслуговування (Predictive Maintenance, PdM) – передбачають проведення технічних заходів на підставі результатів безперервного моніторингу фактичного технічного стану обладнання. CBM фіксує поточний стан і сигналізує про необхідність втручання при досягненні граничних значень контрольованих параметрів. PdM іде далі, прогножуючи майбутній стан обладнання та залишковий ресурс (Remaining Useful Life, RUL), що дозволяє планувати технічне обслуговування завчасно – до настання відмови, але не раніше, ніж це справді необхідно [3].

Порівняльна характеристика трьох стратегій технічного обслуговування наведена у таблиці 1.1.

Таблиця 1. 1 Порівняльна характеристика стратегій технічного обслуговування

Критерій	Реактивне	Планово-попереджувальне	За станом (CBM)	Предиктивне (PdM)
Плановість зупинок	Відсутня	Фіксований графік	За показниками	За прогнозом
Ризик аварійної відмови	Дуже високий	Низький	Низький	Мінімальний
Використання ресурсу	Повне	Неповне (30–50%)	Близьке до повного	Оптимальне
Потреба в моніторингу	Відсутня	Мінімальна	Безперервна	Безперервна+аналіз
Складність впровадження	Низька	Низька	Середня	Висока
Загальна ефективність	Низька	Середня	Висока	Дуже висока

Як видно з таблиці 1.1, предиктивне технічне обслуговування є найбільш ефективною стратегією з точки зору балансу між надійністю обладнання та витратами на технічне обслуговування. Разом із тим воно вимагає найбільш складної інфраструктури – систем безперервного збору даних, алгоритмів обробки сигналів та прогностичних моделей. Саме розробка таких алгоритмів і є предметом даного дослідження.

Впровадження PdM в системах автоматизації енергетичного устаткування дозволяє вирішити наступні завдання:

- завчасно виявляти дефекти на ранніх стадіях їх розвитку до настання критичного стану, що дозволяє уникнути аварійних зупинок та каскадних пошкоджень суміжних контурів регулювання;
- оптимально планувати технічне обслуговування без порушення технологічного процесу, використовуючи прогноз залишкового ресурсу для вибору зручного часового вікна для ремонту;
- скорочувати прямі витрати на ремонт за рахунок усунення дефектів на початковій стадії, коли обсяг ремонтних робіт та вартість замінних деталей є мінімальними;
- підвищувати загальну ефективність роботи автоматизованої системи керування за рахунок мінімізації незапланованих простоїв та підтримання обладнання в оптимальному технічному стані;
- збільшувати термін служби обладнання шляхом своєчасного усунення дефектів, що запобігає їх прогресуванню та передчасному виходу обладнання з ладу;
- накопичувати статистику несправностей та режимів роботи обладнання, що дозволяє вдосконалювати діагностичні алгоритми та підвищувати точність прогнозування залишкового ресурсу.

Реалізація концепції PdM потребує ефективних методів збору та аналізу діагностичної інформації. Серед різних типів діагностичних сигналів, що формуються в процесі роботи обертового енергетичного устаткування, особливе

місце займають вібраційні сигнали. Їх визначальна перевага полягає у тому, що вони є прямим відображенням динамічних процесів у всіх механічних вузлах машини – підшипниках, валах, роторі, з'єднувальних муфтах – і доступні для неінвазивного вимірювання за допомогою акселерометрів, що встановлюються на зовнішній поверхні корпусу без будь-якого розкриття або зупинки машини [5].

Сучасні методи цифрової обробки вібраційних сигналів у поєднанні з алгоритмами машинного навчання відкривають принципово нові можливості для автоматизованої діагностики стану приводних механізмів безпосередньо в контурі системи керування. Саме розробці та дослідженню таких методів і присвячена дана дисертаційна робота.

Таким чином, задача предиктивної діагностики обертового енергетичного устаткування на основі нейромережевого аналізу вібраційних сигналів є невід'ємною складовою сучасних інтелектуальних систем автоматизації промислових об'єктів. Її успішне розв'язання забезпечує реалізацію принципів Industry 4.0 на рівні польових пристроїв автоматики та сприяє переходу від традиційних реактивних стратегій обслуговування до інтелектуального управління виробничими активами.

1.2. Методи моніторингу стану обертового устаткування

Моніторинг стану обладнання (Condition Monitoring, CM) є технічною основою концепції предиктивного технічного обслуговування. Під цим поняттям розуміють сукупність заходів з безперервного або періодичного вимірювання фізичних параметрів машини в процесі її роботи з метою виявлення змін, що свідчать про розвиток несправності або погіршення технічного стану. Ефективність системи CM визначається вибором відповідних методів вимірювання, типів датчиків та алгоритмів обробки і класифікації отриманої інформації [1].

Вибір методу моніторингу для конкретного застосування визначається рядом факторів: типом устаткування та характером очікуваних несправностей, умовами експлуатації, вимогами до чутливості та оперативності виявлення дефектів, доступністю місць встановлення датчиків, а також економічними міркуваннями – вартістю обладнання, впровадження та обслуговування системи. У промисловій

практиці для діагностики обертового устаткування застосовується широкий спектр методів, що класифікуються за природою вимірюваних фізичних параметрів. Розглянемо найбільш поширені з них.

1.2.1. Термографічний контроль

Термографічний контроль базується на вимірюванні теплового (інфрачервоного) випромінювання поверхні обладнання за допомогою спеціалізованих тепловізійних камер. Фізичною основою методу є закон Стефана-Больцмана, згідно з яким будь-яке тіло з температурою вище абсолютного нуля випромінює електромагнітні хвилі, інтенсивність яких однозначно пов'язана з температурою поверхні. Сучасні тепловізори здатні вимірювати температурні поля у діапазоні від -40°C до $+2000^{\circ}\text{C}$ з роздільною здатністю до $0,1^{\circ}\text{C}$, формуючи кольорові теплові карти об'єкта, що обстежується [15].

Застосування термографії до діагностики електродвигунів ґрунтується на тому, що більшість несправностей супроводжується підвищенням температури уражених вузлів: дефекти підшипників збільшують тертя і, відповідно, тепловиділення; погіршення якості змащення також призводить до перегріву підшипникових вузлів; часткові пробої ізоляції обмоток спричиняють локальні теплові аномалії на поверхні корпусу двигуна. Таким чином, аналіз теплових полів дозволяє непрямо оцінити стан внутрішніх вузлів машини [15].

До переваг термографічного методу відносяться: повна безконтактність вимірювань – обстеження виконується з відстані без будь-яких маніпуляцій з обладнанням; можливість одночасного охоплення великої кількості вузлів одним знімком; наочність результатів, що робить їх доступними для інтерпретації навіть нефаховим персоналом; можливість проведення обстежень без зупинки обладнання та в складних умовах доступу. Цей метод особливо ефективний для виявлення несправностей електричного характеру – перегріву контактів, місць погіршення ізоляції, нерівномірного навантаження фаз.

Разом із тим термографія має суттєві обмеження. По-перше, чутливість методу до ранніх стадій механічних дефектів є відносно невисокою: температурне поле на поверхні корпусу реагує лише на досить значний розвиток дефекту, коли

підвищення тертя стає помітним. Дефекти підшипників, наприклад, на ранніх стадіях не спричиняють помітного нагріву і виявляються термографією лише при значному зношуванні. По-друге, результати вимірювань суттєво залежать від зовнішніх умов: температури навколишнього повітря, протягів, наявності теплових джерел поряд з об'єктом. По-третє, метод потребує прямої видимості поверхні обладнання, що не завжди забезпечується в умовах промислових об'єктів, де обладнання може бути закрите кожухами або розташоване у важкодоступних місцях [15].

1.2.2. Аналіз мастильних матеріалів

Аналіз мастильних матеріалів, або трибологічний аналіз, є потужним методом діагностики механічних несправностей, що базується на дослідженні хімічного та фізичного стану мастила або мастильного матеріалу, що використовуються у вузлах тертя машини. Під час роботи будь-якого механізму відбувається зношування поверхонь, що труться, і продукти зношування – металеві частинки – потрапляють до мастила. Аналіз концентрації, розміру, форми та елементного складу цих частинок дозволяє не лише виявити факт зношування, але й ідентифікувати його джерело та оцінити інтенсивність [13].

Для аналізу мастил застосовуються різноманітні методи: спектрометрія (атомно-емісійна або рентгенофлуоресцентна) дозволяє визначити елементний склад продуктів зношування і встановити, яка саме деталь зношується; ферографія досліджує форму та розмір феромагнітних частинок, розділяючи їх за розміром; оптична мікроскопія забезпечує візуальний аналіз частинок та їх класифікацію за морфологічними ознаками; аналіз в'язкості мастила свідчить про ступінь його деградації та необхідність заміни.

Перевагою методу є висока інформативність – він здатний виявляти деградацію підшипників кочення та зубчастих передач на ранніх стадіях, задовго до появи помітних змін у вібраційних або шумових характеристиках. Проте метод має суттєве обмеження, що робить неможливим його застосування у системах безперервного автоматизованого моніторингу: він вимагає фізичного відбору проб мастила та проведення лабораторного аналізу, що є трудомістким процесом і

передбачає значні часові затримки між виникненням дефекту та отриманням результатів. Зазвичай цей метод використовується виключно як допоміжний при проведенні планових технічних оглядів.

1.2.3. Аналіз струму статора

Аналіз струму статора двигуна (Motor Current Signature Analysis, MCSA) є методом діагностики, що ґрунтується на спектральному аналізі сигналу струму статора електродвигуна. Фізичну основу методу складає той факт, що будь-які механічні або електромагнітні несправності двигуна призводять до зміни характеру споживаного струму, який модулюється на характерних частотах, пов'язаних із типом і ступенем розвитку несправності [16].

Зокрема, поломка або тріщина у стержнях короткозамкненої обмотки ротора асинхронного двигуна проявляється у спектрі струму у вигляді бокових смуг навколо основної частоти живлення f_1 . Ці бокові смуги розташовуються на частотах $f_1(1 \pm 2ks)$, де $k = 1, 2, 3, \dots$ – порядок гармоніки, s – ковзання двигуна. Дефекти підшипників також спричиняють характерні модуляції струму, проте їх амплітуда, як правило, значно менша, ніж наприклад у вібраційних сигналах, що ускладнює виявлення дефектів на ранніх стадіях [16].

Ключовою перевагою MCSA є можливість реалізації без встановлення додаткових датчиків – сигнали струму живлення доступні у шафі автоматики. Це суттєво знижує вартість впровадження системи моніторингу, особливо на об'єктах, де вже встановлені засоби обліку електроенергії. Разом із тим MCSA має обмеження: чутливість до механічних дефектів, особливо підшипникових, є значно нижчою порівняно з вібраційним аналізом. Точність діагностики також знижується при наявності електромагнітних завад у мережі живлення та при суттєво змінному навантаженні на вал двигуна.

1.2.4. Акустична емісія

Акустична емісія (АЕ) є методом неруйнівного контролю, що реєструє пружні хвилі у діапазоні від 20 кГц до 1 МГц, що виникають безпосередньо у матеріалі конструктивних елементів при мікроструктурних перетвореннях – розвитку та зростанні тріщин, поверхневому руйнуванню, пластичній деформації,

зміні фазового стану матеріалу. Принципова відмінність АЕ від інших методів полягає в тому, що сигнали АЕ генеруються безпосередньо місцем дефекту, а не є відгуком конструкції на зовнішнє збудження, як у вібраційному аналізі [15].

Для діагностики підшипників кочення АЕ є найбільш чутливим із доступних методів: вона здатна виявляти одиничні удари тіл кочення об дефектну ділянку доріжки кочення навіть при мікроскопічних розмірах пошкодження, коли жоден інший метод ще не реагує. Це робить АЕ особливо цінним інструментом для виявлення зародкових стадій пошкоджень, що є критичним для систем, де будь-яка відмова підшипника недопустима через ризики безпеки.

Проте практичне промислове застосування АЕ суттєво ускладнюється. По-перше, метод вимагає спеціалізованого широкосмугового п'єзоелектричного обладнання та складних систем попередньої обробки сигналів із фільтрацією завад у верхньочастотному діапазоні. По-друге, сигнали АЕ є дуже слабкими і легко маскуються промисловими завадами – шумом сусіднього обладнання, гідроударами в трубопроводах, електричними завадами. По-третє, забезпечення надійного акустичного контакту між датчиком і корпусом обладнання вимагає ретельної підготовки поверхні. Сукупність цих факторів робить широкомасштабне промислове впровадження АЕ технічно і економічно складним, що обмежує її застосування переважно відповідальними об'єктами атомної, аерокосмічної та нафтохімічної промисловості.

1.2.5. Вібраційний аналіз

Вібраційний аналіз є найбільш поширеним, технологічно зрілим і практично ефективним методом діагностики обертового устаткування, що підтверджується десятиліттями успішного промислового застосування. Метод ґрунтується на вимірюванні та аналізі механічних коливань корпусу машини – прискорення, швидкості або переміщення – за допомогою датчиків (акселерометрів, велосиметрів або датчиків переміщення), що встановлюються на зовнішній поверхні корпусу у безпосередній близькості від підшипникових опор [5].

Фізична основа методу полягає в тому, що будь-який дефект механічного або електромагнітного характеру змінює динамічні процеси у відповідних вузлах

машини і, через механічні зв'язки конструкції, відображається у характері коливань корпусу. Зокрема, розвиток дефекту на доріжці кочення підшипника призводить до появи періодичних імпульсних ударів при проходженні тіл кочення через пошкоджену ділянку, що формує характерні частотні складові у спектрі вібрацій. Дисбаланс ротора спричиняє відцентрові сили, що обертаються разом із ротором і збуджують коливання на частоті обертання. Розцентровка осей призводить до знакозмінних навантажень на підшипники з подвоєною частотою обертання [6].

Перевагами вібраційного аналізу, що обумовлюють його домінуюче положення серед методів СМ, є такі:

- висока чутливість до широкого спектру механічних та електромагнітних дефектів – підшипників, валів, роторів, з'єднань, обмоток – у рамках єдиного вимірювального каналу;
- здатність виявляти дефекти на ранніх стадіях розвитку, коли пошкодження ще є мінімальним і не впливає на продуктивність обладнання;
- неінвазивність вимірювань – датчики встановлюються на зовнішній поверхні корпусу без розкриття або зупинки машини;
- придатність для безперервного моніторингу в режимі реального часу, що дозволяє відстежувати динаміку розвитку дефекту;
- добре розроблена теоретична база для інтерпретації характерних частот дефектів, що дозволяє ідентифікувати конкретний несправний вузол;
- відносно невисока вартість акселерометрів та систем реєстрації порівняно з обладнанням для акустичної емісії або рентгенографічного контролю;
- можливість інтеграції у системи АСУ ТП без суттєвого ускладнення апаратної частини завдяки стандартним аналоговим або цифровим інтерфейсам.

Порівняльна характеристика основних методів моніторингу стану обертового устаткування за ключовими критеріями наведена у таблиці 1.2.

Таблиця 1. 2 Порівняльна характеристика методів моніторингу стану обертового устаткування

Метод	Чутливість (ранні стадії)	Безперервний моніторинг	Вартість впровадження	Складність аналізу	Типові об'єкти
Вібраційний аналіз	Висока	Так	Низька	Середня	Підшипники, вали, ротор
Термографія	Низька	Так	Середня	Низька	Обмотки, контакти
Аналіз мастил	Висока	Ні	Висока	Висока	Підшипники, редуктори
MCSA	Середня	Так	Дуже низька	Середня	Ротор, підшипники
Акустична емісія	Дуже висока	Так	Висока	Дуже висока	Тріщини, руйнування

Аналіз таблиці 1.2 підтверджує, що вібраційний аналіз є оптимальним методом для задач моніторингу стану обертового енергетичного устаткування в системах автоматизації, забезпечуючи найбільш збалансоване поєднання чутливості, можливості безперервного моніторингу, вартості та практичності впровадження. Цим і обумовлений його вибір як основи системи предиктивної діагностики, що розробляється у даній роботі.

1.3. Вібраційний аналіз як основний метод діагностики

Теоретичною основою вібраційної діагностики є взаємозв'язок між динамічними процесами у внутрішніх вузлах машини та коливаннями її корпусу, що реєструються зовнішніми датчиками. Вібраційний сигнал електродвигуна є суперпозицією великої кількості компонентів різної природи, кожен із яких несе інформацію про стан відповідного вузла або процесу [5]. Завдання вібраційної

діагностики полягає у виділенні та ідентифікації компонентів, характерних для несправного стану, на фоні нормальних вібрацій та шуму.

Математично вібраційний сигнал $x(t)$ справного електродвигуна може бути представлений як:

$$x(t) = x_{bal}(t) + x_{bear}(t) + x_{em}(t) + x_{noise}(t), \quad (1.1)$$

де $x_{bal}(t)$ – компонента дисбалансу, що визначається неоднорідністю розподілу маси ротора; $x_{bear}(t)$ – компонента від роботи підшипників кочення; $x_{em}(t)$ – електромагнітна компонента, зумовлена взаємодією ротора і статора; $x_{noise}(t)$ – широкосмуговий шум, що включає випадкові механічні та акустичні збудження.

Кожна компонента вібраційного сигналу проявляється у певному частотному діапазоні. Компонента дисбалансу зосереджена на частоті обертання ротора та її гармоніках, компоненти підшипників – на характеристичних частотах дефектів, що визначаються геометрією підшипника, а електромагнітна компонента – на частотах, кратних частоті мережі живлення. Ця частотна локалізація дозволяє ідентифікувати джерело вібрацій шляхом аналізу спектру сигналу. Характеристичні частоти основних типів дефектів розглянуто у підрозділах 1.3.1 та 1.3.2.

Реєстрація вібраційних сигналів здійснюється п'єзоелектричними акселерометрами, що встановлюються на зовнішній поверхні корпусу електродвигуна поблизу підшипникових опор. Основним вимірюваним параметром є віброприскорення $a(t)$, що вимірюється в одиницях прискорення вільного падіння (g , де $1g = 9,81 \text{ м/с}^2$). Вибір віброприскорення обумовлений його високою чутливістю до високочастотних імпульсних компонент, характерних для ранніх стадій механічних дефектів. Вимірювання можуть проводитися у трьох напрямках: радіальному (перпендикулярному осі вала), осьовому (паралельному осі вала) та тангенціальному (дотичному до поверхні). Кожен напрямок несе інформацію про різні типи дефектів, і вибір осей вимірювання визначається конкретною діагностичною задачею. Обґрунтування вибору осей для даного дослідження наведено у підрозділі 3.2.3.

При розвитку несправності до рівняння (1.1) додається компонента $x_{fault}(t)$, що несе в собі характерну сигнатуру конкретного типу дефекту. Завдання діагностики полягає у виявленні та ідентифікації цієї компоненти.

1.3.1. Характеристичні частоти дефектів підшипників

Підшипники кочення є найбільш поширеним джерелом несправностей електродвигунів, що визначає важливість їх точної діагностики. Геометрія підшипника кочення однозначно визначає набір характеристичних частот, на яких проявляються дефекти різних його елементів. Знання цих частот є необхідною умовою правильної інтерпретації спектру вібрацій.

Частота проходження тіл кочення через дефект на зовнішньому кільці (Ball Pass Frequency Outer, BPFO) визначається як:

$$BPFO = \frac{N_b}{2} f_r \left(1 - \frac{d}{D} \cos \alpha \right) \quad (1.2)$$

Частота проходження тіл кочення через дефект на внутрішньому кільці (Ball Pass Frequency Inner, BPFI) визначається як:

$$BPFI = \frac{N_b}{2} f_r \left(1 + \frac{d}{D} \cos \alpha \right) \quad (1.3)$$

Частота обертання сепаратора (Fundamental Train Frequency, FTF) відповідає власній частоті сепаратора і визначається як:

$$FTF = \frac{f_r}{2} \left(1 - \frac{d}{D} \cos \alpha \right) \quad (1.4)$$

Частота обертання тіл кочення (Ball Spin Frequency, BSF) є характеристичною частотою дефекту на поверхні тіла кочення і визначається як:

$$BSF = \frac{D}{2d} f_r \left(1 - \left(\frac{d}{D} \right)^2 (\cos \alpha)^2 \right), \quad (1.5)$$

де у формулах (1.2)–(1.5): N_b – кількість тіл кочення; f_r – частота обертання ротора (Гц); d – діаметр тіла кочення (мм); D – діаметр кола розташування центрів тіл кочення (мм); α – кут контакту тіл кочення з доріжками.

Наявність дефекту на відповідному елементі підшипника проявляється у спектрі вібрацій у вигляді підвищених амплітуд на відповідній характеристичній частоті та її гармоніках. На початкових стадіях розвитку дефекту ці компоненти можуть мати малу амплітуду і перекриватися шумом, що ускладнює їх виявлення

методами простого спектрального аналізу та обумовлює необхідність застосування більш досконалих методів обробки сигналів.

1.3.2. Характеристичні ознаки інших типів несправностей

Дисбаланс ротора є одним із найбільш поширених дефектів, що виникає внаслідок нерівномірного розподілу маси ротора відносно осі обертання. Причинами дисбалансу можуть бути технологічні відхилення при виготовленні, налипання забруднень на лопатки або диски, нерівномірний знос, а також деформація ротора внаслідок термічних або механічних впливів. У спектрі вібрацій дисбаланс проявляється у вигляді домінуючої амплітуди на основній частоті обертання f_r та незначних амплітуд на її вищих гармоніках ($2f_r, 3f_r, \dots$). Характерною ознакою дисбалансу є переважно радіальний напрям вібрацій і їх синусоїдальна форма у часовій області [5].

Розцентровка (зміщення осей) виникає при відсутності збігу осей обертання ротора двигуна та механізму, що приводиться. Розрізняють паралельну розцентровку – коли осі паралельні, але не збігаються – та кутову – коли осі перетинаються під кутом. Змішана розцентровка є комбінацією обох видів і є найбільш поширеною в практиці. У спектрі вібрацій розцентровка характеризується підвищеними амплітудами на частотах f_r та $2f_r$, при цьому амплітуда $2f_r$, як правило, більша за f_r , що дозволяє відрізнити розцентровку від дисбалансу. При кутовій розцентровці також суттєво зростають осьові вібрації на частоті f_r .

Вигин вала є деформацією геометричної осі вала, при якій вона відхиляється від прямолінійної форми. Невеликий залишковий вигин, що не виявляється при статичному контролі, може суттєво впливати на динамічні характеристики машини при обертанні. У спектрі вібрацій вигин вала проявляється домінуванням амплітуди на частоті $2f_r$, що часто супроводжується підвищеними осьовими вібраціями. Вигин вала поблизу центра прольоту між підшипниками дає більшу амплітуду на f_r , ніж $2f_r$, що відрізняє його від розцентровки.

Механічне послаблення кріплень фундаментних болтів або з'єднань конструктивних елементів призводить до нелінійної поведінки системи. При

наявності люфтів в елементах конструкції виникають удари у місцях втрати контакту, що спричиняють збагачений гармоніками спектр вібрацій. Характерною ознакою механічного послаблення є поява субгармонік ($0.5f_r, 1.5f_r$) та великої кількості вищих гармонік ($3f_r, 4f_r, 5f_r, \dots$) у спектрі вібрацій, часто з нерівномірними амплітудами.

Пошкодження стержнів ротора є специфічним дефектом асинхронних електродвигунів. Обриви або тріщини у стержнях короткозамкненої обмотки ротора порушують симетрію його магнітного поля, що призводить до нерівномірного розподілу електромагнітних сил і пульсацій крутного моменту. У вібраційному сигналі ці пульсації проявляються як амплітудна модуляція на частоті $2sf_1$, де s – ковзання двигуна, f_1 – частота живлення. Це породжує характерні бокові смуги навколо вищих гармонік частоти обертання у спектрі вібрацій.

На рисунках 1.3 та 1.4 наведено характерні спектри вібраційних сигналів електродвигуна для семи досліджуваних режимів роботи. Аналіз спектрів дозволяє виявити характерні ознаки кожного типу несправності: домінування основної частоти обертання f_r при дисбалансі, переважання другої гармоніки $2f_r$ при зміщенні осей та вигині вала, появу субгармонік при механічному послабленні, бокових смуг $\pm 2sf_1$ при пошкодженні стержнів ротора, а також характеристичних частот BPFO та їх гармонік з боковими смугами $\pm f_r$ при дефекті підшипника.

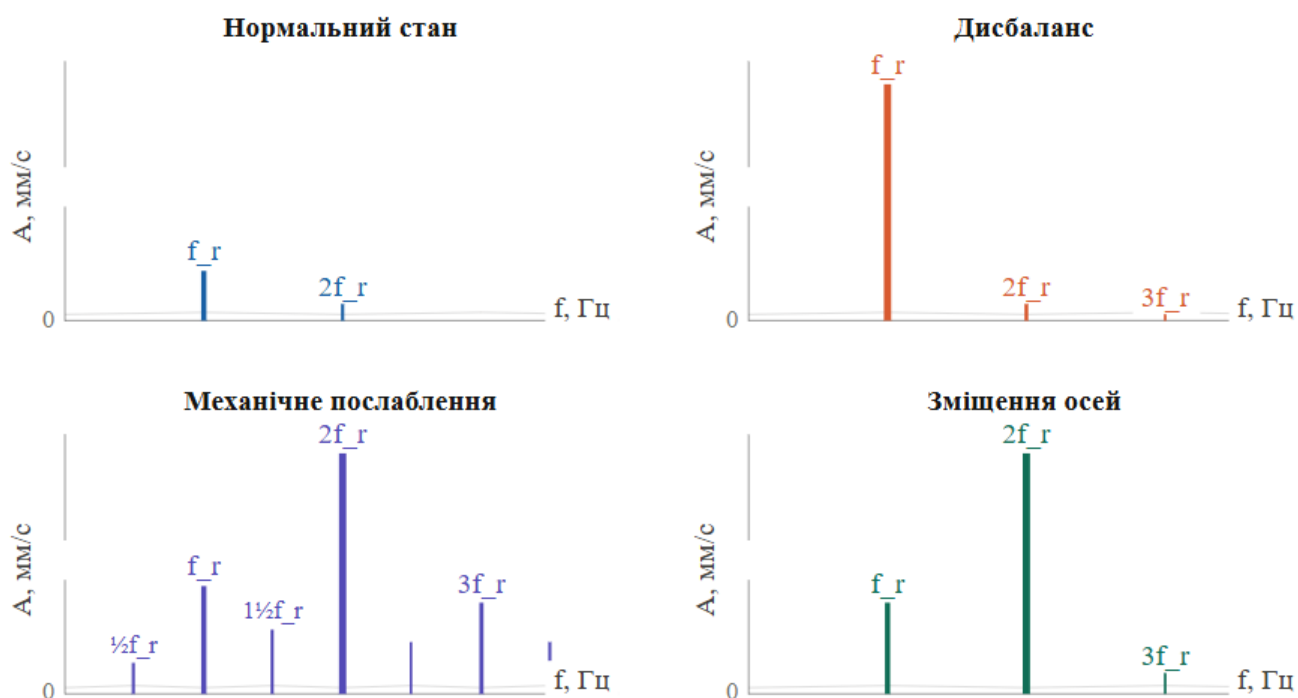


Рисунок 1. 3 Характерні спектри вібраційних сигналів електродвигуна при нормальному стані, дисбалансі, механічному послабленні та зміщенні осей

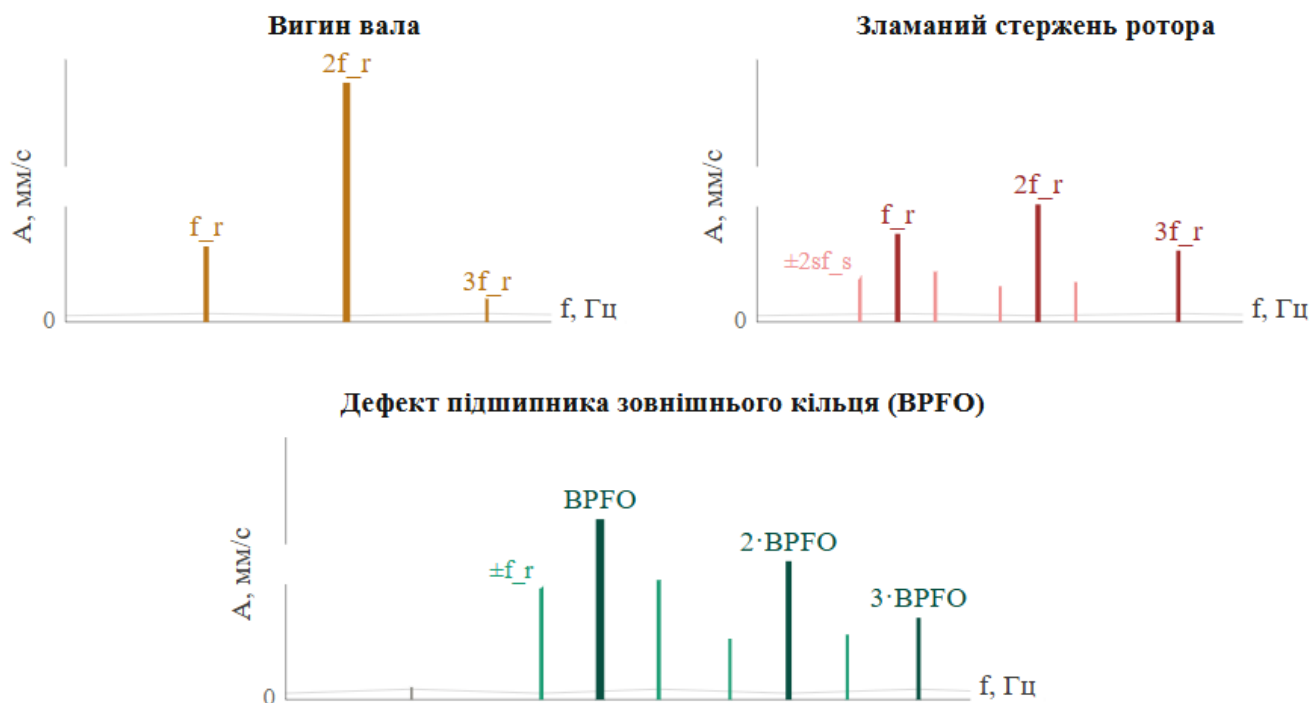


Рисунок 1. 4 Характерні спектри вібраційних сигналів електродвигуна при вигині вала, пошкодженні стержня ротора та дефекті підшипника зовнішнього кільця (BPFO)

Таким чином, різні типи несправностей мають характерні спектральні ознаки, що теоретично дозволяє ідентифікувати дефект за його частотною сигнатурою. Проте на практиці ці ознаки можуть перекриватися, а їх інтерпретація вимагає значного досвіду та експертних знань, що обумовлює доцільність застосування автоматизованих методів класифікації.

1.3.3. Аналіз вібраційних сигналів у часовій області

Процес вібраційної діагностики складається з декількох послідовних етапів: реєстрація сигналу, його попередня обробка, аналіз та класифікація стану. На кожному з цих етапів застосовуються специфічні методи, вибір яких суттєво впливає на точність та надійність кінцевого діагнозу [8].

Аналіз у часовій області є найпростішим і найшвидшим методом і передбачає безпосереднє дослідження форми сигналу прискорення $a(t)$ у часі. Для кількісної оцінки стану обладнання з вібраційного сигналу обчислюються статистичні показники, що характеризують різні аспекти його динамічної поведінки.

Загальний енергетичний рівень вібрацій оцінюється середньоквадратичним значенням (Root Mean Square, RMS). Форму розподілу амплітуд характеризують куртозис (Kurtosis, K), що відображає гостровершинність розподілу, та асиметрія (Skewness, Skew), що вказує на його несиметричність. Наявність пікових складових оцінюється пік-фактором (Crest Factor, CF), форм-фактором (Shape Factor, SF) та імпульс-фактором (Impulse Factor, IF).

Сукупність зазначених показників формує вектор ознак для подальшої класифікації стану обладнання.

Основним обмеженням аналізу у часовій області є те, що статистичні показники не дозволяють розрізнити різні типи дефектів між собою – вони лише фіксують факт відхилення від норми, але не надають інформації про характер несправності. Крім того, ці показники залежать від режиму роботи машини, що ускладнює їх використання при змінних умовах. Повний математичний апарат статистичних показників та формування розширеного вектора ознак наведено у підрозділі 2.1.

1.3.4. Аналіз вібраційних сигналів у частотній області

Аналіз вібраційного сигналу у часовій області дозволяє виявити загальне відхилення від норми, проте не дає інформації про те, які саме частотні складові змінились. Для ідентифікації характеру несправності необхідний перехід до частотної області, який дозволяє розкласти складний вібраційний сигнал на елементарні гармонічні складові та виявити характерні частоти дефектів – гармоніки ротора, частоти підшипникових дефектів та бокові смуги модуляції.

Математичною основою такого переходу є перетворення Фур'є, яке представляє сигнал як суму синусоїдальних коливань різних частот та амплітуд. Оскільки вібраційний сигнал реєструється у цифровій формі, для його аналізу застосовується дискретне перетворення Фур'є (ДПФ). Пряме обчислення ДПФ є обчислювально затратним, тому на практиці використовується його прискорена версія – алгоритм швидкого перетворення Фур'є (FFT), який дає математично ідентичний результат за значно меншу кількість операцій [5, 8].

Важливою особливістю спектрального аналізу є ефект витоку спектру. Він виникає через те, що будь-який реальний сигнал має кінцеву тривалість – фактично це еквівалентно множенню нескінченного сигналу на прямокутне вікно, яке різко обрізає його на краях. Таке різке обрізання спричиняє розмиття частотних піків у спектрі: енергія однієї частоти «витікає» на сусідні частотні відліки, що ускладнює ідентифікацію близько розташованих компонент. Для зменшення цього ефекту перед обчисленням перетворення Фур'є сигнал множиться на спеціальну вікнову функцію, яка плавно зводить його амплітуду до нуля на краях. Існують різні типи вікон – Хеннінга, Хеммінга, Блекмана, Блекмана-Харріса – кожне з яких забезпечує різний компроміс між шириною головного піку та рівнем бокових пелюсток у спектрі. Вибір конкретного вікна залежить від задачі та обговорюється у підрозділі 2.1.

Принциповим обмеженням FFT є відсутність у результуючому спектрі інформації про те, коли саме у часі виникає та чи інша частотна складова – результатом є усереднений спектр за всю тривалість сигналу. Для стаціонарних сигналів, частотний склад яких не змінюється у часі, це не є проблемою. Проте при

змінних режимах роботи обладнання або при наявності перехідних процесів це обмеження стає суттєвим.

Короткочасне перетворення Фур'є (STFT) вирішує цю проблему, розбиваючи сигнал на короткі перекривні часові фрагменти і обчислюючи спектр для кожного фрагмента окремо. Перед обчисленням спектру кожен фрагмент також множиться на вікнову функцію – так само як і при звичайному FFT – для зменшення витоку спектру. Результатом STFT є спектрограма – двовимірне зображення, де горизонтальна вісь відповідає часу, вертикальна – частоті, а інтенсивність кольору – амплітуді відповідної частотної складової у даний момент часу.

Спектрограми є інформативним представленням для задач нейромережевої класифікації, оскільки різні типи дефектів формують на них характерні візуальні патерни. Горизонтальні лінії відповідають стаціонарним гармонікам ротора, вертикальні спалахи через рівні інтервали – ударним імпульсам при дефектах підшипників, а амплітудна модуляція горизонтальних ліній – пошкодженням стержнів ротора [5].

Разом з тим, STFT має обмеження у вигляді компромісу між часовою та частотною роздільною здатністю: збільшення тривалості вікна підвищує точність визначення частоти, але знижує точність визначення моменту часу, і навпаки. Вибір оптимальних параметрів вікна є важливим етапом при підготовці спектрограм для задач класифікації. Детальний математичний апарат ДПФ, STFT, порівняння вікнових функцій та обґрунтування вибору параметрів спектрального аналізу наведено у підрозділі 2.1.

1.4. Огляд існуючих досліджень та публікацій

1.4.1. Класичні методи обробки вібраційних сигналів у задачах діагностики

Розвиток методів вібраційної діагностики обертового устаткування на основі машинного навчання бере початок із кінця 1990-х – початку 2000-х років, коли накопичення обчислювальних потужностей зробило практично реалізованим застосування відносно складних алгоритмів класифікації до великих масивів вібраційних даних. Перші роботи цього напрямку зосереджувались на побудові класифікаторів, що використовують ознаки, вручну виділені зі спектрів вібрацій –

амплітуди на характеристичних частотах, ширини піків, відношення сигнал-шум на частотах дефектів. Такий підхід вимагав значної галузевої експертизи при формуванні набору ознак, проте при правильному виборі ознак забезпечував достатньо високу точність класифікації [14].

Систематичний огляд методів машинного навчання, що застосовуються до діагностики машин, проведений у [14], охоплює роботи за період понад двох десятиліть і виявляє наступні тенденції. По-перше, метод опорних векторів (SVM) тривалий час залишався домінуючим підходом завдяки теоретичній обґрунтованості, ефективності на вибірках обмеженого розміру та стійкості до перенавчання. По-друге, ансамблеві методи – випадковий ліс, градієнтний бустинг – показали вищу точність порівняно з SVM на великих вибірках. По-третє, нейронні мережі поступово витіснили класичні методи ML із найвищих позицій у міру збільшення доступних обсягів навчальних даних. По-четверте, незважаючи на різноманітність підходів, більшість робіт фокусується на діагностиці дефектів підшипників, залишаючи менш дослідженими комплексну класифікацію різноманітних дефектів електродвигуна.

У роботі [17] запропоновано метод класифікації дефектів підшипників на основі ознак часової та частотної областей, включаючи середнє значення, дисперсію, асиметрію, куртозис, медіану та діапазон сигналу. Для підвищення якості навчальної вибірки застосовано підхід Median-Based Outlier Detection (MOD) для виявлення та виключення аномальних зразків, характеристики яких зумовлені зовнішніми факторами, а не дефектами. Показано, що правильне попереднє очищення даних суттєво підвищує точність класифікатора. Разом із тим авторами не розглядалась класифікація схожих за проявами дефектів різного ступеня розвитку – одна з найскладніших практичних задач вібраційної діагностики.

Дослідження [18] є важливим з точки зору аналізу взаємозв'язку між параметрами реєстрації сигналів та точністю діагностики. Автори порівняли результати чотирьох алгоритмів ML – SVM, дерев рішень, MLP та LSTM – при різних частотах дискретизації вібраційних сигналів асинхронного двигуна. Ключовий висновок полягає у тому, що нейронні мережі, особливо глибокі

архітектури, суттєво чутливіші до зменшення частоти дискретизації, ніж класичні методи. При зниженні частоти дискретизації із 48 кГц до 12 кГц точність MLP знижувалась на 15–20%, тоді як SVM – лише на 5–7%. Водночас при достатній частоті дискретизації нейронні мережі суттєво перевершували SVM за точністю класифікації.

У [20] проведено систематичний огляд методів візуалізації вібраційних сигналів для діагностики обертових машин. Розглянуто підходи перетворення одновимірних часових сигналів у двовимірні зображення в часовій, частотній та часово-частотній областях. Автори виявили, що часово-частотні представлення – спектрограми STFT та вейвлет-скалограми – забезпечують найбільш інформативне двовимірне представлення для наступної класифікації засобами CNN. Проведено також кількісне порівняння роздільної здатності різних перетворень і виявлено оптимальні параметри STFT для задач діагностики підшипників.

Підхід, заснований на застосуванні комплексного вейвлету Морле для отримання часово-частотних характеристик вібраційних сигналів із подальшим перетворенням на кольорові RGB-зображення для класифікації засобами CNN, запропоновано у [19]. Перетворення вейвлету Морле є особливо ефективним для виявлення перехідних та нестационарних компонент сигналів, що характерні для ударних процесів у пошкоджених підшипниках. Отримані кольорові часово-частотні зображення відрізняються наочністю і дозволяють чітко диференціювати різні типи дефектів за характером розподілу інтенсивності. Проте автори відзначили обмеженість тестової вибірки та залежність результатів від правильного вибору параметрів базисного вейвлету.

1.4.2. Методи машинного навчання у задачах діагностики обертового устаткування

Застосування методів машинного навчання до задач діагностики обертового устаткування пройшло кілька еволюційних етапів. На першому етапі домінували класичні лінійні класифікатори та SVM, що навчались на вручну сформованих наборах ознак. На другому – ансамблеві методи та нелінійні ядрові класифікатори. На третьому, сучасному – методи глибокого навчання, що автоматично виділяють

ознакові представлення з «сирих» даних [14].

Метод опорних векторів, незважаючи на появу потужніших конкурентів, зберігає практичну значущість у задачах діагностики завдяки своїм унікальним властивостям. SVM будує оптимальну гіперплощину, що максимізує відстань між класами у просторі ознак, використовуючи лише опорні вектори – найближчі до межі поділу зразки. Це забезпечує хорошу узагальнювальну здатність навіть на вибірках обмеженого розміру, що є важливим для промислових застосувань, де дані про несправності можуть бути нечисленними [14].

Алгоритм Random Forest продемонстрував ефективність у ряді задач промислової діагностики. У [22] показано, що ансамблевий метод Random Forest у поєднанні з вейвлет-перетворенням для виділення ознак струмових сигналів забезпечує високу точність класифікації дефектів підшипників індукційних двигунів. Метод є стійким до нерелевантних ознак та забезпечує оцінку важливості ознак, що корисно для аналізу інформативності різних характеристик сигналу. Проте алгоритм залишається залежним від якості вручну сформованого набору ознак і не здатен автоматично адаптуватись до нових типів несправностей.

Підхід на основі навчання без учителя, запропонований у [21], є особливо актуальним для промислових застосувань, де марковані дані про несправні стани можуть бути недоступними або вкрай нечисленними. Автори побудували модель виявлення аномалій для двох типів двигунів – нового та промислового зі значним напрацюванням – навчаючи модель виключно на даних нормального режиму роботи. Відхилення від засвоєного нормального патерну класифікується як аномалія. Метод показав задовільні результати при стандартизованих умовах роботи, однак не здатен ідентифікувати конкретний тип несправності – лише сигналізувати про відхилення від норми. Крім того, точність методу суттєво залежить від повноти навчальних даних щодо нормального режиму: якщо під час навчання не було представлено деяких нормальних режимів роботи, вони можуть надалі неправильно класифікуватись як аномалії.

Принциповим обмеженням усіх класичних методів ML є залежність від ручного формування ознак. Якість класифікації повністю визначається вибором

ознак, що вимагає глибокої галузевої експертизи та є трудомістким процесом. При появі нових типів несправностей весь процес розробки системи діагностики доводиться проходити заново. Подолання цього обмеження і стало головною мотивацією для переходу до методів глибокого навчання [14].

1.4.3. Методи глибокого навчання та нейромережеві архітектури

Революція глибокого навчання, що розпочалася у комп'ютерному зорі та обробці природної мови, закономірно поширилась на задачі промислової діагностики. Здатність глибоких нейронних мереж автоматично виділяти ієрархічні ознакові представлення безпосередньо з «сирих» даних відкрила можливість побудови систем діагностики, що не потребують ручного проектування ознак. При наявності достатнього обсягу маркованих навчальних даних глибокі мережі стабільно перевершують класичні методи ML у точності класифікації [9].

Систематичний аналіз архітектур глибокого навчання, що застосовуються у задачах предиктивного технічного обслуговування, проведений у [9], охоплює чотири основні класи: глибокі мережі переконань (DBN), автоенкодера (AE), згорткові нейронні мережі (CNN) та рекурентні нейронні мережі (RNN). Встановлено, що CNN найбільш ефективні для класифікації вібраційних сигналів, поданих у вигляді двовимірних зображень; RNN та LSTM – для аналізу часових залежностей у послідовних даних; автоенкодера – для виявлення аномалій без учителя. Гібридні архітектури, що поєднують переваги кількох класів, формують найбільш перспективний напрям розвитку.

Застосування CNN до задач діагностики підшипників розглянуто у [10]. Запропонована двогілкова архітектура окремо обробляє часові та частотні характеристики вібраційних сигналів у двох паралельних гілках CNN, після чого отримані ознакові вектори конкатенуються та подаються до повнозв'язних шарів для класифікації. Така архітектура забезпечує комплементарне використання інформації з різних областей сигналу. Водночас двогілкова структура суттєво збільшує кількість параметрів моделі та час навчання, що є практичним обмеженням при обмежених обчислювальних ресурсах.

Підхід до діагностики несправностей асинхронного двигуна на основі

поєднання часового та часово-частотного аналізу вібраційних сигналів із CNN запропоновано у [23]. Автори перетворюють одновимірний вібраційний сигнал у двовимірні спектрограми STFT, які подаються на вхід попередньо навченої CNN із дообученням (transfer learning) для класифікації несправностей ротора при різних умовах навантаження. Застосування transfer learning дозволяє досягати прийнятної точності класифікації навіть при обмеженому обсязі маркованих даних – що є типовою практичною проблемою промислових систем діагностики. Разом із тим розглянутий підхід зосереджений переважно на дефектах ротора і не забезпечує комплексної класифікації всього спектру можливих несправностей електродвигуна, що обмежує його застосовність у повноцінних системах предиктивного обслуговування.

Рекурентні нейронні мережі із LSTM є природним інструментом для аналізу вібраційних сигналів як часових послідовностей. Архітектура LSTM із трьома типами вентилів дозволяє мережі вибірково зберігати або забувати інформацію з попередніх часових кроків, уникаючи проблеми зникнення градієнта. Авторами [11] запропоновано гібридну архітектуру LSTM-FCN, що поєднує паралельно LSTM-гілку для аналізу часових залежностей та повнозв'язну CNN-гілку для виділення локальних ознак. Порівняльне дослідження показало перевагу LSTM-FCN над чистими LSTM та CNN на багатовимірних вібраційних даних.

Важливим напрямом є також розробка інтерпретованих методів глибокого навчання. У [12] запропоновано WaveletKernelNet – нейронну мережу зі вбудованими вейвлет-ядрами у першому шарі замість стандартних ядер CNN. Вейвлет-ядра мають конкретний фізичний сенс – вони виявляють компоненти сигналу на конкретних часово-частотних смугах, що відповідають характеристикам певних типів дефектів. Завдяки цьому активації нейронів першого шару можуть бути безпосередньо інтерпретовані фахівцями-механіками. Це суттєво підвищує довіру до результатів діагностики та полегшує верифікацію моделі. Запропонована мережа показала вищу точність та кращу узагальнювальну здатність порівняно зі стандартними CNN на кількох стандартних наборах даних.

Гібридна архітектура MSCNN-LSTM, що поєднує багатомасштабну CNN із

LSTM та механізмом уваги, запропонована та досліджена у [9]. Ключовою відмінністю MSCNN від стандартної CNN є використання паралельних гілок із ядрами різних розмірів (3 та 5 одиниць для одновимірної CNN), що дозволяє одночасно виділяти ознаки різного масштабу. Отримані ознакові вектори об'єднуються та передаються до блоку LSTM, який враховує часові залежності між послідовними вікнами сигналу. Механізм уваги зважає на вихідні вектори LSTM відповідно до їх діагностичної значущості, що підвищує стійкість класифікації. Запропонований метод тестувався на наборі вібраційних даних із п'ятьма типами несправностей та досяг 100% точності класифікації, підтвердженої аналізом матриці помилок та t-SNE-візуалізацією простору ознак.

Методологія послідовного аналізу вібраційних сигналів у часовій та частотній областях із подальшим формуванням спектрограм STFT та їх класифікацією засобами CNN розроблена у [5]. Дослідження охоплює сім режимів роботи двигуна – нормальний стан та шість типів несправностей – та демонструє поступове покращення діагностичних можливостей при переході від часової до частотної та часово-частотної областей. Результати підтверджують, що аналіз у часовій області є найменш інформативним (візуальні відмінності між сигналами різних дефектів практично відсутні), FFT дозволяє ідентифікувати несправності за характерними частотами, а STFT-спектрограми у поєднанні з CNN забезпечують найвищу точність класифікації.

Аналіз розглянутих публікацій дозволяє виявити кілька систематичних прогалин у дослідженнях. По-перше, переважна більшість робіт орієнтована на діагностику дефектів підшипників, тоді як комплексна класифікація різноманітних типів несправностей – дисбалансу, розцентровки, вигину вала, механічного послаблення, пошкоджень стержнів ротора та підшипникових дефектів – у рамках єдиної системи залишається недостатньо вивченою. По-друге, систематичний порівняльний аналіз методів різної складності у рамках єдиної дослідницької методології та на єдиному наборі даних представлений лише в одиничних роботах.

1.5. Висновки до розділу

1.5.1. Сучасний стан та тенденції у діагностиці обертового устаткування

Аналіз сучасних тенденцій у сфері промислової автоматизації підтверджує зростаючу актуальність задачі предиктивної діагностики обертового енергетичного устаткування. Розвиток концепції Industry 4.0 та перехід промислових підприємств до інтелектуального управління виробничими активами формують запит на системи безперервного моніторингу стану обладнання з автоматичним виявленням ознак деградації та прогнозуванням залишкового ресурсу. Електродвигуни як ключові елементи приводних механізмів у системах автоматичного керування – теплових пунктів, насосних станцій, вентиляційних установок – є об'єктами першочергової уваги, оскільки їх відмова безпосередньо порушує функціональність відповідних контурів регулювання. Статистика IEEE підтверджує, що дефекти підшипників, обмоток та ротора є найбільш поширеними причинами відмов і забезпечують разом понад 90% усіх несправностей. Традиційні стратегії – реактивне та планово-попереджувальне обслуговування – є субоптимальними: перше призводить до аварійних зупинок, друге – до нераціонального використання ресурсу. Концепція PdM, заснована на безперервному моніторингу та прогнозуванні, є об'єктивно необхідним напрямом розвитку систем автоматизації.

1.5.2. Методи обробки вібраційних сигналів

Серед методів моніторингу стану обертового устаткування – термографії, аналізу мастил, MCSA, акустичної емісії та вібраційного аналізу – останній посідає провідне місце завдяки оптимальній сукупності характеристик: висока чутливість до широкого класу механічних та електромагнітних дефектів, придатність для безперервного моніторингу в режимі реального часу, неінвазивність вимірювань та відносно низька вартість впровадження. Класичні методи обробки вібраційних сигналів утворюють ієрархію зростаючої інформативності: аналіз у часовій області забезпечує оперативну інтегральну оцінку, FFT дозволяє ідентифікувати характерні частоти дефектів, STFT поєднує часову та частотну інформацію у вигляді спектрограм, що несуть значно більше діагностичної інформації. Спектрограми є природним вхідним форматом для згорткових нейронних мереж,

що виявляють характерні просторові патерни у двовимірних зображеннях.

1.5.3. Нейромережеві підходи до діагностики

Методи глибокого навчання – CNN та гібридні архітектури CNN-LSTM – демонструють суттєву перевагу над класичними підходами ML завдяки автоматичному виділенню ієрархічних ознакових представлень без необхідності ручного проектування ознак. CNN ефективно розпізнають характерні часово-частотні патерни у спектрограмах вібраційних сигналів, тоді як LSTM забезпечують моделювання часової динаміки розвитку дефектів. Гібридні архітектури MSCNN-LSTM, що поєднують багатомасштабне виділення ознак із аналізом часових залежностей та механізмом уваги, демонструють найвищу точність класифікації серед розглянутих підходів. Водночас більшість досліджень зосереджена на діагностиці підшипників, що залишає відкритим питання комплексної класифікації різноманітних несправностей електродвигуна.

1.5.4. Подальші напрями досліджень

На підставі проведеного огляду літератури виявлено наступні невирішені проблеми, що визначають напрям даного дисертаційного дослідження. По-перше, відсутній систематичний порівняльний аналіз підходів різної складності – від аналізу у часовій та частотній областях через CNN на спектрограмах STFT до гібридної архітектури MSCNN-LSTM – у рамках єдиної дослідницької методології та на єдиному наборі даних. Це унеможливорює об'єктивне порівняння підходів та обґрунтований вибір оптимального. По-друге, більшість існуючих робіт обмежуються діагностикою одного-двох типів дефектів, не розглядаючи комплексну класифікацію повного спектру несправностей електродвигуна, характерних для реальних умов експлуатації в системах автоматизації. По-третє, питання впливу гіперпараметрів нейронних мереж на точність класифікації різних типів дефектів та методи їх оптимізації розглядаються лише в одиничних роботах.

Усунення зазначених недоліків є метою даної дисертаційної роботи, що передбачає розробку та порівняльне дослідження методів предиктивної діагностики обертового енергетичного устаткування – від класичних підходів до гібридної нейромережевої архітектури на основі MSCNN-LSTM – з

обґрунтуванням вибору оптимального рішення та оцінкою практичних можливостей його інтеграції в системи автоматизації промислових об'єктів.

Розділ 2. Методи обробки вібраційних сигналів та нейромережеві архітектури

2.1. Математичні основи аналізу вібраційних сигналів

Для побудови ефективної системи класифікації стану необхідно детально розглянути математичний апарат трьох рівнів аналізу вібрації: часового, частотного та часово-частотного. У розділі 1 ці методи розглянуто з огляду на їх діагностичну інформативність. У даному підрозділі буде розглянуто повний математичний апарат кожного методу в тій формі, що безпосередньо використовується у розробленій системі.

2.1.1. Часова область. Статистичні показники, формування вектора ознак та базовий класифікатор

Аналіз у часовій області є первинним етапом обробки вібраційних сигналів і забезпечує швидку оцінку загального стану обладнання без значних обчислювальних витрат. Основу аналізу складають статистичні показники, що кількісно описують розподіл амплітуд сигналу [24].

Середньоквадратичне значення (RMS) є головним енергетичним показником рівня вібрацій і визначається як:

$$RMS = \sqrt{\frac{\sum x_i^2}{N}}, \quad (2.1)$$

де N – кількість відліків сигналу; x_i – значення i -го відліку віброприскорення; підсумовування виконується за всіма відліками від $i = 1$ до N . Фізично RMS відображає середню потужність вібраційного сигналу. Значення RMS є основним нормованим параметром оцінювання технічного стану за стандартами ISO 10816 та ISO 20816, що встановлюють граничні рівні вібрацій для різних класів обладнання. Підвищення RMS свідчить про загальне погіршення стану, але не дозволяє ідентифікувати конкретний тип дефекту [24].

Для опису наступних показників введемо загальні позначення. Середнє арифметичне значення сигналу:

$$\bar{x} = \frac{\sum x_i}{N} \quad (2.2)$$

Стандартне відхилення сигналу:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}} \quad (2.3)$$

Максимальне абсолютне значення сигналу:

$$|x|_{max} = \max |x_i|, i = 1, \dots, N \quad (2.4)$$

Куртозис (Kurtosis, K) характеризує гостровершинність розподілу амплітуд сигналу, тобто наскільки розподіл відрізняється від нормального за концентрацією значень навколо середнього та у «хвостах»:

$$K = \frac{\sum (x_i - \bar{x})^4}{N \sigma^4} \quad (2.5)$$

Для нормального розподілу $K = 3$. Поява ударних імпульсів при розвитку дефектів підшипників збільшує куртозис до 6–8 і вище. Куртозис є найбільш чутливим статистичним показником до ранніх стадій пошкоджень, однак при значному розвитку дефекту, коли удари стають безперервними, K знову наближається до 3 [24].

Пік-фактор (Crest Factor, CF) визначається як відношення максимального абсолютного значення сигналу до його середньоквадратичного значення:

$$CF = \frac{|x|_{max}}{RMS} \quad (2.6)$$

Для сигналу з нормальним розподілом амплітуд CF знаходиться в діапазоні від 2,5 до 4,5. Значення $CF > 6$ свідчить про наявність виражених імпульсних складових у сигналі. На відміну від куртозису, пік-фактор зберігає діагностичну інформативність і при розвинених дефектах, коли ударні імпульси стають регулярними [24].

Асиметрія (Skewness, S) характеризує несиметричність розподілу амплітуд відносно середнього значення:

$$S = \frac{\sum (x_i - \bar{x})^3}{N \sigma^3} \quad (2.7)$$

Для симетричного розподілу $S = 0$. Ненульова асиметрія може свідчити про односторонній характер механічних ударів або нелінійність пружної системи опори.

Форм-фактор (Shape Factor, SF) характеризує відхилення форми розподілу амплітуд сигналу та визначається як відношення середньоквадратичного значення

до середнього абсолютного:

$$SF = \frac{N * RMS}{\sum |x_i|} \quad (2.8)$$

Імпульс-фактор (Impulse Factor, IF) визначається як відношення максимального абсолютного значення до середнього абсолютного та характеризує виразність поодиноких імпульсів у сигналі:

$$IF = \frac{N * |x|_{max}}{\sum |x_i|} \quad (2.9)$$

Підвищення IF свідчить про наявність нерегулярних ударних навантажень, характерних для ранніх стадій руйнування підшипників або послаблення кріплень [25].

Порівняльна характеристика статистичних показників за діагностичною інформативністю наведена у таблиці 2.1.

Таблиця 2. 1 Діагностична інформативність статистичних показників часової області

Показник	Формула	Чутливість	Обмеження
RMS	$RMS = \sqrt{\frac{\sum x_i^2}{N}}$	Загальний рівень вібрацій	Не розрізняє типи дефектів
Куртозис(K)	$K = \frac{\sum (x_i - \bar{x})^4}{N \sigma^4}$	Ранні стадії підшипників	Знижується при розвиненому дефекті
Пік-фактор(CF)	$CF = \frac{ x _{max}}{RMS}$	Імпульсні компоненти	Залежить від навантаження
Асиметрія (S)	$S = \frac{\sum (x_i - \bar{x})^3}{N * \sigma^3}$	Несиметричні удари	Низька специфічність
Форм-фактор(SF)	$SF = \frac{N * RMS}{\sum x_i }$	Зміна форми сигналу	Низька специфічність

Продовження таблиці 2.1

Показник	Формула	Чутливість	Обмеження
Імпульс-фактор(IF)	$IF = \frac{N * x _{max}}{\sum x_i }$	Нерегулярні удари	Чутливий до шуму

Для класифікації стану обладнання класичними методами необхідно перетворити вібраційний сигнал у числовий вектор ознак – компактний набір кількісних характеристик, за якими класифікатор розрізняє типи стану. Таке перетворення забезпечує стиснення інформації: замість тисяч дискретних відліків сигналу класифікатор працює з вектором із невеликої кількості числових значень, кожне з яких відображає певну властивість сигналу.

Вектор ознак може включати як показники часової області (формули 2.1–2.9), так і характеристики частотної області. Інформативною частотною ознакою є середня частота спектра, яка характеризує розподіл енергії сигналу у частотній області:

$$f_c = \frac{\sum (f_k) * S[k]}{\sum S[k]} \quad (2.10)$$

де $S[k]$ – значення спектру потужності на k -му частотному відліку; f_k – відповідна частота. Зміщення середньої частоти у бік вищих значень може свідчити про розвиток механічних дефектів, при яких зростає інтенсивність високочастотних вібраційних компонент.

Об'єднаний вектор ознак формується шляхом конкатенації часових та частотних показників:

$$f = [f_{time} | f_{freq}] \quad (2.11)$$

де f_{time} – вектор статистичних показників часової області; f_{freq} – вектор частотних ознак, що може включати амплітуди спектру на характеристичних частотах дефектів (підрозділ 1.3.1), середню частоту спектра та інші спектральні параметри; символ «|» позначає конкатенацію (послідовне об'єднання) двох векторів. Конкретний склад та розмірність вектора визначаються на етапі

експериментального дослідження залежно від діагностичної задачі та характеристик вхідних даних (підрозділ 3.4).

Оскільки показники часової та частотної областей мають різну фізичну природу та різний числовий діапазон, перед класифікацією необхідна нормалізація ознак. Без нормалізації ознаки з більшим числовим діапазоном домінуватимуть при обчисленні відстані між зразками, а ознаки з малим діапазоном фактично не впливатимуть на результат. Для вирівнювання масштабів застосовується метод *z-score*, що приводить кожну ознаку до нульового середнього та одиничної дисперсії:

$$\tilde{f}_i = \frac{(f_i - \mu_i)}{\sigma_i} \quad (2.12)$$

де f_i – значення i -ї ознаки; μ_i та σ_i – середнє значення та стандартне відхилення i -ї ознаки. Параметри μ_i та σ_i обчислюються виключно на навчальній вибірці та застосовуються як до навчальних, так і до тестових даних. Така процедура запобігає витоку інформації з тестової вибірки на етап навчання, що є обов'язковою вимогою коректного оцінювання якості класифікації.

Для класифікації стану обладнання за сформованим та нормалізованим вектором ознак у даній роботі як базовий метод порівняння використано алгоритм k найближчих сусідів (*k-Nearest Neighbors*, *k-NN*). Вибір цього алгоритму обумовлений його простотою, інтерпретованістю результатів та відсутністю етапу навчання у класичному розумінні, що дозволяє використовувати його як нижню межу якості для подальшого порівняння з нейромережевими моделями.

Принцип роботи *k-NN* полягає у прямому порівнянні нового зразка з усіма зразками навчальної вибірки. Нехай ϵ навчальна вибірка, що складається з N зразків, кожен з яких описується вектором ознак f_i . Для класифікації нового зразка з вектором ознак f_{new} алгоритм обчислює відстань від f_{new} до кожного зразка навчальної вибірки f_i за евклідовою метрикою:

$$d(f_{new}, f_i) = \sqrt{\sum (f_{new}[j] - f_i[j])^2}, j = 1 \dots N_f \quad (2.13)$$

де d – евклідова відстань, що визначає міру подібності між двома зразками: чим менше значення d , тим більш схожими є зразки за своїми ознаками; f_{new} – вектор ознак нового зразка, клас якого потрібно визначити; f_i – вектор ознак i -го

зразка навчальної вибірки з відомим класом; $f_{new}[j]$ та $f_i[j]$ – значення j -ї ознаки відповідних векторів; N_f – загальна розмірність вектора ознак, тобто кількість числових характеристик, за якими порівнюються зразки; підсумовування виконується за всіма ознаками від $j = 1$ до N_f .

Серед N зразків обираються k найближчих (з найменшими відстанями). Кожен зразок навчальної вибірки має відому мітку класу y_i , і новому зразку присвоюється клас, що зустрічається найчастіше серед k сусідів за правилом більшості голосів:

$$\hat{y} = \underset{c}{\operatorname{argmax}} \sum \mathbb{I}(y_i = c), i \in N_k(f_{new}) \quad (2.14)$$

де $\hat{y}$ – прогнозований клас нового зразка; $\underset{c}{\operatorname{argmax}}$ – операція вибору такого значення класу c , при якому вираз у правій частині досягає максимуму; c – номер класу, що пробігає значення від 0 до кількості класів мінус один; Σ – сума за всіма індексами i , що належать до множини $N_k(f_{new})$; $N_k(f_{new})$ – множина індексів k найближчих сусідів нового зразка f_{new} , визначених за формулою 2.13; y_i – відома мітка класу i -го сусіда; $\mathbb{I}(y_i = c)$ – індикаторна функція, що дорівнює 1, якщо клас i -го сусіда збігається з c , та 0 – якщо ні. Таким чином, для кожного класу c формула підраховує кількість сусідів цього класу серед k найближчих, і новому зразку присвоюється клас з найбільшою кількістю голосів.

Алгоритм k -NN є непараметричним: він не будує внутрішньої моделі даних, а зберігає всю навчальну вибірку у пам'яті та звертається до неї при кожній класифікації. Значення параметра k визначає компроміс між стійкістю до шуму та чутливістю до локальних структур у даних: при малих k алгоритм точніше відтворює межі між класами, але є більш чутливим до шумових зразків; при великих k – стійкіший до шуму, але може втрачати здатність розрізняти класи з близькими ознаками.

Суттєвим обмеженням k -NN є його чутливість до розмірності вектора ознак: зі збільшенням кількості ознак відстані між зразками вирівнюються, і здатність алгоритму розрізняти сусідів знижується. Це обмеження, відоме як прокляття розмірності, обумовлює необхідність ретельного відбору найбільш інформативних

ознак для формування вектора. Крім того, обчислювальна складність класифікації зростає лінійно з обсягом навчальної вибірки, оскільки для кожного нового зразка необхідно обчислити відстань до всіх навчальних прикладів.

Важливою умовою коректної роботи k -NN є попередня нормалізація ознак (формула 2.12): без неї ознаки з більшим числовим діапазоном домінуватимуть у обчисленні відстані, що призведе до ігнорування потенційно інформативних ознак з малим діапазоном значень. Вибір конкретного значення k , складу вектора ознак та результати класифікації наведено у підрозділі 3.4.

Таким чином, статистичні показники часової області забезпечують швидку загальну оцінку стану обладнання, проте мають суттєве обмеження: вони характеризують лише амплітудні властивості сигналу і не містять інформації про його частотний склад. Різні типи дефектів можуть спричиняти подібні зміни амплітудних показників, що ускладнює їх диференціацію лише на основі часового аналізу. Для ідентифікації конкретних джерел вібрацій за характеристичними частотами дефектів необхідний перехід до частотної області, математичний апарат якого розглянуто у підрозділі 2.1.2.

2.1.2. Дискретна реалізація FFT та практичні аспекти спектрального аналізу

Статистичні показники часової області, розглянуті у підрозділі 2.1.1, характеризують загальні амплітудні властивості вібраційного сигналу, проте не дають інформації про його частотний склад. Для ідентифікації джерел вібрацій за характеристичними частотами дефектів необхідний перехід до частотної області.

Теоретичною основою такого переходу є перетворення Фур'є, яке представляє сигнал як суперпозицію гармонічних складових різних частот та амплітуд. Для неперервного сигналу перетворення Фур'є визначається як [26]:

$$X(f) = \int x(t)e^{-j2\pi ft} dt, \quad (2.15)$$

де $X(f)$ – комплексний спектр сигналу, що описує амплітуду та фазу гармонічної складової на частоті f ; $x(t)$ – вхідний сигнал у часовій області; f – частота (Гц); j – уявна одиниця ($j^2 = -1$); t – час (с); інтегрування виконується за всією тривалістю сигналу.

Оскільки вібраційний сигнал реєструється акселерометрами у цифровій

формі як послідовність дискретних відліків, для його аналізу застосовується дискретне перетворення Фур'є (ДПФ):

$$X[k] = \sum x[n] e^{\frac{-j2\pi kn}{N}}, n = 0, 1, \dots, N-1 \quad (2.16)$$

де $X[k]$ – комплексний спектр на k -му частотному відліку, що описує амплітуду та фазу гармонічної складової на відповідній частоті; $x[n]$ – дискретний вхідний сигнал (значення n -го відліку віброприскорення); N – загальна кількість відліків сигналу; k – індекс частотного відліку ($k = 0, 1, \dots, N-1$); j – уявна одиниця.

Кожному індексу k відповідає фізична частота:

$$f_k = \frac{k f_s}{N}, \quad (2.17)$$

де f_s – частота дискретизації (Гц). Мінімальна відстань між двома частотами, які можна розрізнити у спектрі, визначається частотною роздільною здатністю:

$$\Delta f = \frac{f_s}{N} \quad (2.18)$$

Збільшення тривалості аналізованого фрагменту (тобто кількості відліків N) підвищує частотну роздільну здатність, що є важливим для розрізнення близько розташованих характеристичних частот дефектів.

Пряме обчислення ДПФ за формулою (2.16) потребує порядку N^2 арифметичних операцій, що при великих обсягах даних є обчислювально затратним. Алгоритм швидкого перетворення Фур'є (Fast Fourier Transform, FFT) дає математично ідентичний результат, але потребує лише порядку $N \log_2 N$ операцій. Наприклад, для сигналу з 4 096 відліків пряме обчислення ДПФ потребує понад 16 мільйонів операцій, тоді як FFT виконує ту саму задачу приблизно за 49 тисяч операцій [26]. Саме завдяки цьому FFT є стандартним інструментом спектрального аналізу.

Для кількісної оцінки розподілу енергії сигналу за частотами використовується спектр потужності (Power Spectral Density, PSD):

$$S[k] = \frac{|X[k]|^2}{N}, \quad (2.19)$$

де $S[k]$ – значення спектру потужності на k -му частотному відліку; $|X[k]|^2$ – квадрат модуля комплексного спектру, що відповідає енергії сигналу на частоті f_k ;

N – кількість відліків. Значення $S[k]$ на характеристичних частотах дефектів (f_r , $2f_r$, BPFO, BPFI та інших, визначених у підрозділі 1.3.1) використовуються як частотні ознаки діагностичного вектора f_{freq} (формула 2.11). Також на основі $S[k]$ обчислюється середня частота спектра f_c (формула 2.10).

Важливою особливістю дискретного спектрального аналізу є ефект витоку спектру (spectral leakage). При обчисленні ДПФ аналізується фрагмент сигналу кінцевої тривалості, що фактично еквівалентно множенню нескінченного сигналу на прямокутне вікно – функцію, яка дорівнює одиниці в межах фрагменту і нулю поза ним. У частотній області таке множення відповідає згортці спектру сигналу зі спектром прямокутного вікна. Оскільки спектр прямокутного вікна містить значні бокові пелюстки, енергія кожної частотної складової «витікає» на сусідні частотні відліки: піки у спектрі розмиваються, а слабкі компоненти можуть бути замасковані боковими пелюстками потужних компонент. Для задач вібраційної діагностики це є критичним, оскільки амплітуди дефектних компонент (наприклад, BPFO) можуть бути на 40–80 дБ нижчими за амплітуди основних гармонік ротора.

Для зменшення витоку спектру перед обчисленням ДПФ сигнал множиться на спеціальну вікнову функцію $w[n]$, яка плавно зводить амплітуду сигналу до нуля на краях фрагменту:

$$x_w[n] = x[n] * w[n], \quad (2.20)$$

де $x[n]$ – вхідний дискретний сигнал; $w[n]$ – вікнова функція довжиною N , значення якої плавно зростають від нуля на початку фрагменту до максимуму в центрі та знову спадають до нуля на кінці; $x_w[n]$ – сигнал після застосування вікна, який подається на вхід ДПФ замість вихідного сигналу $x[n]$.

Вибір конкретної вікнової функції визначає компроміс між двома характеристиками: шириною головної пелюстки (що впливає на частотну роздільну здатність – здатність розрізняти дві близькі частоти) та рівнем бокових пелюсток (що визначає динамічний діапазон – здатність виявляти слабкі компоненти поруч із потужними). Вузька головна пелюстка забезпечує кращу роздільну здатність, але високі бокові пелюстки можуть маскувати слабкі складові.

Широка головна пелюстка знижує роздільну здатність, але суттєво пригнічує бокові пелюстки. Порівняння основних вікнових функцій наведено у таблиці 2.2.

Таблиця 2. 2 Порівняльна характеристика вікнових функцій

Вікно	Ширина головної пелюстки (відн.)	Рівень бокових пелюсток (дБ)	Динамічний діапазон (дБ)
Прямокутне	1,0	−13	13
Хеннінга	1,6	−31	31
Хеммінга	1,4	−43	43
Блекмана	1,7	−58	58
Блекмана-Харріса	2,0	−92	92

Прямокутне вікно (фактично відсутність вікна) має найвужчу головну пелюстку, але найвищий рівень бокових пелюсток (−13 дБ), що робить його непридатним для аналізу сигналів з компонентами різної амплітуди. Вікно Блекмана-Харріса забезпечує найвищий динамічний діапазон – 92 дБ, що є критично важливим при різниці амплітуд між гармоніками ротора та дефектними компонентами у 60–80 дБ. Ширша головна пелюстка цього вікна дещо знижує частотну роздільну здатність, проте для задач класифікації спектрограм нейронною мережею це є прийнятним компромісом [26]. Вікнова функція застосовується як при звичайному спектральному аналізі FFT, так і при формуванні спектрограм STFT (підрозділ 2.1.3), де кожен часовий фрагмент сигналу множиться на вікно перед обчисленням спектру.

Теорема Котельникова-Найквіста встановлює нижню межу частоти дискретизації для коректного відтворення сигналу: $f_s \geq 2f_{max}$, де f_{max} – максимальна частота, присутня у сигналі. При недотриманні цієї умови виникає ефект аліасингу – накладання високочастотних компонент на низькочастотні, що спотворює спектр і унеможлиблює коректну ідентифікацію характеристичних частот дефектів. Для запобігання аліасингу перед аналогово-цифровим перетворенням встановлюється антиаліасинговий фільтр нижніх частот із

частотою зрізу $\frac{f_s}{2}$ [26].

При об'єднанні даних, зареєстрованих з різними частотами дискретизації, необхідна процедура передискретизації (resampling) – приведення сигналів до єдиної цільової частоти f_s^* . Передискретизація з раціональним коефіцієнтом $\frac{p}{q}$ реалізується послідовною інтерполяцією вхідного сигналу у p разів та децимацією у q разів із застосуванням антиаліасингового фільтра нижніх частот для запобігання спектральним спотворенням. Математично передискретизований сигнал визначається як:

$$a^*[m] = \sum a[n] h_{lp}[mp - nq], n = 0, 1, \dots, L - 1, \quad (2.21)$$

де $a[n]$ – вхідний сигнал довжиною L відліків; $a^*[m]$ – вихідний сигнал довжиною $L^* = \left\lfloor L \frac{p}{q} \right\rfloor$; h_{lp} – імпульсна характеристика антиаліасингового фільтра нижніх частот із частотою зрізу $f_c = \min\left(\frac{1}{p}, \frac{1}{q}\right) \frac{f_s}{2}$. Коефіцієнти p та q визначаються як $p = \frac{f_s^*}{\text{НСД}(f_s, f_s^*)}$ та $q = \frac{f_s}{\text{НСД}(f_s, f_s^*)}$, де f_s – вихідна частота дискретизації, НСД – найбільший спільний дільник [26]. Практичне застосування процедури передискретизації до узгодження даних із двох датасетів описано у підрозділі 3.2.

Таким чином, перетворення Фур'є та його дискретна реалізація FFT забезпечують перехід від часового представлення вібраційного сигналу до частотного, що дозволяє ідентифікувати характеристичні частоти дефектів та формувати частотні ознаки діагностичного вектора. Проте результатом FFT є усереднений спектр за всю тривалість сигналу, що не містить інформації про часову динаміку частотних компонент. Для одночасного аналізу часової та частотної структури сигналу у підрозділі 2.1.3 розглянуто короткочасне перетворення Фур'є (STFT).

2.1.3. Розширений математичний апарат STFT для формування спектрограм

Спектральний аналіз на основі FFT, розглянутий раніше, дає усереднений спектр за всю тривалість сигналу і не відображає зміну частотного складу у часі. Для одночасного аналізу часової та частотної структури вібраційного сигналу застосовується короткочасне перетворення Фур'є (Short-Time Fourier Transform,

STFT).

Перед обчисленням STFT вхідний сигнал може потребувати попередньої нормалізації амплітуд. Якщо дані надходять з різних вимірювальних систем, абсолютні значення віброприскорення можуть відрізнятися навіть для аналогічних режимів роботи обладнання внаслідок різної чутливості та калібрування акселерометрів. Для усунення цього впливу кожен фрагмент сигналу нормалізується методом z-score:

$$\tilde{a}[n] = \frac{a[n] - \mu_a}{\sigma_a} \quad (2.22)$$

де $a[n]$ – вхідний дискретний сигнал довжиною N відліків; $\mu_a = \frac{1}{N} \sum a[n]$ – середнє значення сигналу; σ_a – стандартне відхилення; $\tilde{a}[n]$ – нормалізований сигнал із нульовим середнім та одиничною дисперсією. На відміну від нормалізації вектора ознак, де параметри обчислюються для кожної ознаки по всій вибірці, тут параметри μ_a та σ_a обчислюються окремо для кожного фрагменту сигналу, що дозволяє коректно об'єднувати дані, зареєстровані різними вимірювальними системами. Практичне застосування цієї процедури описано у підрозділі 3.2.

Для неперервного сигналу STFT визначається як [26]:

$$STFT\{x\}(\tau, f) = \int x(t) * w(t - \tau) * e^{-j2\pi f t} dt \quad (2.23)$$

де $x(t)$ – вхідний сигнал; $w(t - \tau)$ – вікнова функція, зсунута на час τ , що локалізує аналіз у обмеженому часовому інтервалі навколо моменту τ ; f – частота; j – уявна одиниця. Суть STFT полягає у тому, що замість обчислення перетворення Фур'є для всього сигналу одразу, воно обчислюється для послідовних коротких фрагментів, кожен з яких виділяється вікновою функцією.

У дискретному випадку STFT обчислюється для послідовності перекривних фрагментів сигналу:

$$X[m, k] = \sum x[n] * w[n - mH] * e^{-\frac{j2\pi kn}{N_w}}, n = 0, 1, \dots, N_w - 1, \quad (2.24)$$

де $X[m, k]$ – комплексний спектр m -го часового вікна на k -му частотному відліку; $x[n]$ – дискретний вхідний сигнал; $w[n - mH]$ – вікнова функція, зсунута на mH відліків; m – індекс часового вікна ($m = 0, 1, 2, \dots$); H – крок зсуву вікна

(hop size) у кількості відліків, що визначає наскільки кожне наступне вікно зміщене відносно попереднього; N_w – довжина вікна у відліках; k – індекс частотного відліку. Вікнова функція $w[n]$ обирається відповідно до критеріїв, обґрунтованих у підрозділі 2.1.2 (таблиця 2.2).

Загальна кількість часових вікон для сигналу довжиною L відліків:

$$M = \left\lfloor \frac{L - N_w}{H} \right\rfloor + 1, \quad (2.25)$$

де L – довжина сигналу у відліках; N_w – довжина вікна; H – крок зсуву; $\lfloor \cdot \rfloor$ – операція округлення донизу до цілого числа.

Для кожного часового вікна обчислюється спектр потужності, що утворює двовимірний масив – спектрограму:

$$S[m, k] = |X[m, k]|^2, \quad (2.26)$$

де $S[m, k]$ – значення спектру потужності на частоті f_k у момент часу t_m ; $|X[m, k]|^2$ – квадрат модуля комплексного спектру. Вісь m відповідає часу, вісь k – частоті, а значення $S[m, k]$ – енергії сигналу на відповідній частоті у відповідний момент часу.

Значення спектру потужності можуть відрізнятися на кілька порядків: амплітуда основної гармоніки ротора може перевищувати амплітуду дефектної компоненти у сотні або тисячі разів. Для стиснення цього динамічного діапазону та рівномірного відображення як потужних, так і слабких компонент спектрограма перетворюється до логарифмічної шкали [20]:

$$S_{dB}[m, k] = 10 \log_{10}(S[m, k] + \varepsilon), \quad \varepsilon = 10^{-10}, \quad (2.27)$$

де $S_{dB}[m, k]$ – значення спектру потужності у логарифмічній шкалі (децибели); ε – мала додатна константа, що запобігає обчисленню логарифма від нуля. Логарифмічне масштабування є принципово важливим для подальшої обробки нейронною мережею: без нього мережа сприймає лише домінуючі компоненти спектру, тоді як у логарифмічній шкалі компоненти з амплітудою на 40 дБ нижче максимуму (тобто у 100 разів слабші за абсолютним значенням) зберігаються на спектрограмі та залишаються доступними для розпізнавання.

Для забезпечення інваріантності до абсолютного рівня вібрацій спектрограма

нормалізується до діапазону $[0, 1]$:

$$\check{S}[m, k] = \frac{S_{dB}[m, k] - S_{min}}{S_{max} - S_{min}}, \quad (2.28)$$

де S_{min} та S_{max} – мінімальне та максимальне значення S_{dB} у межах даної спектрограми. Нормалізація виконується окремо для кожної спектрограми, що забезпечує незалежність результату від абсолютного рівня вібрацій конкретного вимірювання. Після нормалізації спектрограми з різних вимірювань мають однаковий діапазон значень і відрізняються лише формою розподілу інтенсивності, що є визначальною для ідентифікації типу несправності.

Результатом описаної послідовності перетворень є нормалізована спектрограма $\check{S}[m, k]$ – двовимірне зображення, що поєднує часову та частотну інформацію про вібраційний сигнал у формі, безпосередньо придатній для обробки згортковими нейронними мережами. Приклад нормалізованої спектрограми STFT вібраційного сигналу електродвигуна наведено на рисунку 2.1.

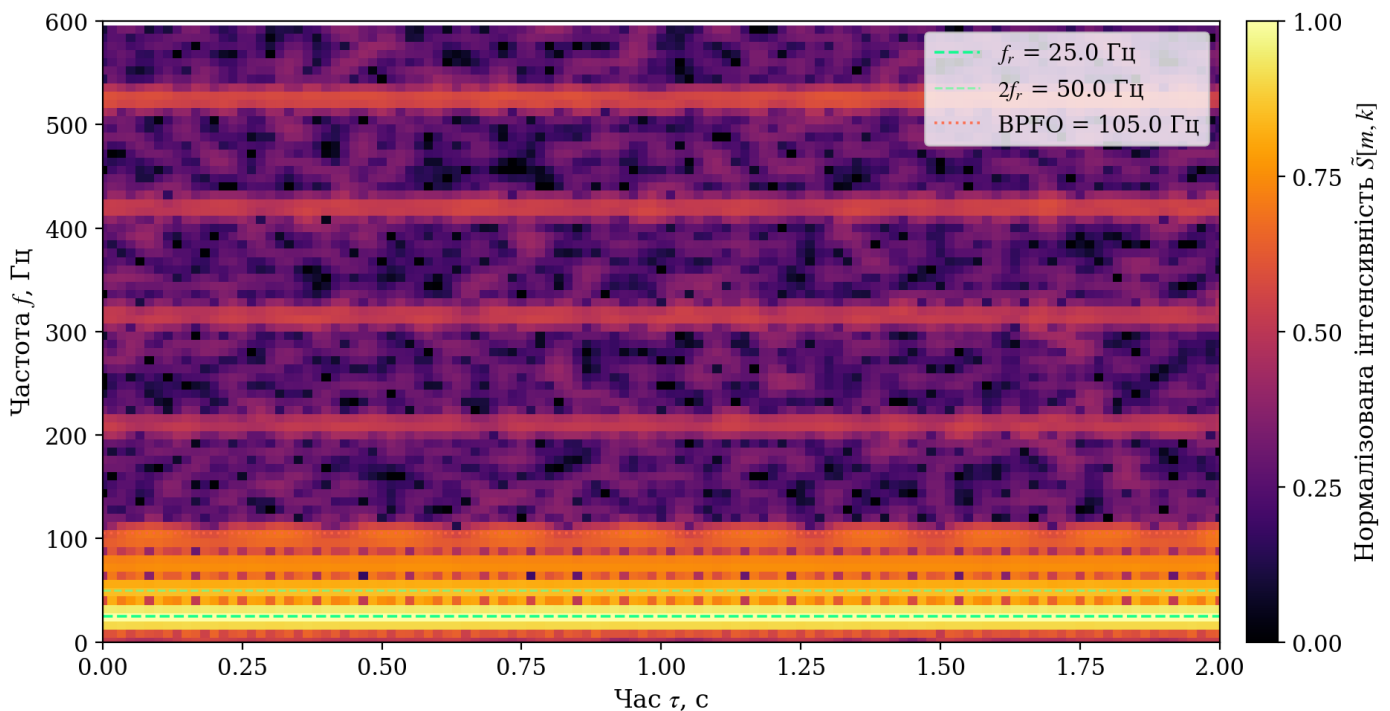


Рисунок 2. 1 Спектрограма STFT вібраційного сигналу електродвигуна

Якість спектрограми суттєво залежить від вибору параметрів STFT – довжини вікна N_w , кроку зсуву H та ступеня перекриття сусідніх вікон. Ці параметри визначають компроміс між часовою та частотною роздільною

здатністю, а їх обґрунтований вибір розглянуто далі.

2.1.4. Обґрунтування та вибір параметрів STFT

Якість спектрограми, сформованої за допомогою STFT, визначається набором параметрів, вибір яких безпосередньо впливає на інформативність вхідних даних для подальшої класифікації. Основними параметрами є довжина вікна N_w , крок зсуву H та ступінь перекриття сусідніх вікон OL .

Фундаментальним обмеженням будь-якого часово-частотного перетворення є принцип невизначеності Гейзенберга-Габора, який встановлює нижню межу добутку часової та частотної роздільності:

$$\sigma_t * \sigma_f \geq \frac{1}{4\pi}, \quad (2.29)$$

де σ_t – невизначеність у часовій області (с); σ_f – невизначеність у частотній області (Гц). Ця нерівність означає, що одночасне підвищення точності визначення і частоти, і моменту часу є принципово неможливим: зменшення невизначеності по одній осі неминуче призводить до збільшення невизначеності по іншій. Для STFT цей компроміс безпосередньо визначається довжиною вікна N_w .

Довжина вікна N_w є ключовим параметром STFT, що визначає баланс між частотною та часовою роздільністю. Частотна роздільність спектрограми обчислюється аналогічно до формули FFT, але з урахуванням того, що перетворення виконується для кожного вікна окремо:

$$\Delta f = \frac{f_s}{N_w}, \quad (2.30)$$

де f_s – частота дискретизації; N_w – довжина вікна у відліках. Зменшення Δf (тобто підвищення частотної роздільності) вимагає збільшення N_w , що одночасно погіршує часову роздільність: довше вікно охоплює більший інтервал часу, і швидкі зміни сигналу всередині вікна усереднюються. Навпаки, коротке вікно забезпечує високу часову роздільність, але не дозволяє розрізнити близько розташовані частотні компоненти.

Крок зсуву H (hop size) визначає, на скільки відліків зміщується вікно при переході від одного часового стовпця спектрограми до наступного. Ступінь перекриття сусідніх вікон OL (overlap) пов'язаний з кроком зсуву та довжиною

вікна:

$$OL = \frac{N_w - H}{N_w}, \quad (2.31)$$

де OL – коефіцієнт перекриття (безрозмірна величина від 0 до 1); N_w – довжина вікна; H – крок зсуву. При $OL = 0$ ($H = N_w$) вікна не перекриваються і кожен відлік сигналу потрапляє лише в одне вікно. При $OL = 0,5$ ($H = \frac{N_w}{2}$) кожен відлік потрапляє у два сусідні вікна. Збільшення перекриття підвищує часову деталізацію спектрограми та зменшує втрату інформації на краях вікон, де вікнова функція послаблює амплітуду сигналу. Проте надмірне перекриття збільшує кількість обчислень та розмір результуючої спектрограми.

Часова роздільність спектрограми визначається інтервалом часу між сусідніми стовпцями:

$$\Delta t = \frac{H}{f_s}, \quad (2.32)$$

де H – крок зсуву (відліки); f_s – частота дискретизації (Гц). Менший крок H (більше перекриття) дає менший Δt , тобто більш детальне відображення зміни спектру у часі.

Загальна тривалість фрагменту сигналу, що відповідає одній спектрограмі:

$$T_{seg} = \frac{N}{f_s}, \quad (2.33)$$

де N – загальна кількість відліків фрагменту; f_s – частота дискретизації. Кількість часових вікон M визначається за формулою 2.25, при цьому кожне вікно відповідає одному стовпцю спектрограми. Кількість частотних рядків дорівнює $\left(\frac{N_w}{2} + 1\right)$ (від нульової частоти до частоти Найквіста $\frac{f_s}{2}$).

При виборі параметрів STFT для задач діагностики необхідно враховувати наступні чинники. По-перше, частотна роздільність Δf має бути достатньою для розрізнення характеристичних частот дефектів, визначених у підрозділі 1.3.1, зокрема частоти обертання ротора та її гармонік, частот підшипникових дефектів BPFO та BPFI, а також бокових смуг модуляції. По-друге, часова роздільність Δt має забезпечувати виявлення перехідних процесів та імпульсних подій, характерних для певних типів дефектів. По-третє, розмір результуючої

спектрограми визначає обсяг вхідних даних для подальшого аналізу та впливає на обчислювальні витрати. По-четверте, при фіксованій частоті дискретизації f_s аналіз усього частотного діапазону від 0 до $\frac{f_s}{2}$ не завжди є доцільним: якщо характеристичні частоти дефектів зосереджені у вузькому діапазоні, обмеження робочої смуги дозволяє підвищити інформативність спектрограми за рахунок концентрації роздільної здатності у діагностично значущому діапазоні.

Конкретні числові значення параметрів STFT, обрані для даного дослідження з урахуванням характеристик вхідних даних та вимог задачі класифікації, наведено у підрозділі 3.3.

2.2. Архітектура згорткових нейронних мереж

Згорткові нейронні мережі (Convolutional Neural Networks, CNN) є класом глибоких нейронних мереж, що спеціалізуються на обробці даних із просторовою структурою – зображень, спектрограм, карт. У контексті даної роботи вхідними даними для CNN є двоканальні спектрограми STFT, сформовані за методикою підрозділу 2.1.3, де два канали відповідають радіальному та осьовому напрямкам вібрації.

Базовим елементом будь-якої нейронної мережі є штучний нейрон – математична функція, що обчислює зважену суму вхідних значень, додає зміщення (bias) та застосовує до результату нелінійну функцію активації. Саме вибір функції активації визначає тип поведінки нейрона та його роль в архітектурі мережі. Організація нейронів у шари та спосіб з'єднання шарів між собою визначає архітектуру мережі.

Ключовою ідеєю CNN є те, що корисні ознаки зображення є локальними та просторово інваріантними [28]. Локальність означає, що для виявлення патерну достатньо аналізувати невелику ділянку зображення, а не все зображення цілком. Просторова інваріантність означає, що один і той самий патерн має бути розпізнаний незалежно від його положення на зображенні. Саме ці властивості реалізуються операцією згортки та принципом спільності ваг, що відрізняє CNN від повнозв'язних мереж.

У даному підрозділі послідовно розглядаються всі компоненти архітектури CNN: операція згортки, функції активації, шари пулінгу, методи нормалізації та регуляризації, а також багатомасштабна модифікація MSCNN.

2.2.1. Операція згортки та принципи побудови CNN

Вхідні дані для CNN – двовимірне зображення, що представляється у вигляді тривимірного тензора (багатовимірного масиву чисел) розміром $H \times W \times C$, де H – висота зображення у пікселях, W – ширина, C – кількість каналів. Для спектрограм STFT: H – кількість частотних рядків, W – кількість часових стовпців, C – кількість каналів вібрації (у даній роботі $C = 2$: радіальний та осьовий). Для звичайних кольорових зображень $C = 3$ (канали R, G, B) [28].

Операція згортки є базовою математичною операцією в CNN. Її суть полягає у наступному: невелика матриця (фільтр або ядро) розміром $k \times k$ послідовно переміщується по вхідному зображенню з кроком s (stride). У кожному положенні обчислюється поелементне множення фільтра на відповідну ділянку зображення та підсумовування результатів. Формально вихідне значення визначається як:

$$Y[i, j] = (\sum \sum X[i * s + m, j * s + n] * K[m, n]) + b, m, n = 0 \dots k - 1, \quad (2.34)$$

де X – вхідне зображення (або карта ознак попереднього шару); K – матриця фільтра (ядра згортки) розміром $k \times k$; b – зміщення (bias), константа що додається до кожного результату; s – крок (stride) переміщення фільтра; $Y[i, j]$ – значення вихідної карти ознак у позиції (i, j) ; m та n – індекси переміщення всередині фільтра. Ваги фільтра K та зміщення b є параметрами мережі, що визначаються автоматично у процесі навчання.

Розмір вихідної карти ознак залежить від режиму обробки країв зображення. При використанні доповнення нулями (padding = 'same') навколо вхідного зображення додаються рядки та стовпці з нульовими значеннями, що дозволяє фільтру проходити по всіх пікселях включаючи крайні:

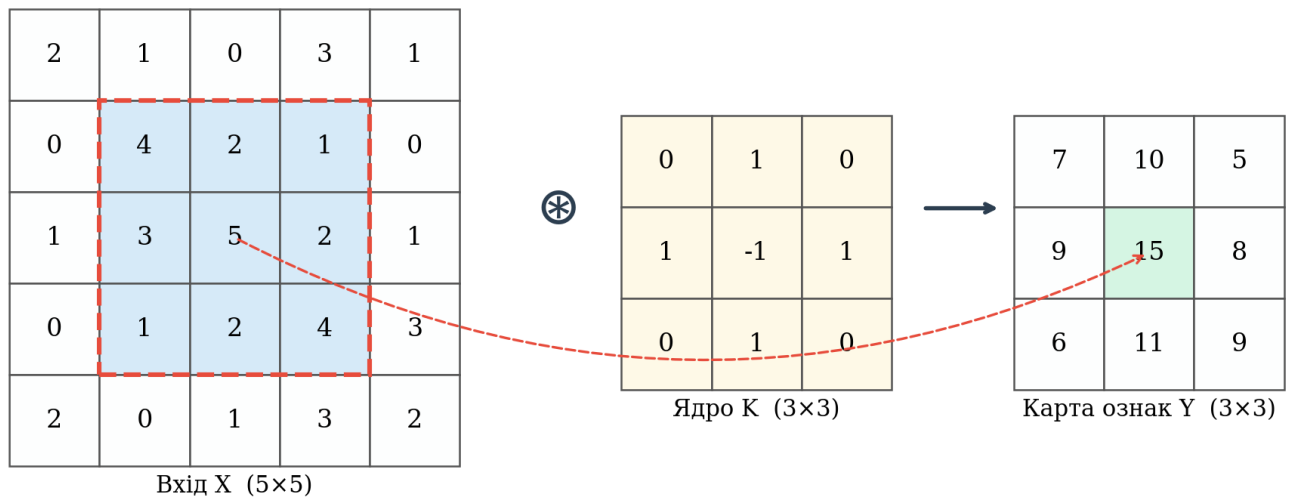
$$H_{out} = H_{in}, W_{out} = W_{in} \text{ (при padding = 'same' та } s = 1) \quad (2.35)$$

Без доповнення нулями (padding = 'valid') фільтр не може розташовуватись на краях зображення, і розмір виходу зменшується:

$$H_{out} = \left\lfloor \frac{H_{in}-k}{s} \right\rfloor + 1, W_{out} = \left\lfloor \frac{W_{in}-k}{s} \right\rfloor + 1, \quad (2.36)$$

де H_{in} , W_{in} – розміри вхідного зображення; k – розмір фільтра; s – крок; $\lfloor \cdot \rfloor$ – округлення донизу. Наприклад, для вхідного зображення 128×128 та фільтра 3×3 без доповнення: $H_{out} = (128 - 3)/1 + 1 = 126$. З доповненням нулями розмір залишається 128×128 .

Один згортковий шар містить не один, а F фільтрів – кожен з яких є незалежним детектором, що навчається виявляти свій тип ознаки (наприклад, горизонтальну лінію, вертикальний імпульс, діагональний перехід). Кожен фільтр обробляє всі вхідні канали і створює одну карту ознак (feature map). Таким чином, вихідний тензор одного згорткового шару має розміри $H_{out} \times W_{out} \times F$. Принцип операції згортки проілюстровано на рисунку 2.2.



$$Y[i,j] = \sum_m \sum_n X[i \cdot s + m, j \cdot s + n] \cdot K[m,n] + b$$

Рисунок 2. 2 Принцип операції двовимірної дискретної згортки

Важливою властивістю операції згортки є спільність ваг (weight sharing): один і той самий фільтр з одним і тим самим набором ваг застосовується до всіх позицій зображення. Це означає, що кількість параметрів залежить від розміру фільтра та кількості каналів, а не від розміру зображення. Загальна кількість навчальних параметрів одного згорткового шару:

$$P_{conv} = F(k * k * C_{in} + 1), \quad (2.37)$$

де F – кількість фільтрів; k – розмір ядра; C_{in} – кількість вхідних каналів;

доданок 1 відповідає зміщенню b для кожного фільтра. Наприклад, шар з $F = 32$ фільтрами розміром $k = 3$ для двоканального входу ($C_{in} = 2$): $P_{conv} = 32 \cdot (3 \cdot 3 \cdot 2 + 1) = 608$ параметрів. Це значно менше ніж повнозв'язний шар з аналогічним входом, що забезпечує ефективність CNN на великих зображеннях.

Принцип вибору кількості шарів та фільтрів визначається задачею. Кожен наступний згортковий шар виділяє ознаки вищого рівня абстракції: перші шари розпізнають найпростіші елементи (краї, градієнти інтенсивності), середні – їх комбінації (текстури, локальні патерни), глибокі – складні структури, характерні для конкретних типів несправностей. Кількість фільтрів зазвичай збільшується від шару до шару (наприклад, $32 \rightarrow 64 \rightarrow 128$), оскільки глибші шари потребують більше детекторів для відображення більшої різноманітності складних ознак [28].

Функція активації застосовується поелементно до виходу кожного згорткового шару і вносить нелінійність у модель. Без нелінійної активації послідовність згорткових шарів математично еквівалентна одному шару, що унеможливило б виділення складних ознак. У прихованих шарах CNN стандартною є функція ReLU (Rectified Linear Unit) [28]:

$$ReLU(x) = \max(0, x) \quad (2.38)$$

ReLU залишає без змін усі позитивні значення (які відповідають «виявленому» патерну) і замінює нулем усі від'ємні (відсутність патерну). Переваги ReLU: обчислювальна простота (одне порівняння), відсутність проблеми зникнення градієнта для позитивних значень, розрідженість активацій (частина нейронів має нульовий вихід, що діє як природна регуляризація).

Повнозв'язний шар (Fully Connected, FC) використовується у кінцевій частині мережі для об'єднання просторових ознак, виділених згортковими шарами, у єдиний вектор рішення. Кожен нейрон повнозв'язного шару з'єднаний з усіма елементами вхідного вектора:

$$z_j = \left(\sum w_{ij} x_i \right) + b_j, \quad i = 1 \dots N_{in}, \quad (2.39)$$

де x_i – елементи вхідного вектора ознак; w_{ij} – ваги з'єднань між i -м входом та j -м нейроном; b_j – зміщення нейрона j ; z_j – вихідне значення (логіт) j -го нейрона;

N_{in} – розмірність вхідного вектора.

Для перетворення вектора логітів у вектор ймовірностей класів застосовується функція Softmax. Вона гарантує, що всі вихідні значення додатні та їх сума дорівнює одиниці:

$$p_i = \frac{e^{z_i}}{\sum e^{z_j}}, j = 1 \dots N_c, \quad (2.40)$$

де z_i – логіт для класу i ; e – основа натурального логарифма; N_c – кількість класів. Компонент p_i інтерпретується як ступінь впевненості мережі в тому, що вхідний зразок належить до класу i . Клас з найбільшим p_i обирається як прогнозований.

2.2.2. Шари пулінгу, нормалізації та регуляризації

Шар максимального пулінгу (Max Pooling) є операцією просторового стиснення карти ознак. Вхідна карта ознак ділиться на непересічні блоки розміром $p \times p$ пікселів (найчастіше $p = 2$), і з кожного блоку у вихідну карту записується лише максимальне значення. Таким чином вихідна карта виявляється у p^2 разів меншою просторово: при $p = 2$ вхідна карта розміром $H \times W$ стискається до $\left(\frac{H}{2}\right) \times \left(\frac{W}{2}\right)$ [29]:

$$Y_{max}[i, j] = \max\{X[i * p + m, j * p + n]\}, m, n = 0 \dots (p - 1), \quad (2.41)$$

де X – вхідна карта ознак; Y_{max} – вихідна зменшена карта ознак; i та j – індекси позиції у вихідній карті ознак; p – розмір вікна пулінгу; m та n – індекси переміщення всередині вікна.

Ця операція виконує одразу кілька важливих функцій. По-перше, вона скорочує обсяг обчислень: після пулінгу 2×2 кожен наступний шар обробляє вчетверо менше значень. По-друге, вона забезпечує часткову інваріантність до просторових зсувів: якщо певний патерн зсунувся на 1 піксель, максимум у вікні 2×2 все одно буде виявлений. По-третє, зменшення просторових розмірів при незмінній кількості каналів збільшує відносне рецептивне поле нейронів у глибших шарах – кожен нейрон фактично «бачить» все більшу ділянку вхідного зображення.

Глобальний середній пулінг (Global Average Pooling, GAP) є окремим

випадком, коли вікно охоплює всю карту ознак розміром $H \times W$. Для кожного каналу c обчислюється середнє значення по всіх $H \cdot W$ просторових позиціях:

$$Y_{gap}[c] = \frac{1}{HW} \sum \sum X[i, j, c], i = 0 \dots (H - 1), j = 0 \dots (W - 1), \quad (2.42)$$

де $Y_{gap}[c]$ – значення виходу для каналу c ; X – вхідний тензор розміром $H \times W \times C$; i та j – індекси просторових позицій у карті ознак; H та W – просторові розміри карти ознак (висота та ширина); C – кількість каналів; c – індекс каналу; $\frac{1}{HW}$ – нормалізація на загальну кількість просторових позицій.

GAP перетворює тензор $H \times W \times C$ у вектор із C елементів – по одному числу на кожен канал. Наприклад, тензор розміром $8 \times 8 \times 256$ стискається до вектора із 256 елементів. Ця операція замінює операцію вирівнювання (flatten) разом з великим повнозв'язним шаром, що значно зменшує кількість параметрів і ризик перенавчання [30].

Пакетна нормалізація (Batch Normalization, BN) – метод стабілізації і прискорення навчання глибоких нейронних мереж. Проблема, яку він вирішує, полягає у наступному: коли дані проходять через кілька шарів мережі, розподіл значень активацій може суттєво змінюватись від шару до шару і від пакету до пакету. Це явище уповільнює навчання і вимагає дуже малих швидкостей навчання. BN нормалізує активації кожного шару по поточному міні-пакету навчальних зразків (міні-пакет, або batch – група з B зразків, що обробляється одночасно) [29]:

$$\mu_B = \frac{1}{B} \sum x_i; \sigma_B^2 = \frac{1}{B} \sum (x_i - \mu_B)^2 \quad (2.43)$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}; y_i = \gamma \hat{x}_i + \beta, \quad (2.44)$$

де B – розмір міні-пакету (кількість зразків у групі); x_i – вхідне значення для i -го зразка ($i = 1, \dots, B$); μ_B – середнє значення по пакету; σ_B^2 – дисперсія по пакету; підсумовування у обох формулах виконується за всіма зразками міні-пакету від $i = 1$ до B ; $\epsilon \approx 10^{-5}$ – мала константа для уникнення ділення на нуль; $\hat{x}_i$ – нормалізоване значення з нульовим середнім та одиничною дисперсією; γ і β – навчальні параметри масштабу і зміщення, що дозволяють мережі скоригувати нормалізований розподіл; y_i – вихідне значення.

Результатом BN є те, що незалежно від того, які дані надходять на вхід шару, їх розподіл на виході завжди залишається стабільним. Це дозволяє використовувати більші швидкості навчання і суттєво прискорює збіжність. BN зазвичай застосовується після кожного згорткового шару, до функції активації ReLU [29].

Dropout є методом регуляризації, що запобігає перенавчанню – ситуації, коли мережа «заучує» конкретні навчальні приклади замість того щоб навчитись узагальнювати закономірності. Механізм Dropout полягає у тому, що під час кожного кроку навчання кожен нейрон з ймовірністю p_{drop} незалежно «вимикається» – його вихід прирівнюється до нуля. Виходи активних нейронів масштабуються, щоб сумарна потужність сигналу залишалась незмінною [29]:

$$y_i = \frac{x_i m_i}{(1 - p_{drop})}, \quad m_i \sim \text{Bernoulli}(1 - p_{drop}), \quad (2.45)$$

де y_i – вихідне значення нейрона після застосування Dropout; x_i – вхідне значення i -го нейрона; i – індекс нейрона у шарі; m_i – бінарна випадкова маска, що приймає значення 0 (нейрон вимкнено) або 1 (нейрон працює), що генерується незалежно для кожного нейрона на кожній ітерації навчання; p_{drop} – ймовірність вимикання нейрона; $(1 - p_{drop})$ – ймовірність того що нейрон залишається активним; ділення на $(1 - p_{drop})$ компенсує зменшення середнього виходу шару внаслідок вимикання частини нейронів; символ $\sim$ означає «розподілений за»; $\text{Bernoulli}(1 - p_{drop})$ – розподіл Бернуллі з параметром $(1 - p_{drop})$, за яким маска m_i приймає значення 1 з ймовірністю $(1 - p_{drop})$ та значення 0 з ймовірністю p_{drop} .

Наприклад, при $p_{drop} = 0,5$ приблизно половина нейронів вимикається на кожному кроці навчання. Це змушує мережу навчатись розподіленому представленню ознак і не покладатись на будь-яку конкретну підмножину нейронів. Під час тестування Dropout вимикається – всі нейрони активні, а ваги вже підготовлені до роботи в такому режимі завдяки масштабуванню.

Компоненти архітектури CNN, розглянуті у підрозділах 2.2.1 та 2.2.2,

узагальнено у таблиці 2.3 із зазначенням призначення кожного компонента, форми вхідного та вихідного тензорів.

Таблиця 2. 3 Компоненти CNN: призначення та характеристики перетворення тензорів

Шар	Вхід	Вихід	Призначення та ефект
Згортка (F фільтрів $k \times k$)	$H \times W \times H_{in}$	$H_{out} \times W_{out} \times F$	Виявлення F різних локальних патернів у вхідному зображенні. Кількість параметрів: $F \cdot (k^2 \cdot C_{in} + 1)$
Batch Normalization	$H \times W \times C$	$H \times W \times C$	Стабілізація розподілу активацій. Розмір не змінюється. Параметри: $2 \cdot C$ (γ та β для кожного каналу)
ReLU	$H \times W \times C$	$H \times W \times C$	Нелінійність: від'ємні значення (відсутність патерну) замінюються нулем. Параметрів немає

Продовження таблиця 2.3

Max Pooling $p \times p$	$H \times W \times C$	$(H/p) \times (W/p) \times C$	Просторове стиснення: з кожного вікна $p \times p$ зберігається лише максимум. Інваріантність до малих зсувів. Параметрів немає
Global Avg Pooling	$H \times W \times C$	C (вектор)	Стиснення просторових карт до глобального вектора ознак. Замінює flatten+FC, суттєво зменшуючи кількість параметрів. Параметрів немає
Dropout (p_{drop})	вектор N	вектор N	Регуляризація: p_{drop} частка нейронів випадково вимикається під час навчання. Запобігає перенавчанню. Параметрів немає

Продовження таблиця 2.3

FC-шар (N_{out} нейронів)	Вектор N_{in}	Вектор N_{out}	Нелінійна комбінація всіх ознак для класифікації. Параметрів: $N_{in} \cdot N_{out} + N_{out}$
Softmax (N_c класів)	Вектор N_c (логіти)	Вектор N_c (ймов.)	Перетворення оцінок класів у вектор ймовірностей: усі значення від 0 до 1, сума дорівнює 1. Параметрів немає

Розглянуті компоненти об'єднуються у типову архітектуру CNN наступним чином. Вхідне зображення послідовно проходить через N згорткових блоків, де N – загальна кількість блоків, що визначається складністю задачі. Кожен блок складається з фіксованої послідовності шарів: згортка (Conv), пакетна нормалізація (BN), функція активації (ReLU) та пулінг (Pool). Після кожного блоку просторові розміри карти ознак зменшуються внаслідок пулінгу, а кількість каналів зростає внаслідок збільшення кількості фільтрів. Після останнього N -го блоку тензор має розміри $H_N \times W_N \times F_N$, де H_N та W_N – просторові розміри після N послідовних пулінгів, F_N – кількість фільтрів останнього шару. Глобальний середній пулінг (GAP) перетворює цей тензор у вектор розмірності $d = F_N$, тобто по одному числу на кожен канал. Цей вектор подається на повнозв'язний шар (FC) з функцією активації ReLU, а потім на вихідний шар з функцією Softmax, що формує вектор ймовірностей класів p розмірності N_c . Загальну архітектуру CNN від вхідного

зображення до вектора ймовірностей класів показано на рисунку 2.3.

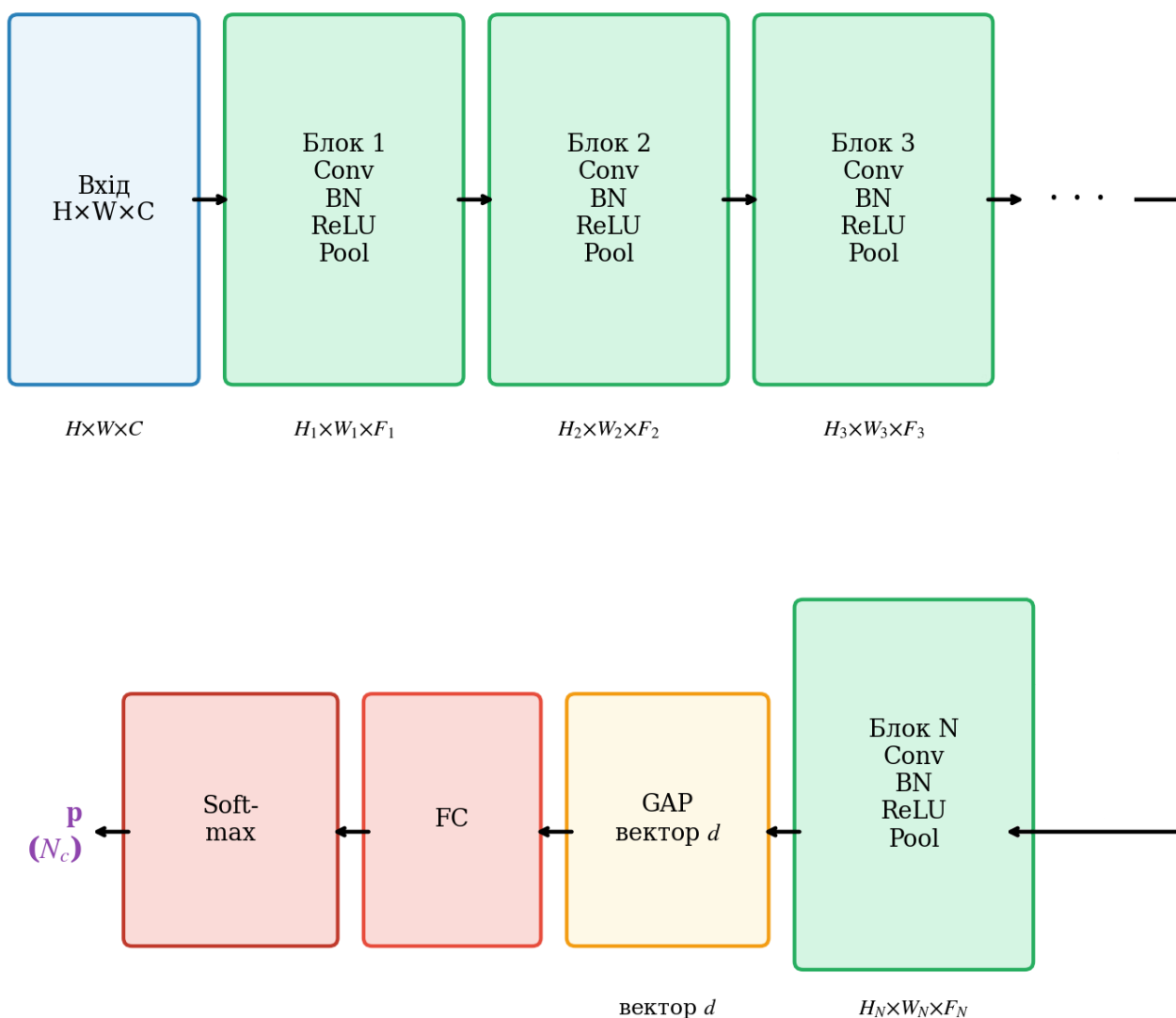


Рисунок 2. 3 Загальна архітектура CNN для задачі класифікації зображень

Таким чином, розглянуті компоненти – пулінг для зменшення просторових розмірів, пакетна нормалізація для стабілізації навчання та Dropout для запобігання перенавчанню – разом із операцією згортки та функціями активації (підрозділ 2.2.1) утворюють повний набір будівельних блоків для побудови архітектури CNN довільної глибини та складності.

2.2.3. Багатомасштабна CNN з паралельними гілками (MSCNN)

Стандартна CNN з однаковим розміром ядра у всіх шарах аналізує зображення лише в одному просторовому масштабі – тобто «бачить» патерни лише певного розміру. Однак у реальних задачах класифікації зображення може

одночасно містити значущі ознаки різних масштабів: дрібні локальні деталі і великомасштабні структури, що охоплюють значну площу. Наприклад, малим ядром 3×3 зручно виявляти вузькі точкові патерни, а великим ядром 5×5 – широкі протяжні структури. Якщо для аналізу доступний лише один розмір ядра, частина інформації неминуче втрачається [31].

У задачах вібраційної діагностики ця проблема є особливо актуальною: дефекти підшипників проявляються на спектрограмі як вузькі локальні імпульси, що потребують дрібного ядра для виявлення, тоді як дисбаланс ротора створює широкі горизонтальні смуги, які краще виявляються ядром більшого розміру.

Концепція багатомасштабної CNN (Multi-Scale CNN, MSCNN) вирішує цю проблему шляхом паралельного оброблення вхідного зображення кількома незалежними гілками, кожна з яких використовує ядра різного розміру. Обидві гілки отримують однакове вхідне зображення, але аналізують його через «вікна» різного розміру. Результати обох гілок у вигляді карт ознак мають однакові просторові розміри (завдяки доповненню нулями) і об'єднуються шляхом конкатенації – послідовного «склеювання» по осі каналів. Після об'єднання класифікатор отримує розширений ознаковий опис, що одночасно містить ознаки обох масштабів [23]. Архітектуру MSCNN з двома паралельними гілками і схему об'єднання ознак ілюструє рисунок 2.4.

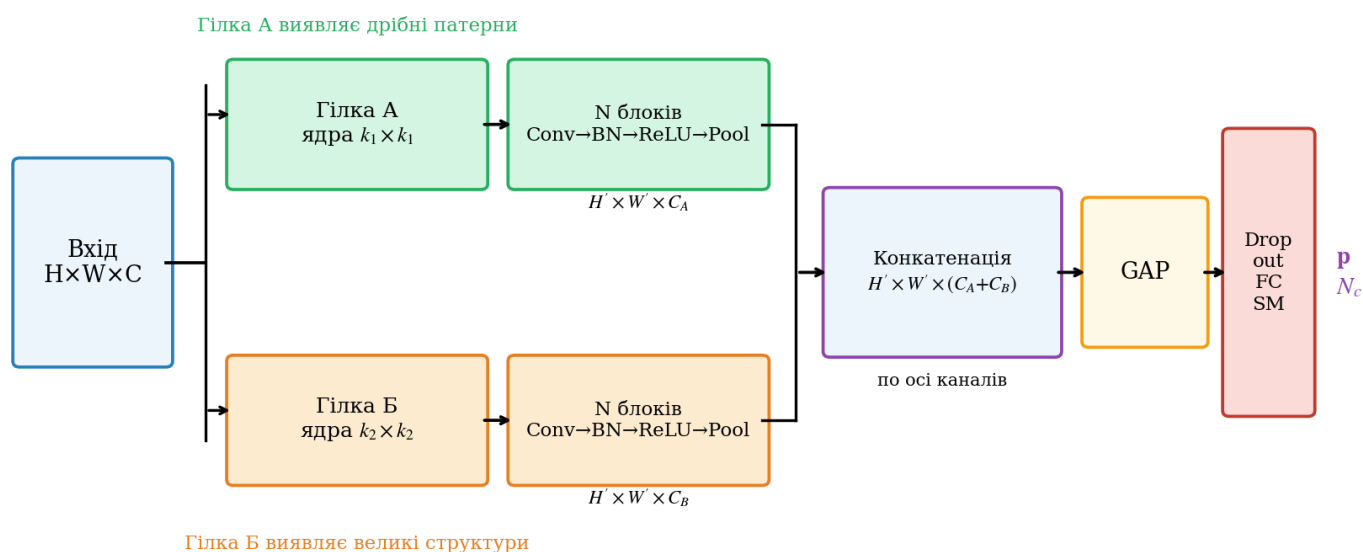


Рисунок 2. 4 Архітектура MSCNN з двома паралельними гілками і схема об'єднання ознак

Математично конкатенація двох тензорів X_A розміром $H \times W \times C_A$ та X_B розміром $H \times W \times C_B$ створює результуючий тензор X_{cat} розміром $H \times W \times (C_A + C_B)$, у якому перші C_A каналів відповідають виходу гілки А, а наступні C_B каналів – виходу гілки Б:

$$X_{cat}[i, j, c] = \begin{cases} X_A[i, j, c], c = 1, \dots, C_A \\ X_B[i, j, c - C_A], c = C_A + 1, \dots, C_A + C_B \end{cases} \quad (2.46)$$

де i та j – просторові індекси; c – індекс каналу у результуючому тензорі; $(c - C_A)$ – зсунутий індекс каналу для гілки Б.

Переваги MSCNN порівняно зі стандартною CNN ілюструє наступний приклад. Нехай маємо задачу класифікації з двома типами несправностей: тип 1 проявляється як вузька вертикальна смуга шириною 2 пікселі на спектрограмі, а тип 2 – як широка горизонтальна зона підвищеної яскравості заввишки 10 пікселів. Ядро 3×3 добре виявляє перший патерн, але погано «бачить» другий у своєму обмеженому полі зору. Ядро 5×5 добре виявляє другий патерн, але може пропустити перший через «розмиття» локальних деталей. MSCNN з паралельними ядрами 3×3 і 5×5 виявляє обидва патерни одночасно і після об'єднання класифікатор має повну картину [32]. Принцип виявлення ознак різного масштабу ядрами 3×3 та 5×5 на прикладі фрагменту спектрограми наведено на рисунку 2.5.

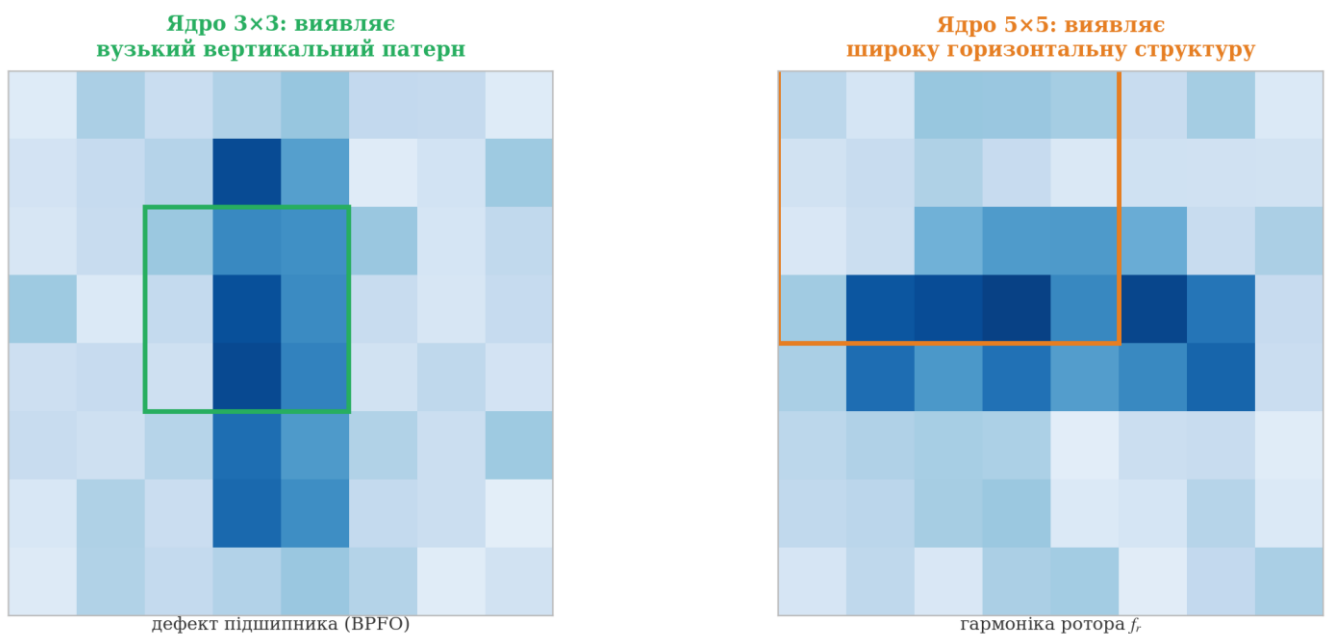


Рисунок 2. 5 Виявлення ознак різного просторового масштабу ядрами 3×3 та 5×5

Порівняння стандартної CNN та MSCNN за ключовими характеристиками наведено у таблиці 2.4 [31, 34].

Таблиця 2. 4 Порівняння стандартної CNN та MSCNN

Характеристика	Стандартна CNN	MSCNN
Розмір ядер згортки	Однаковий у всіх шарах (наприклад, 3×3 або 5×5)	Різний у паралельних гілках (наприклад, 3×3 в одній і 5×5 в іншій)
Масштаб аналізу	Один масштаб – лише локальні або лише крупні ознаки	Одночасно кілька масштабів – і локальні деталі, і великомасштабні структури
Вхід	Одне зображення $\rightarrow$ одна гілка	Одне зображення $\rightarrow$ паралельно кілька гілок
Вихід ознакового блоку	Тензор $H' \times W' \times C$	Тензор $H' \times W' \times (C_A + C_B)$ після конкатенації – удвічі більший ознаковий опис
Кількість параметрів	Менша – одна гілка	Більша – кожна гілка навчається незалежно
Ефективність на різномасштабних задачах	Нижча – ядро фіксованого розміру пропускає ознаки інших масштабів	Вища – паралельні ядра різних розмірів охоплюють повний спектр масштабів

Таким чином, у підрозділі 2.2 розглянуто математичний апарат і принцип роботи CNN як методу автоматичного виділення просторових ознак із двовимірних зображень. Показано що кожен компонент архітектури виконує конкретну функцію: операція згортки виявляє локальні патерни, пулінг стискає просторові

розміри і забезпечує інваріантність до зсувів, пакетна нормалізація стабілізує навчання, Dropout запобігає перенавчанню, а Softmax перетворює вихід мережі у вектор ймовірностей класів. Архітектура MSCNN з паралельними гілками розширює стандартну CNN, дозволяючи одночасно аналізувати ознаки різного просторового масштабу, що є важливою перевагою при класифікації спектрограм вібраційних сигналів. Проте MSCNN обробляє кожну спектрограму окремо і не враховує часову послідовність зміни стану обладнання. Для аналізу часових залежностей у послідовностях вимірювань у підрозділі 2.3 розглянуто рекурентні нейронні мережі.

2.3. Рекурентні нейронні мережі та механізм уваги

Згорткові нейронні мережі, розглянуті у підрозділі 2.2, є потужним інструментом для виявлення просторових патернів у зображеннях. Однак вони обробляють кожне вхідне зображення незалежно – не зберігаючи жодної інформації про попередні спостереження. У задачах діагностики обертового устаткування це є суттєвим обмеженням: стан двигуна змінюється поступово, і для точного визначення типу та стадії розвитку несправності важливо враховувати не лише поточну спектрограму, а й динаміку змін у часі – як еволюціонував сигнал протягом останніх секунд, хвилин або годин. Саме для аналізу таких часових залежностей призначені рекурентні нейронні мережі [36].

2.3.1. Рекурентні нейронні мережі та проблема зникнення градієнта

Рекурентна нейронна мережа (Recurrent Neural Network, RNN) – це архітектура, що обробляє послідовності даних, зберігаючи «пам'ять» про попередні елементи послідовності у вигляді прихованого стану (hidden state). На відміну від звичайної нейронної мережі, де інформація передається лише у прямому напрямку від входу до виходу, RNN має зворотні зв'язки: вихід мережі на кожному кроці подається назад як частина входу на наступному кроці [36].

Формально, на кожному часовому кроці t рекурентна мережа отримує вхідний вектор x_t і обчислює прихований стан h_t на основі поточного входу та попереднього прихованого стану:

$$h_t = \tanh(W_h h_{t-1} + W_x x_t + b), \quad (2.47)$$

де h_t – прихований стан на кроці t , що є «стислим» описом усієї передісторії послідовності аж до поточного кроку; h_{t-1} – прихований стан на попередньому кроці; x_t – вхідний вектор на кроці t ; W_h – матриця ваг зворотного зв'язку (від попереднього прихованого стану); W_x – матриця ваг від входу; b – вектор зміщень; $\tanh$ – гіперболічний тангенс як функція активації, що обмежує значення h_t діапазоном $(-1, 1)$. Вихідне значення мережі на кроці t обчислюється з прихованого стану:

$$y_t = W_y h_t + b_y, \quad (2.48)$$

де y_t – вихідний вектор на кроці t ; W_y – матриця ваг вихідного шару; b_y – вектор зміщень вихідного шару.

Графічно RNN можна зобразити як нейронну мережу, «розгорнуту» у часі: на кожному кроці t є однакова мережа з тими самими вагами W_h , W_x , що обробляє черговий елемент послідовності та передає прихований стан наступному кроку. Саме спільність ваг дозволяє RNN застосовуватись до послідовностей довільної довжини [36]. Архітектуру RNN у компактному та розгорнутому у часі вигляді наведено на рисунку 2.6.

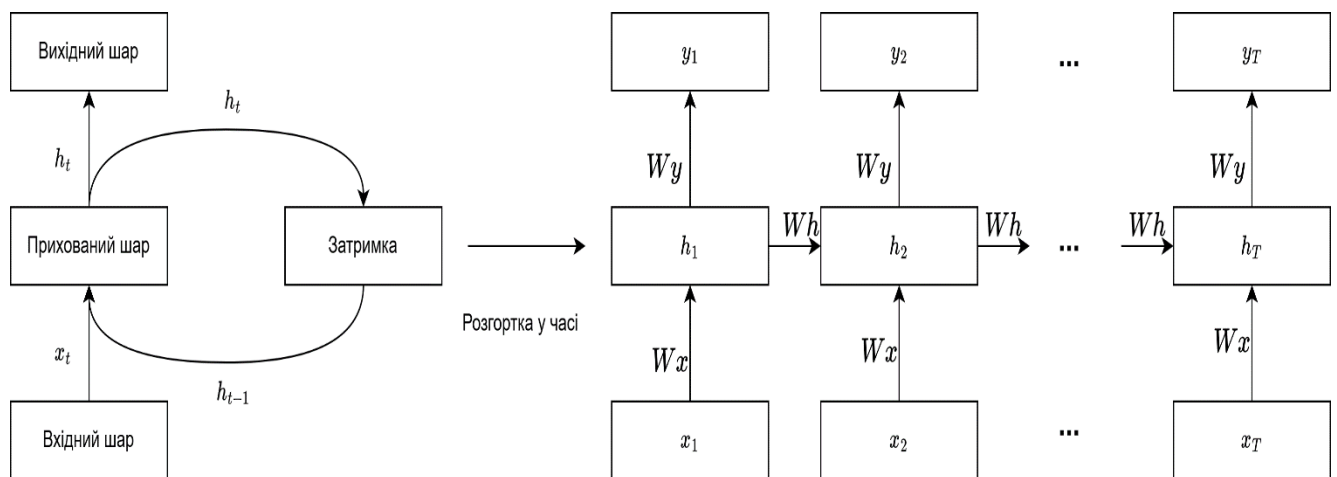


Рисунок 2. 6 Архітектура рекурентної нейронної мережі у компактному та розгорнутому у часі вигляді

Незважаючи на теоретичну здатність враховувати довгострокові залежності, стандартні RNN на практиці страждають від критичної проблеми – зникнення

градієнта (vanishing gradient problem). Щоб зрозуміти суть цієї проблеми, необхідно звернутись до принципу навчання нейронних мереж. Мережа навчається шляхом зворотного поширення помилки (backpropagation): похибка на виході «поширюється» назад через усі шари і часові кроки, коригуючи ваги. При цьому градієнт помилки на кожному кроці множиться на матрицю ваг W_h . Якщо найбільше власне значення цієї матриці, позначене λ , менше 1, то при розповсюдженні через T кроків градієнт зменшується як $\lambda^T \rightarrow 0$. Наприклад, при $\lambda = 0,9$ і $T = 50$ кроках градієнт зменшується до $0,9^{50} \approx 0,005$, тобто у 200 разів – мережа практично перестає враховувати події, що відбулись більше ніж 10–20 кроків тому [37].

Для задачі діагностики це означає наступне: якщо ознаки несправності розвивались протягом багатьох часових вікон, стандартна RNN «забуває» ранні ознаки до моменту класифікації. Це є принциповим обмеженням, яке не може бути усунуто простим збільшенням кількості шарів або нейронів. Вирішення цієї проблеми було запропоновано Хохрайтером і Шмідхубером у 1997 році у вигляді архітектури LSTM [37]. Порівняння стандартної RNN та LSTM за ключовими характеристиками наведено у таблиці 2.5.

Таблиця 2. 5 Порівняння стандартної RNN та LSTM

Характеристика	Стандартна RNN	LSTM
Пам'ять про минулі події	Короткострокова (10–20 кроків)	Довгострокова (сотні кроків)
Проблема зникнення градієнта	Присутня, навчання важке	Значно послаблена завдяки вентилям
Внутрішній стан	Один вектор h_t	Два вектори: h_t (прихований) та c_t (клітинний)
Механізм управління пам'яттю	Відсутній	Три вентиля: забуття, вхідний, вихідний

Продовження таблиці 2.5

Кількість параметрів	Менша	У ~ 4 рази більша при однаковому розмірі h
----------------------	-------	---

Як видно з таблиці 2.5, архітектура LSTM усуває основне обмеження стандартної RNN завдяки спеціальному механізму керування пам'яттю – вентилям, що дозволяють мережі самостійно визначати яку інформацію зберігати, а яку відкидати. Детальний опис архітектури LSTM та вентиляного механізму наведено далі.

2.3.2. Архітектура LSTM та вентиляний механізм

LSTM (Long Short-Term Memory – довга короткострокова пам'ять) є спеціалізованою архітектурою рекурентної нейронної мережі, що вирішує проблему зникнення градієнта шляхом введення спеціального механізму керування пам'яттю – вентилів (gates). Ідея вентилів полягає у тому, що мережа сама навчається вирішувати: яку інформацію зберігати в пам'яті, яку забувати, яку приймати з нового входу і яку видавати на вихід. Це принципово відрізняє LSTM від стандартної RNN, де «пам'ять» є лише пасивним накопиченням усіх попередніх входів без жодного механізму відбору [37].

В архітектурі LSTM кожна клітина (cell) містить два вектори стану: клітинний та прихований.

Клітинний стан c_t (cell state) – довгострокова пам'ять, що проходить через усі часові кроки майже без змін, якщо вентиля це дозволяють. Саме клітинний стан забезпечує «канал» для передачі градієнта без затухання.

Прихований стан h_t (hidden state) – короткострокова пам'ять, що є виходом клітини на поточному кроці і передається на вхід наступного кроку, а також використовується як вихід всього шару LSTM.

Для керування потоком інформації LSTM використовує три вентиля – вхідний (input gate), вентиль забуття (forget gate) і вихідний (output gate). Кожен вентиль є нейронним шаром зі своїми навчальними ваговими матрицями і

функцією активації сигмоїд, що видає значення від 0 до 1. Функція сигмоїд визначається як:

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad (2.49)$$

де x – вхідне значення; e – основа натурального логарифма. Функція сигмоїд перетворює будь-яке вхідне значення $x \in (-\infty, +\infty)$ у діапазон $(0, 1)$ і застосовується до кожного значення незалежно. У контексті вентилів LSTM вихідне значення $\sigma(x)$ інтерпретується як ступінь «відкритості» вентиля: значення близьке до 0 означає «повністю заблоковано» (інформація не проходить), значення близьке до 1 – «повністю пропущено» (інформація проходить без змін) [38].

Вентиль забуття (forget gate) визначає, яку частину попереднього клітинного стану c_{t-1} слід зберегти, а яку «забути». Він аналізує попередній прихований стан h_{t-1} та поточний вхід x_t :

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f), \quad (2.50)$$

де σ – функція сигмоїд; $[h_{t-1}, x_t]$ – конкатенація (послідовне об'єднання) векторів h_{t-1} і x_t ; W_f – матриця навчальних ваг вентиля забуття; b_f – вектор зміщень. Результат f_t є вектором значень від 0 до 1: компонент, близький до 0, означає «забути цю частину пам'яті», близький до 1 – «зберегти». Наприклад, якщо мережа аналізуючи послідовність спектрограм виявила що сигнал повернувся до норми після короткочасного стрибка, вентиль забуття може «видалити» інформацію про той стрибок з клітинного стану.

Вхідний вентиль (input gate) визначає, яку нову інформацію слід записати в клітинний стан. Він складається з двох частин: вхідного вентиля i_t та вектора-кандидата g_t :

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i), \quad (2.51)$$

$$g_t = \tanh(W_g * [h_{t-1}, x_t] + b_g), \quad (2.52)$$

де i_t – вектор значень від 0 до 1, що визначає «скільки» нової інформації записати; W_i та b_i – навчальні параметри вхідного вентиля; g_t – вектор-кандидат зі значеннями від -1 до 1 , що визначає «яку саме» нову інформацію записати; W_g та b_g – навчальні параметри вектора-кандидата. Таким чином, вхідний вентиль i_t

визначає «скільки» записати, а вектор-кандидат g_t визначає «що саме» записати.

Оновлення клітинного стану відбувається як поєднання «забування» старого і «запису» нового:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t, \quad (2.53)$$

де $\odot$ – поелементне множення (добуток Адамара); $f_t \odot c_{t-1}$ – частина попереднього клітинного стану, що зберігається; $i_t \odot g_t$ – нова інформація, що записується. Ця формула є ключовою: вона показує, що клітинний стан оновлюється через додавання, а не множення на матрицю ваг – саме це забезпечує стабільну передачу градієнта через тривалі послідовності без затухання [37].

Вихідний вентиль (output gate) визначає, яку частину клітинного стану слід «прочитати» і видати як прихований стан h_t :

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o), \quad (2.54)$$

де W_o – матриця навчальних ваг вихідного вентиля; b_o – вектор зміщень.

Прихований стан обчислюється як:

$$h_t = o_t \odot \tanh c_t, \quad (2.55)$$

де o_t – вихідний вентиль зі значеннями від 0 до 1; $\tanh c_t$ – нормалізований клітинний стан у діапазоні $(-1, 1)$. Прихований стан h_t є виходом LSTM-клітини на кроці t : він передається як вхід на наступний крок і використовується як вихідний вектор для подальшої обробки або класифікації [38]. Внутрішню архітектуру клітини LSTM з трьома вентилями наведено на рисунку 2.7.

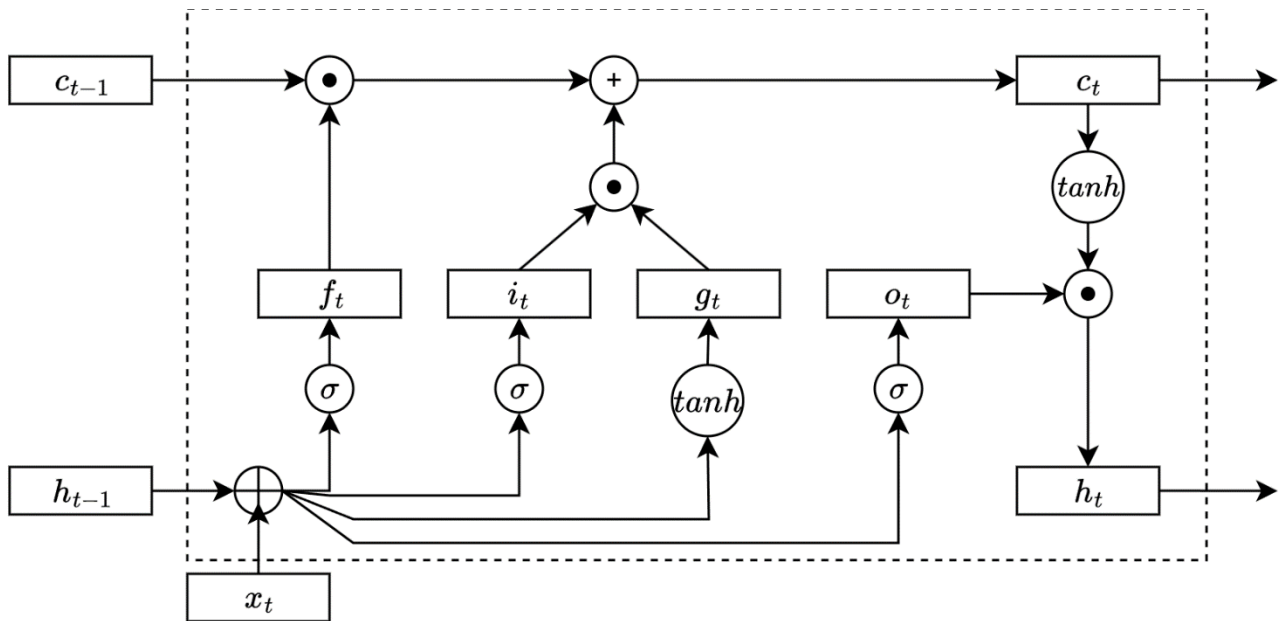


Рисунок 2. 7 Внутрішня архітектура клітини LSTM з трьома вентилями

Загальна кількість навчальних параметрів одного шару LSTM з розміром прихованого стану h та розміром входу d :

$$P_{LSTM} = 4h(d + h + 1), \quad (2.56)$$

де h – розмір прихованого стану (кількість нейронів LSTM); d – розмірність вхідного вектора; коефіцієнт 4 відображає те, що LSTM містить чотири незалежних нейронних шари (три вентиля та вектор-кандидат), кожен зі своїми ваговими матрицями. Наприклад, при $h = 64$ і $d = 128$: $P_{LSTM} = 4 \cdot 64 \cdot (128 + 64 + 1) = 49408$ параметрів.

2.3.3. Обробка часових послідовностей у LSTM

Для розуміння того, як LSTM працює з часовими послідовностями, розглянемо деталі проходження даних через шар LSTM. Вхідними даними є послідовність векторів $X = [x_1, x_2, \dots, x_T]$, де x_t – вектор ознак на кроці t ($t = 1, 2, \dots, T$) розмірності d ; T – довжина послідовності; d – розмірність вхідного вектора [36].

На початку обробки ініціалізуються початкові стани: $h_0 = 0$ (нульовий вектор) і $c_0 = 0$. Потім для кожного кроку t від 1 до T мережа послідовно виконує формули (2.50)–(2.55), отримуючи оновлені h_t і c_t .

Шар LSTM може працювати у двох режимах, що визначається параметром

`return_sequences`. У режимі `return_sequences = False` шар повертає тільки останній прихований стан h_T , який містить стиснений опис усієї передісторії послідовності. Цей режим використовується коли для подальшої обробки достатньо одного підсумкового вектора. У режимі `return_sequences = True` шар повертає послідовність усіх прихованих станів $[h_1, h_2, \dots, h_T]$. Цей режим є необхідним коли подальший шар потребує інформацію з кожного часового кроку окремо – зокрема, для механізму уваги (підрозділ 2.3.4), який аналізує приховані стани всіх кроків для визначення найбільш інформативних моментів часу.

Розглянемо конкретний приклад того, як LSTM обробляє послідовність спектрограм вібраційного сигналу. Нехай маємо T послідовних спектрограм, знятих через рівні інтервали часу. Згорткова мережа MSCNN (підрозділ 2.2.3) обробляє кожен спектрограму окремо і видає вектор ознак розмірності d (після Global Average Pooling). Ці T векторів утворюють послідовність $X = [x_1, \dots, x_T]$. LSTM обробляє цю послідовність крок за кроком: на першому кроці він аналізує ознаки першої спектрограми і формує початковий прихований стан h_1 ; на другому кроці бере ознаки другої спектрограми і попередній стан h_1 , формуючи h_2 з урахуванням вже побаченого; і так далі до h_T . Фінальний стан h_T відображає не лише поточну спектрограму, а й увесь часовий контекст – як змінювались ознаки впродовж послідовності. Саме це робить LSTM принципово кращим за ізольовану класифікацію окремих спектрограм [38]. Схему обробки часової послідовності шаром LSTM наведено на рисунку 2.8.

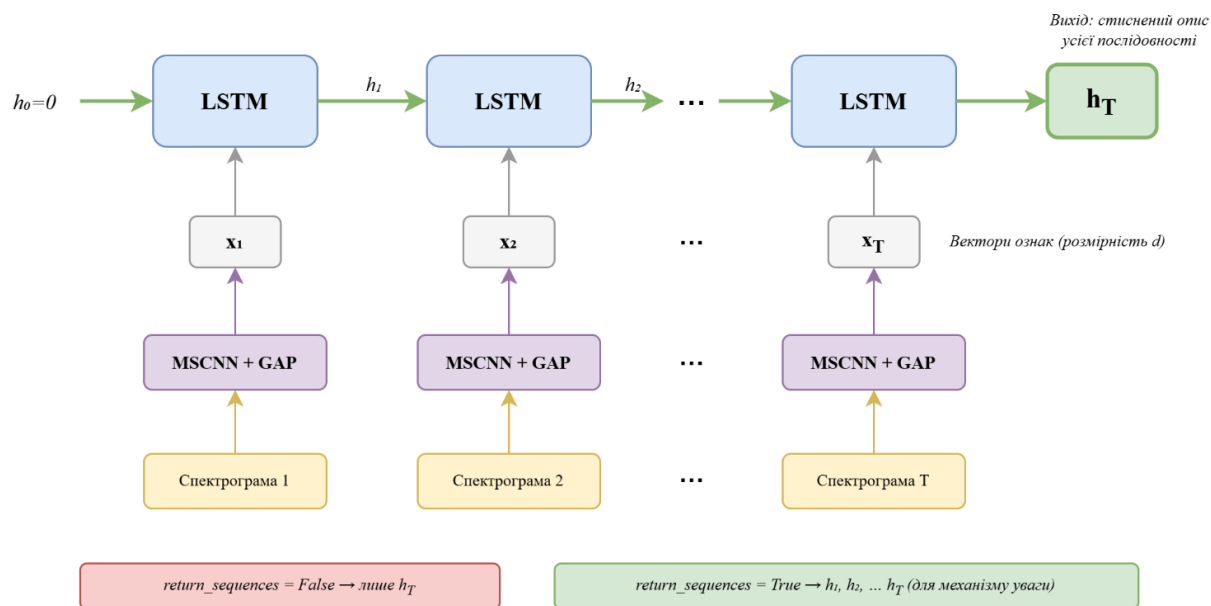


Рисунок 2. 8 Обробка часової послідовності шаром LSTM

Важливою модифікацією є двонаправлений LSTM (Bidirectional LSTM, BiLSTM), де дві незалежні LSTM-мережі обробляють послідовність у прямому та зворотному напрямках, а їх виходи конкатенуються. BiLSTM є переважним у задачах аналізу збережених записів, де доступна вся послідовність цілком. Для задач онлайн-діагностики, де рішення приймається у реальному часі і майбутні спостереження недоступні, використовується стандартний однонаправлений LSTM [39]. У даній роботі застосовується однонаправлений LSTM.

2.3.4. Механізм уваги (Attention Mechanism)

Механізм уваги (Attention Mechanism) є надбудовою над LSTM, що вирішує ще одне обмеження рекурентних мереж – «вузьке місце» у вигляді одного фінального прихованого стану h_T . При обробці довгої послідовності весь її зміст повинен бути «стиснений» в один вектор h_T фіксованого розміру, що є надмірно жорстким обмеженням для складних послідовностей. Механізм уваги дозволяє мережі при формуванні класифікаційного рішення «звертатись» до прихованих станів усіх T кроків послідовності, а не лише до останнього. Для цього LSTM має працювати у режимі $return_sequences = True$ (підрозділ 2.3.3), повертаючи приховані стани $[h_1, h_2, \dots, h_T]$ для всіх часових кроків [40].

Ідея механізму уваги полягає у наступному: різні часові кроки послідовності

можуть мати різний діагностичний «внесок» у поточну класифікацію. Наприклад, якщо дефект підшипника проявляється у вигляді короткого сплеску ознак на кроці $t = 5$, а решта кроків є «нормальними», то для правильної класифікації мережа повинна «приділити увагу» саме кроку 5, а не розподіляти вагу рівномірно між усіма T кроками. Механізм уваги навчається обчислювати такі «ваги уваги» автоматично [40].

Математично механізм уваги обчислює скалярний оцінковий бал (attention score) e_t для кожного прихованого стану h_t . Базовий варіант (адитивна увага) використовує навчальний вектор v_a для обчислення оцінки:

$$e_t = v_a^T \tanh(W_a \cdot h_t + b_a), \quad (2.57)$$

де e_t – оцінковий бал для кроку t , що відображає «важливість» цього кроку; h_t – прихований стан LSTM на кроці t ; W_a – навчальна матриця ваг розміром $h \times h$; b_a – вектор зміщень; $\tanh$ – функція активації; v_a^T – транспонований навчальний вектор (верхній індекс T тут означає операцію транспонування, а не кількість часових кроків), що перетворює результат функції $\tanh$ у скалярне значення e_t .

Оцінки e_t нормалізуються функцією Softmax (формула 2.40) для отримання вагового розподілу уваги:

$$\alpha_t = \frac{e^{e_t}}{\sum e^{e_\tau}}, \tau = 1 \dots T, \quad (2.58)$$

де α_t – вага уваги для кроку t , що показує яку частку уваги мережа приділяє цьому кроку; e_t – оцінковий бал для кроку t , обчислений за формулою (2.57); e_τ – оцінковий бал для кроку τ ; τ – індекс підсумовування, що пробігає значення від 1 до T ; T – загальна кількість часових кроків. Ділення експоненти для одного кроку на суму експонент усіх кроків (функція Softmax) гарантує що всі ваги α_t додатні та їх сума дорівнює одиниці: $\alpha_t \in (0, 1)$ та $\sum \alpha_t = 1$. Таким чином, вектор ваг α є розподілом уваги по часових кроках, де більше значення α_t означає що крок t є більш важливим для поточного класифікаційного рішення.

Виходом механізму уваги є зважена сума прихованих станів:

$$z = \sum \alpha_t h_t, t = 1 \dots T, \quad (2.59)$$

де z – контекстний вектор розмірності h , що є «зваженим резюме» всієї

послідовності з акцентом на найбільш інформативних кроках. Саме цей вектор z подається на повнозв'язний класифікаційний шар замість одного фінального прихованого стану h_T [40]. Схему механізму уваги поверх шару LSTM наведено на рисунку 2.9.

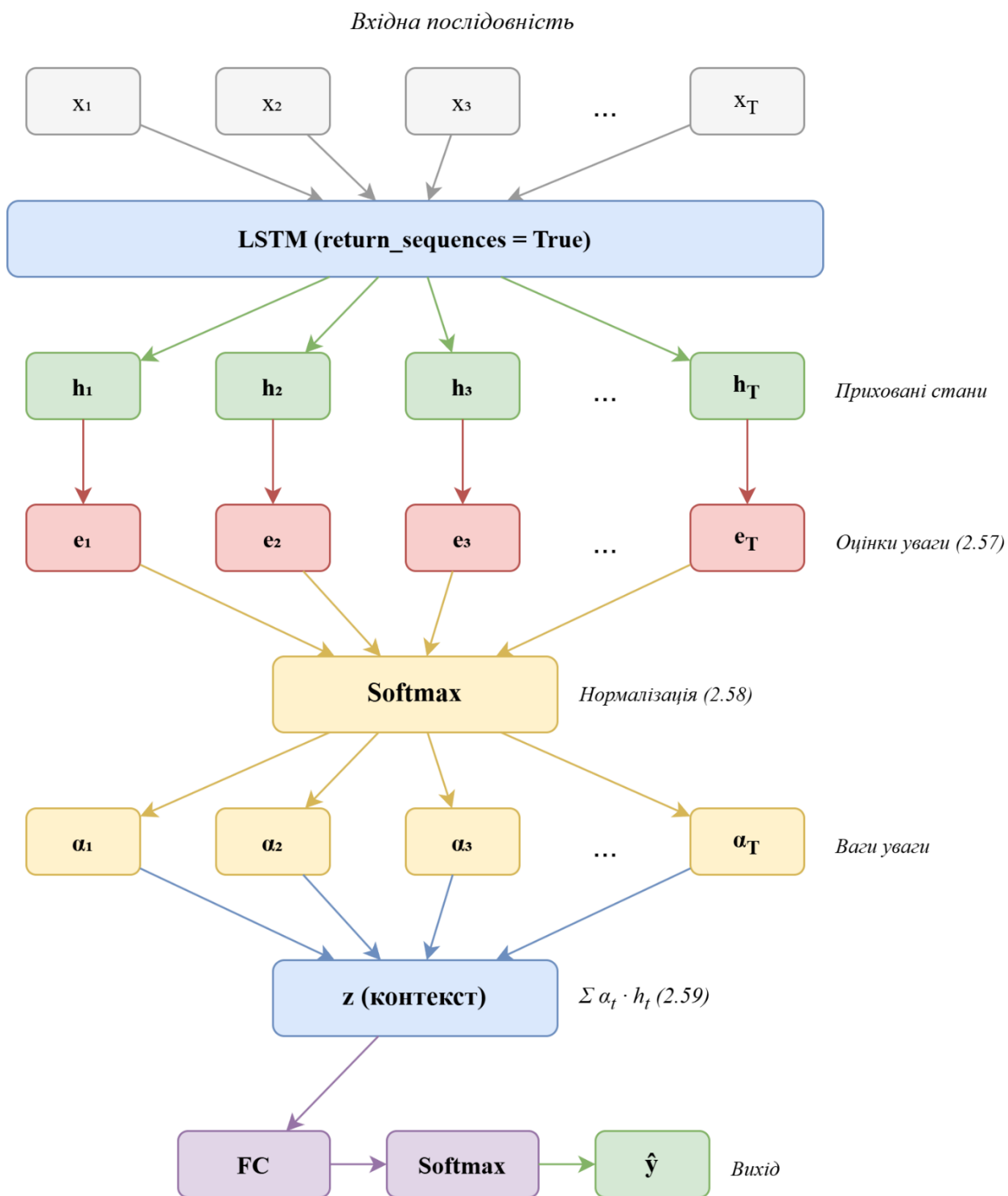


Рисунок 2. 9 Механізм уваги поверх шару LSTM: від прихованих станів до контекстного вектора

Важливим є питання про те, що саме відображають ваги уваги α_t в контексті діагностики вібраційних сигналів. Вага α_t показує, наскільки «важливим» є t -й

часовий крок для поточного класифікаційного рішення. Якщо розвивається дефект підшипника і характерні ознаки з'явилися кілька кроків тому, механізм уваги присвоїть вищі ваги тим крокам, де ці ознаки проявились найчіткіше. Це підвищує точність класифікації і водночас забезпечує часткову інтерпретованість – дослідник може проаналізувати ваги уваги і визначити, які моменти часу найбільше вплинули на прийняте рішення [41].

Описаний механізм уваги працює у часовому вимірі – він визначає значущість кожного часового кроку послідовності, тому називається часовою увагою (Temporal Attention). Він принципово відрізняється від механізму уваги каналів (Channel Attention), що розглядається у підрозділі 2.4.3 і працює у вимірі каналів просторових ознак, визначаючи значущість кожної гілки багатомасштабної мережі MSCNN.

Порівняння підходу на основі механізму уваги з використанням лише фінального прихованого стану LSTM наведено у таблиці 2.6.

Таблиця 2. 6 Порівняння підходів до агрегації виходів LSTM

Характеристика	Використання h_T (останній стан)	Механізм уваги (зважена сума всіх h_t)
Інформація з послідовності	Лише з останнього кроку	З усіх T кроків з різними вагами
Здатність фокусуватись на ключових моментах	Відсутня – всі кроки впливають однаково через рекурентний зв'язок	Присутня – ваги α_t автоматично виділяють найбільш інформативні кроки
Ефективність на довгих послідовностях	Знижується – інформація з ранніх кроків може бути «розмита»	Зберігається – прямий доступ до h_t будь-якого кроку

Продовження таблиці 2.6

Інтерпретованість	Низька	Вища – карта уваги α показує які моменти часу вплинули на рішення
Додаткові параметри	0	W_a, b_a, v_a – невелика кількість ($\sim h^2 + 2h$)

Таким чином, у підрозділі 2.3 розглянуто математичний апарат рекурентних нейронних мереж і механізму часової уваги. Стандартні RNN страждають від проблеми зникнення градієнта, яку вирішує архітектура LSTM через трьохвентильний механізм керування пам'яттю. LSTM обробляє послідовності векторів ознак, накопичуючи часовий контекст у прихованому стані, і може повертати стани всіх часових кроків для подальшого аналізу. Механізм часової уваги розширює LSTM, дозволяючи мережі вибірково фокусуватись на найбільш інформативних часових кроках – що є особливо важливим при аналізі поступово розвиваючихся дефектів обертового устаткування. У підрозділі 2.4 розглядається поєднання згорткової мережі MSCNN з LSTM і механізмом уваги у єдину гібридну архітектуру.

2.4. Гібридна архітектура MSCNN-LSTM

Розглянуті у підрозділах 2.2 та 2.3 методи – MSCNN і LSTM з механізмом уваги – є потужними інструментами, що вирішують принципово різні задачі: MSCNN виявляє просторові патерни у двовимірних зображеннях, тоді як LSTM моделює часові залежності у послідовностях. Поєднання обох підходів в єдину гібридну архітектуру дозволяє отримати систему, яка одночасно розпізнає характерні частотні патерни у спектрограмах та аналізує їх еволюцію у часі [43].

2.4.1. Обґрунтування гібридної архітектури

Задача предиктивної діагностики обертового устаткування включає два взаємопов'язані аспекти: ідентифікацію типу поточного стану (класифікація) та відстеження динаміки зміни цього стану з часом (прогнозування деградації). Огляд

сучасних підходів до діагностики обертового устаткування засобами штучного інтелекту [48] показує, що жоден із методів – CNN або LSTM окремо – не є достатнім для вирішення обох задач одночасно.

Стандартна CNN (або MSCNN) ефективно класифікує поточний стан, оскільки кожна спектрограма є «моментальним знімком» системи. Проте вона не бачить картину в динаміці: для неї не має значення, чи була попередня спектрограма ідентичною чи принципово відмінною. Ця обмеженість є критичною для задач предиктивного технічного обслуговування [48].

Розглянемо приклад: дефект підшипника на ранній стадії може формувати на спектрограмі патерн з малою амплітудою, практично нерозрізнений від шуму. Якщо аналізувати кожен спектрограму ізольовано, така рання стадія буде регулярно класифікуватись як «норма». Однак якщо відстежити 20–30 послідовних спектрограм, можна помітити слабке, але стійке наростання характерного патерну – і саме LSTM здатний виявити цю тенденцію. Таким чином, LSTM виступає не як заміник MSCNN, а як доповнення, що додає часовий вимір до просторового аналізу [44].

Суто LSTM-підхід також має обмеження: якщо подавати на LSTM сирі вібраційні сигнали або послідовності статистичних ознак, мережа змушена самотійно навчитись виявляти складні частотно-просторові залежності, що вимагає значно більших обсягів навчальних даних. MSCNN попередньо виконує ефективне виділення просторових ознак зі спектрограм і передає LSTM вже компактний і інформативний опис кожного моменту часу, а не сирі дані [43].

Порівняння переваг та обмежень трьох підходів – MSCNN окремо, LSTM окремо та гібридної архітектури – наведено у таблиці 2.7.

Таблиця 2. 7 Порівняння підходів до аналізу вібраційних сигналів для задач діагностики [43, 44, 48]

Характеристика	MSCNN окремо	LSTM окремо	MSCNN-LSTM (гібрид)
Виявлення просторових патернів у спектрограмах	Висока ефективність	Слабка	Висока (через блок MSCNN)
Аналіз часової динаміки сигналу	Відсутній	Висока ефективність	Висока (через блок LSTM)
Виявлення ранніх стадій дефектів	Обмежена – ізольована класифікація	Середня – без просторових ознак	Висока – враховує накопичення ознак у часі
Вимоги до навчальних даних	Помірні	Значні	Помірні – MSCNN стискає вхід для LSTM
Складність навчання	Середня	Висока	Вища, але виправдана точністю

Як видно з таблиці 2.7, гібридна архітектура MSCNN-LSTM поєднує переваги обох підходів та усуває їхні основні обмеження. Принцип технічного з'єднання MSCNN та LSTM розглянуто у наступному пункті.

2.4.2. Принцип з'єднання MSCNN і LSTM

MSCNN обробляє одну спектрограму і видає один вектор ознак розмірності d . Для аналізу зміни стану обладнання у часі необхідно обробити не одну, а кілька послідовних спектрограм. Тому з вібраційного сигналу формується вікно з T спектрограм, знятих через рівні проміжки часу Δt . Кожна з T спектрограм окремо проходить через MSCNN, і для кожної мережа видає свій вектор ознак: для першої

спектрограми – x_1 , для другої – x_2 , і так далі до x_T . Ці T векторів утворюють часову послідовність $X = [x_1, x_2, \dots, x_T]$ (підрозділ 2.3.3), яка подається на вхід шару LSTM для аналізу часової динаміки [43, 44].

При цьому для обробки всіх T спектрограм використовується одна і та сама MSCNN з одним набором ваг – принцип розділення ваг у часі (weight sharing). Це означає, що замість навчання T окремих мереж з окремими параметрами єдина MSCNN послідовно застосовується до кожної спектрограми. Розділення ваг забезпечує два ефекти: по-перше, загальна кількість параметрів моделі не залежить від довжини послідовності T ; по-друге, кожен параметр згорткової частини фактично навчається на $T * N_{train}$ зразках, що ефективно збільшує обсяг навчальних даних для MSCNN у T разів та знижує ризик перенавчання на обмежених вибірках. Принцип з'єднання MSCNN і LSTM через послідовність ознакових векторів наведено на рисунку 2.10.

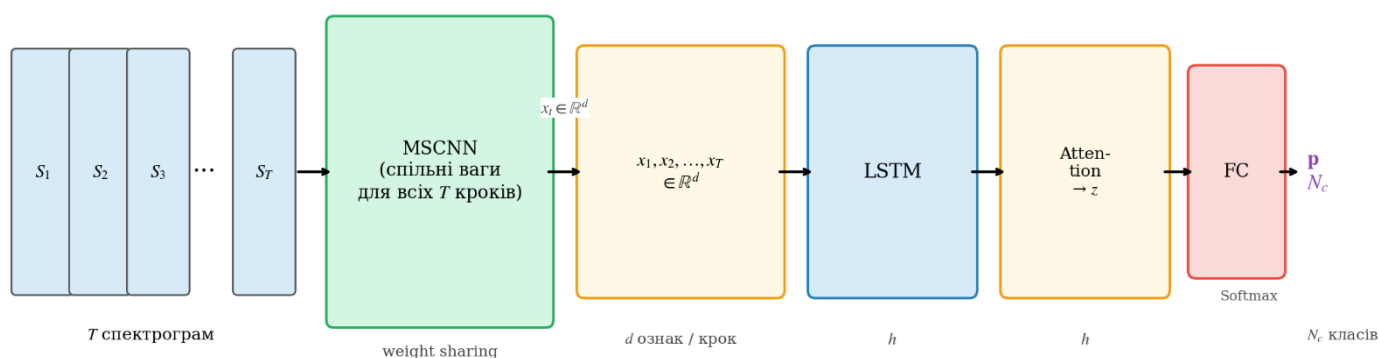


Рисунок 2. 10 Принцип з'єднання MSCNN і LSTM через послідовність ознакових векторів

Послідовність $X = [x_1, x_2, \dots, x_T]$ подається на вхід шару LSTM, який обробляє її крок за кроком у режимі `return_sequences = True`, формуючи приховані стани $[h_1, h_2, \dots, h_T]$. Механізм часової уваги обчислює контекстний вектор z як зважену суму всіх прихованих станів. Цей вектор подається на повнозв'язний класифікаційний блок і Softmax для формування вектора ймовірностей класів p розмірності N_c [43].

Вся система – MSCNN і LSTM – навчається спільно (end-to-end training). Градієнт помилки поширюється від виходу Softmax через весь ланцюжок: FC → Attention → LSTM → MSCNN. Завдяки цьому MSCNN навчається виділяти не просто будь-які ознаки, а саме ті, що є найбільш інформативними для LSTM при аналізі часових тенденцій. Порівняно з послідовним навчанням (спочатку окремо MSCNN, потім окремо LSTM), end-to-end підхід досягає вищої точності класифікації, оскільки обидва блоки оптимізуються під єдину цільову функцію [45]. Повну архітектуру гібридної моделі MSCNN-LSTM з позначенням розмірностей наведено на рисунку 2.11.

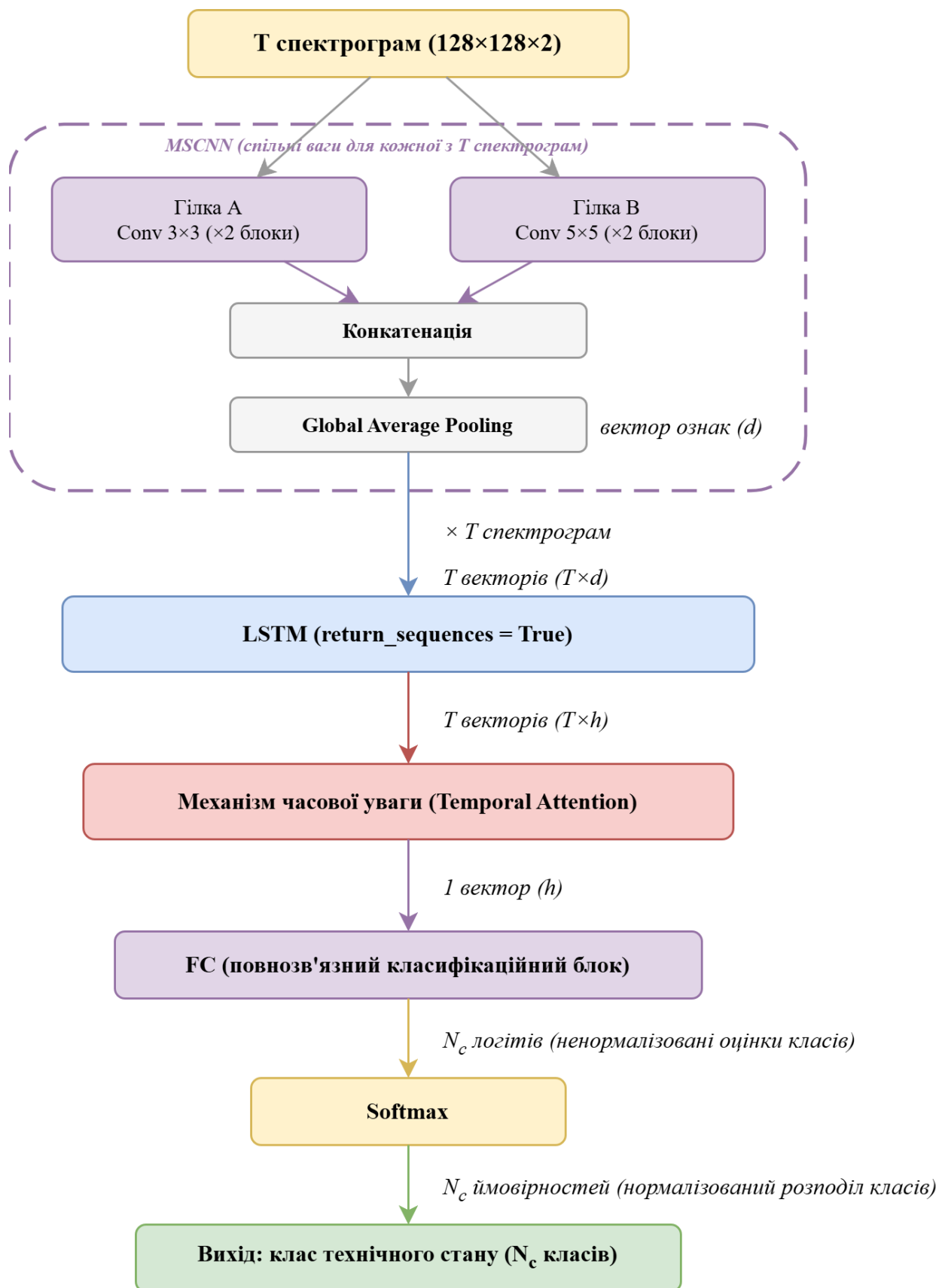


Рисунок 2. 11 Повна архітектура гібридної моделі MSCNN-LSTM

Описана гібридна архітектура забезпечує наскрізну обробку від вхідних спектрограм до вектора ймовірностей класів, де кожен компонент виконує свою функцію: MSCNN виділяє просторові ознаки, LSTM аналізує їх часову динаміку, а механізм уваги фокусується на найбільш діагностично значущих моментах часу. У підрозділі 2.4.3 розглянуто запропоноване удосконалення цієї архітектури – блок міжгілкової уваги каналів.

2.4.3. Запропонований блок міжгілкової уваги каналів

У базовій архітектурі MSCNN результати двох паралельних гілок з різними розмірами ядер об'єднуються операцією конкатенації (формула 2.46), яка надає рівну вагу всім каналам результуючого тензора. Це означає, що незалежно від типу вхідного зразка – чи це дефект підшипника з локальними імпульсними патернами, чи дисбаланс ротора з широкими спектральними смугами – внесок обох гілок у подальшу класифікацію є однаковим. Проте, як показано у підрозділі 1.3.2, різні типи несправностей генерують вібраційні патерни принципово різного просторового масштабу на спектрограмі: дефекти підшипників проявляються вузькими локалізованими імпульсами, що ефективніше захоплюються дрібними ядрами згортки (3×3), тоді як дисбаланс та зміщення осей формують протяжні спектральні смуги, що краще виявляються крупними ядрами (5×5). Фіксоване рівноважне об'єднання гілок не дозволяє моделі адаптувати баланс між масштабами залежно від характеру вхідного сигналу.

Для усунення цього обмеження у даній роботі запропоновано доповнити архітектуру MSCNN блоком міжгілкової уваги каналів (Channel Attention), що автоматично зважує внесок кожного каналу об'єданого тензора залежно від характеристик конкретного вхідного зразка. Блок розміщується безпосередньо після операції конкатенації гілок і перед шаром Global Average Pooling (формула 2.42), що дозволяє здійснити адаптивне зважування каналів до стиснення просторової інформації.

Запропонований блок виконує три послідовні операції: стиснення (Squeeze), збудження (Excitation) та масштабування (Scale).

Операція стиснення (Squeeze) перетворює кожен двовимірну карту ознак

каналу в одне скалярне значення, що є глобальним описом активації даного каналу. Для цього до кожного каналу c об'єднаного тензора F розміром $H \times W \times C$ (де $C = C_A + C_B$ – сумарна кількість каналів обох гілок) застосовується глобальний середній пулінг, аналогічно до формули GAP:

$$u[c] = \frac{1}{HW} \sum \sum F[i, j, c], \quad i = 0 \dots (H - 1), \quad j = 0 \dots (W - 1), \quad c = 0 \dots (C - 1), \quad (2.60)$$

де $u[c]$ – скалярне значення середньої активації каналу c ; $F[i, j, c]$ – значення карти ознак у позиції (i, j) каналу c ; H та W – просторові розміри карти ознак; i та j – просторові індекси; C – загальна кількість каналів після конкатенації ($C = C_A + C_B$); c – індекс каналу від 0 до $C-1$. Результатом є вектор u розмірності C , кожен елемент якого характеризує середній рівень активації відповідного каналу по всій площі карти ознак. Позначення u обрано для відрізнення від контекстного вектора z механізму часової уваги (підрозділ 2.3.4).

Операція збудження (Excitation) обчислює вагові коефіцієнти значущості для кожного каналу на основі вектора u . Для моделювання нелінійних залежностей між каналами та обмеження кількості додаткових параметрів використовується двошаровий повнозв'язний блок із проміжним стисненням розмірності:

$$s = \sigma(W_2 * RELU(W_1 * u + b_1) + b_2), \quad (2.61)$$

де W_1 – матриця першого шару розміром $\left(\frac{C}{r}\right) \times C$, що стискає вектор u у r разів; r – коефіцієнт стиснення; b_1 – вектор зміщень першого шару розмірності $\frac{C}{r}$; ReLU – функція активації (формула 2.38), що забезпечує нелінійність; W_2 – матриця другого шару розміром $C \times \left(\frac{C}{r}\right)$, що відновлює розмірність до C ; b_2 – вектор зміщень другого шару розмірності C ; σ – функція сигмоїд (формула 2.49), що обмежує вихідні значення діапазоном $(0, 1)$; s – результуючий вектор вагових коефіцієнтів розмірності C .

Коефіцієнт стиснення r визначає компроміс між виразністю та кількістю параметрів: менше r дає більше параметрів і вищу виразність, більше r – менше параметрів і сильнішу регуляризацию. У даній роботі обрано $r = 4$, що забезпечує достатню виразність при мінімальному збільшенні складності моделі.

Результуючий вектор s розмірності C містить вагові коефіцієнти значущості для кожного каналу: значення $s[c]$, близьке до 1, означає, що канал c є інформативним для поточного зразка і його вплив слід зберегти; значення, близьке до 0, – що канал є малоінформативним і його вплив слід послабити.

Операція масштабування (Scale) поелементно множить кожен канал карти ознак на відповідний ваговий коефіцієнт:

$$F_{CA}[i, j, c] = s[c] * F[i, j, c], \quad (2.62)$$

де F_{CA} – масштабована карта ознак; $s[c]$ – ваговий коефіцієнт каналу c з формули (2.61); $F[i, j, c]$ – вихідна карта ознак; i та j – просторові індекси; c – індекс каналу. Масштабована карта F_{CA} далі подається на шар Global Average Pooling. Оскільки перші C_A каналів тензора F відповідають гілці з ядрами 3×3 , а наступні C_B – гілці з ядрами 5×5 , блок фактично визначає відносну значущість дрібномасштабної та крупномасштабної гілок для кожного конкретного вхідного зразка.

Кількість додаткових навчальних параметрів, що вносить запропонований блок:

$$P_{CA} = 2 * C * \frac{C}{r} + \frac{C}{r} + C, \quad (2.63)$$

де P_{CA} – кількість параметрів блоку уваги каналів; C – загальна кількість каналів після конкатенації ($C = C_A + C_B$); r – коефіцієнт стиснення. Перший доданок $2 * C * \frac{C}{r}$ відповідає вагам двох матриць W_1 та W_2 , другий $\frac{C}{r}$ – зміщенням b_1 , третій C – зміщенням b_2 . Наприклад, при $C = 128$ (64 канали від кожної гілки) та $r = 4$: $P_{CA} = 2 \cdot 128 \cdot 32 + 32 + 128 = 8\,352$ параметри, що складає менше 5% від загальної кількості параметрів гібридної моделі і не створює ризику перенавчання.

Слід зазначити, що запропонований блок міжгілкової уваги каналів є самостійним механізмом, що не пов'язаний із механізмом часової уваги (Temporal Attention, підрозділ 2.3.4): перший працює на рівні просторових карт ознак MSCNN і визначає значущість кожного каналу (масштабу), другий – на рівні виходів шару LSTM і визначає значущість кожного часового кроку. Результати навчання та тестування гібридної моделі із запропонованим блоком, а також аналіз розподілу

вагових коефіцієнтів за класами несправностей наведено у підрозділі 3.6.

2.4.4. Навчання та оптимізація гіперпараметрів

Навчання гібридної моделі MSCNN-LSTM є процесом ітеративної оптимізації параметрів мережі з метою мінімізації цільової функції втрат. Для задачі багатокласової класифікації стандартною функцією втрат є категоріальна крос-ентропія (Categorical Cross-Entropy), що вимірює різницю між передбаченим розподілом ймовірностей p і істинним розподілом q [44]:

$$L = -(\sum(q_i * \log p_i)), \quad i = 1 \dots N_c, \quad (2.64)$$

де L – значення функції втрат; q_i – ймовірність i -го класу в навчальному прикладі ($q_i = 1$ для істинного класу, 0 для решти – так зване «one-hot» кодування); p_i – передбачена мережею ймовірність i -го класу (вихід Softmax); N_c – кількість класів; i – індекс класу. Крос-ентропія особливо чутлива до ситуацій коли мережа впевнено передбачає хибний клас: наприклад, при $p_i = 0,05$ і $q_i = 1$ штраф становить $-\log(0,05) \approx 3$, що стимулює мережу уникати таких помилок.

Оптимізація параметрів здійснюється алгоритмом Adam (Adaptive Moment Estimation) – адаптивним методом градієнтного спуску, що автоматично коригує швидкість навчання для кожного параметра окремо [46]. Алгоритм підтримує два ковзних моменти градієнтів:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla_t \quad (2.65)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \nabla_t^2, \quad (2.66)$$

де ∇_t – градієнт функції втрат по параметрах мережі на кроці t (позначення ∇_t обрано для відрізнення від вектора-кандидата g_t у формулах LSTM, підрозділ 2.3.2); m_t – перший момент (ковзне середнє градієнтів); v_t – другий момент (ковзне середнє квадратів градієнтів); $\beta_1 = 0,9$ і $\beta_2 = 0,999$ – стандартні коефіцієнти затухання [46]. На перших ітераціях навчання моменти m_t та v_t зміщені до нуля, тому застосовується корекція зміщення:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}; \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad (2.67)$$

де $\hat{m}_t$ та $\hat{v}_t$ – скориговані моменти; β_1^t та β_2^t – коефіцієнти β_1 та β_2 у степені t (номер ітерації).

На основі скоригованих моментів обчислюється оновлення кожного параметра мережі. Нехай θ – будь-який навчальний параметр мережі (елемент ваги фільтра, зміщення, елемент матриці LSTM тощо). Правило оновлення цього параметра на кроці t :

$$\theta_{t+1} = \theta_t - \left(\frac{\alpha \cdot \hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}} \right), \quad \varepsilon \approx 10^{-8}, \quad (2.68)$$

де θ_t – значення параметра мережі на кроці t ; θ_{t+1} – оновлене значення; α – швидкість навчання (learning rate); ε – мала константа для уникнення ділення на нуль. Фізичний зміст: параметри з великими стійкими градієнтами оновлюються повільніше ($\hat{v}_t$ велике), а з малими або нестабільними – швидше. Це забезпечує адаптивну швидкість навчання і робить Adam стійким до шумних градієнтів, що є типовим при навчанні на реальних вібраційних даних [46].

Процес навчання організовано як міні-пакетний стохастичний градієнтний спуск: на кожній ітерації з навчальної вибірки вибирається міні-пакет з B зразків, обчислюється функція втрат і виконується один крок оновлення параметрів. Одна повна прохідка по всіх навчальних зразках називається епохою (epoch).

Перед початком навчання вхідні дані розподіляються на три частини: навчальну вибірку (training set) – для оптимізації параметрів мережі; валідаційну вибірку (validation set) – для контролю за перенавчанням під час навчання; тестову вибірку (test set) – для фінальної оцінки якості моделі на даних, які не використовувались ні для навчання, ні для контролю. Валідаційна вибірка є ключовим інструментом: якщо якість на навчальній вибірці зростає, а на валідаційній – падає, це є ознакою перенавчання. При розподілі даних застосовується стратифікація – процедура, що забезпечує збереження пропорцій класів у кожній з трьох частин, що є важливим при наявності класів з різною кількістю зразків [44].

Навчання продовжується протягом заздалегідь визначеної кількості епох. Для запобігання перенавчання застосовується механізм збереження найкращої моделі (ModelCheckpoint): після кожної епохи порівнюється значення функції втрат на валідаційній вибірці з попереднім мінімумом, і якщо поточне значення є нижчим

– ваги моделі зберігаються. Для фінального тестування використовуються ваги з мінімальною функцією втрат, а не ваги останньої епохи, що гарантує вибір моделі з найкращою узагальнювальною здатністю [44].

До основних гіперпараметрів гібридної архітектури MSCNN-LSTM належать: швидкість навчання α – визначає розмір кроку оновлення параметрів, надто велике значення призводить до нестабільності, надто мале – до повільної збіжності; розмір міні-пакету B – кількість зразків, що обробляються одночасно, більший пакет дає стабільніші оцінки градієнта, але вимагає більше пам'яті; кількість фільтрів і розміри ядер у MSCNN – визначають ємність просторового блоку виділення ознак; розмір прихованого стану h у LSTM – визначає обсяг «пам'яті» рекурентного блоку; довжина послідовності T – кількість послідовних спектрограм, що подаються на LSTM; ймовірність Dropout p_{drop} – ступінь регуляризації для запобігання перенавчанню.

Конкретні значення гіперпараметрів, обрані для реалізованої системи, та обґрунтування їх вибору наведені у підрозділі 3.6.

2.4.5. Метрики оцінювання якості класифікації

Для об'єктивного оцінювання якості роботи класифікатора застосовується набір стандартних метрик, що базуються на матриці помилок (confusion matrix). Матриця помилок для задачі з N_c класами є квадратною матрицею розміром $N_c \times N_c$, де елемент $M[i, j]$ відображає кількість зразків класу i , що класифіковані моделлю як клас j . Діагональні елементи $M[i, j]$ відповідають правильно класифікованим зразкам кожного класу, а позадіагональні – помилковим класифікаціям [47]. Структуру матриці помилок для задачі багатокласової класифікації наведено на рисунку 2.12.

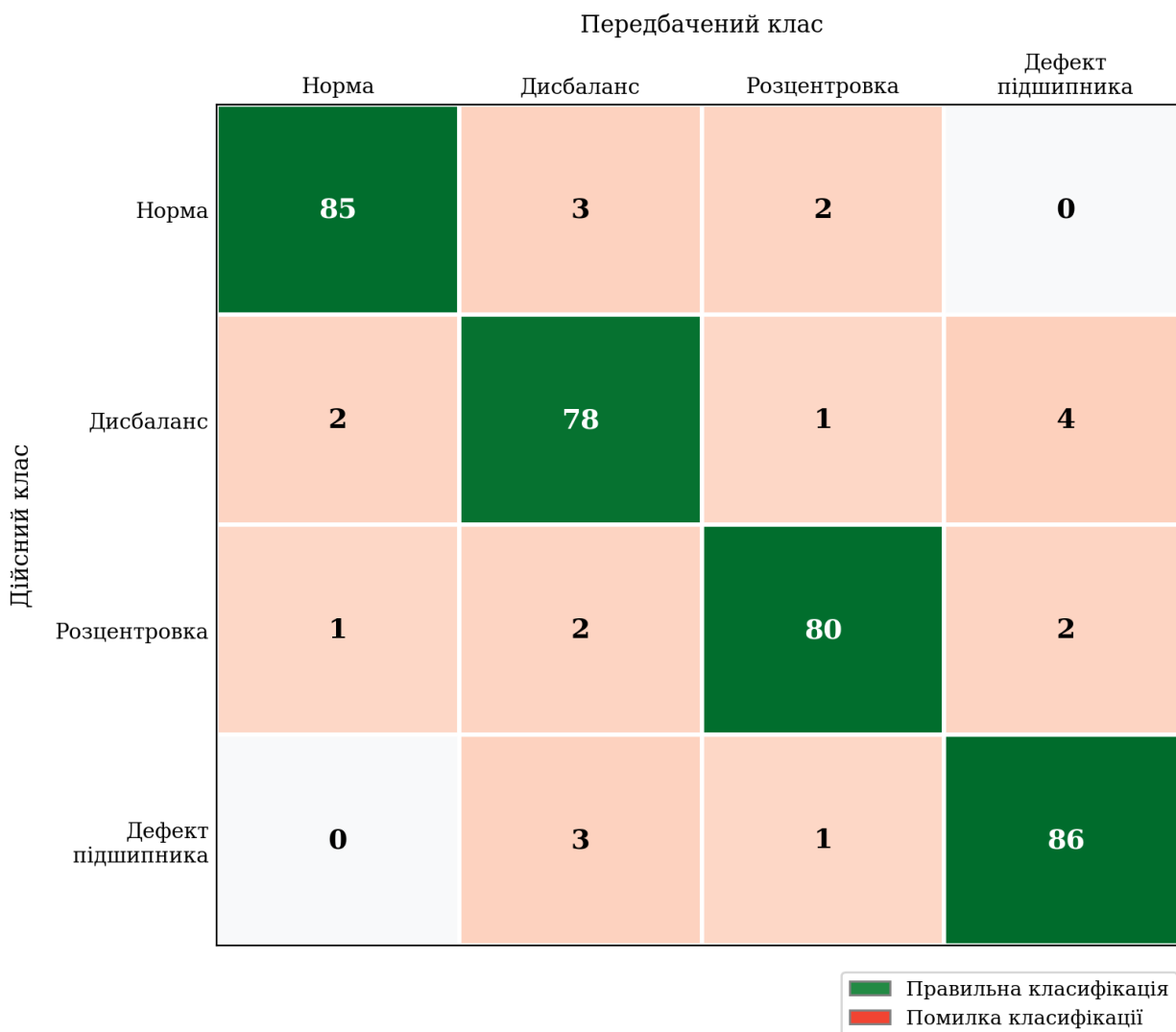


Рисунок 2. 12 Матриця помилок для задачі багатокласової класифікації

Точність (Accuracy) визначається як частка правильно класифікованих зразків серед усіх зразків вибірки:

$$Accuracy = \frac{\sum M[i,i]}{\sum \sum M[i,j]}, \quad (2.69)$$

де $M[i, i]$ – кількість правильно класифікованих зразків класу i (діагональні елементи матриці); $M[i, j]$ – елемент матриці помилок; чисельник – сума діагональних елементів; знаменник – загальна кількість зразків. Точність є інтуїтивно зрозумілою метрикою, проте може бути оманливою при незбалансованих класах: класифікатор, що завжди передбачає найпоширеніший клас, досягне високої точності без жодного корисного навчання [47].

Точність передбачення (Precision) для класу i визначає яка частка зразків,

класифікованих як клас i , дійсно належить до нього:

$$Precision_i = \frac{M[i,i]}{\sum M[j,i]}, j = 1 \dots N_c, \quad (2.70)$$

де $M[i,i]$ – правильно класифіковані зразки класу i ; $\sum M[j,i]$ – загальна кількість зразків, класифікованих моделлю як клас i (включаючи помилкові); j – індекс підсумовування по всіх класах.

Повнота (Recall) для класу i визначає яка частка дійсних зразків класу i знайдена класифікатором:

$$Recall_i = \frac{M[i,i]}{\sum M[i,j]}, j = 1 \dots N_c, \quad (2.71)$$

де $\sum M[i,j]$ – загальна кількість дійсних зразків класу i (включаючи ті що помилково класифіковані як інші класи).

Між точністю передбачення і повнотою існує природний компроміс: підвищення порогу класифікації підвищує Precision, але знижує Recall, і навпаки. Для об'єднання обох метрик в одну використовується F1-міра – гармонічне середнє Precision і Recall [47]:

$$F1_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i}, \quad (2.72)$$

де $Precision_i$ та $Recall_i$ – точність передбачення та повнота для класу i відповідно. F1-міра приймає значення від 0 до 1, де 1 відповідає ідеальній класифікації. Гармонічне середнє є строгим: навіть якщо одна з метрик близька до 1, але інша мала, F1 залишається малою. Це важлива властивість для задач діагностики, де пропущений дефект (низький Recall) є неприпустимим незалежно від того, наскільки висока Precision [47].

Для загальної оцінки якості класифікатора по всіх класах застосовується макро-усереднення (Macro-average):

$$F1_{macro} = \frac{1}{N_c} \sum F1_i, i = 1 \dots N_c, \quad (2.73)$$

де N_c – кількість класів; $F1_i$ – F1-міра для класу i . Макро-усереднення надає рівну вагу кожному класу незалежно від кількості зразків, що є важливим при незбалансованих вибірках – де рідкісний тип несправності є не менш важливим, ніж поширений.

Окрім числових метрик, для якісної оцінки здатності моделі до формування діагностично значущого внутрішнього представлення даних застосовується візуалізація простору ознак методом t-SNE (t-distributed Stochastic Neighbor Embedding). Суть методу полягає у проектуванні багатовимірних векторів ознак на двовимірну площину таким чином, що зразки, близькі у вихідному просторі, залишаються близькими на проекції, а віддалені – розташовуються далеко. Це дозволяє візуально оцінити якість розділення класів: якщо зразки кожного класу утворюють компактний ізольований кластер, модель ефективно виділяє дискримінативні ознаки; перекриття кластерів вказує на недостатню розрізнявальну здатність.

Для побудови проекції використовуються вихідні вектори внутрішнього представлення моделі – як правило, виходи передостаннього шару перед класифікатором Softmax. Основним параметром t-SNE є перплексія (perplexity) – величина, що визначає ефективну кількість сусідів, які враховуються при побудові проекції. Значення перплексії обирається залежно від обсягу вибірки: для малих вибірок (десятки–сотні зразків) рекомендуються значення 5–15, для великих (тисячі) – 30–50 [47]. Результати візуалізації простору ознак наведено у підрозділі 3.6.

Таким чином, у підрозділі 2.4 обґрунтовано доцільність поєднання MSCNN і LSTM в гібридній архітектурі, описано принцип їх з'єднання через часову послідовність ознакових векторів і спільне end-to-end, запропоновано блок міжгілкової уваги каналів для адаптивного зважування ознак різного масштабу, визначено процедуру навчання на основі функції крос-ентропії та алгоритму Adam, а також обґрунтовано набір метрик оцінювання якості класифікації та метод візуалізації простору ознак.

2.5. Методи оцінки залишкового ресурсу та прийняття діагностичних рішень

Класифікація стану обертового устаткування засобами гібридної архітектури MSCNN-LSTM, описаної у підрозділі 2.4, є лише першим кроком у системі предиктивного технічного обслуговування. Кінцева мета – не просто розпізнати

поточний тип стану, але й оцінити наскільки далеко зайшов процес деградації та скільки часу залишається до критичної відмови. Ця задача відома як оцінка залишкового ресурсу (Remaining Useful Life, RUL) і є центральною в системах прогностичного управління технічним станом (Prognostics and Health Management, PHM) [51, 53, 57].

У даному підрозділі розглядається теоретичний апарат перетворення виходу нейромережі на кількісні показники стану та прогноз залишкового ресурсу: інтерпретація вектора ймовірностей класів, побудова індексу здоров'я, оцінка RUL на основі моделі Вейбулла, комбінований підхід з пороговими зонами та теоретична модель прийняття рішень.

2.5.1. Інтерпретація виходу нейромережі

Вихідним результатом роботи гібридної моделі MSCNN-LSTM є вектор ймовірностей $p = [p_0, p_1, \dots, p_{N_c-1}]$, де кожен компонент $p_i \in (0, 1)$ відображає ступінь впевненості мережі в тому, що поточний стан обладнання відповідає i -му класу, а сума всіх компонентів дорівнює 1. Індксація починається з нуля: клас 0 відповідає нормальному стану, класи $1, \dots, 1, \dots, (N_c - 1)$ – різним типам несправностей. Цей вектор є основою для всіх подальших діагностичних рішень і кількісних оцінок стану [47].

Прогнозований клас стану визначається як клас з максимальною ймовірністю:

$$\hat{y} = \operatorname{argmax}(p_i), i = 0, 1, \dots, (N_c - 1), \quad (2.74)$$

де $\hat{y}$ – ідентифікований тип стану (нормальний стан або один із типів несправностей); p_i – ймовірність i -го класу; N_c – кількість класів. Для кількісної оцінки впевненості класифікатора у поточному рішенні вводиться показник довіри (confidence score):

$$C = \max(p_i) * 100\%, i = 0, 1, \dots, (N_c - 1), \quad (2.75)$$

де C – показник впевненості класифікатора у відсотках. Значення $C \in (0, 100\%]$ характеризує наскільки однозначно мережа класифікує поточний стан. Наприклад, якщо $C = 95\%$ – мережа з високою впевненістю ідентифікує

конкретний тип стану. Якщо $C = 45\%$ – ймовірності кількох класів є близькими, що може свідчити про перехідний або змішаний стан деградації [47].

Важливо розрізняти два різних сигнали, які несе вектор p . По-перше, компонент p_0 (ймовірність нормального стану) є прямим показником ступеня відхилення від норми: чим менше p_0 , тим більша ймовірність несправності. По-друге, розподіл ймовірностей між класами несправностей $(p_0, p_1, \dots, p_{N_c-1})$ дозволяє ідентифікувати домінуючий тип дефекту. Обидва сигнали використовуються при побудові індексу здоров'я, що розглядається далі [57].

2.5.2. Індекс здоров'я обладнання

Індекс здоров'я (Health Index, HI) є узагальненим скалярним показником технічного стану обладнання, що відображає ступінь деградації на нормалізованій шкалі від 0 до 1. Значення $HI = 1$ відповідає ідеальному технічному стану без ознак несправностей; $HI = 0$ відповідає критичному стану або відмові. Перевагою HI є те, що він зводить складний багатовимірний вектор ймовірностей до єдиного числа, придатного для моніторингу, відображення на дашборді та як вхідний параметр для моделей прогнозування залишкового ресурсу [52, 55].

Концепція індексу здоров'я широко досліджується в контексті систем прогностичного управління технічним станом [52]. Різні підходи до його побудови передбачають як пряме використання виходу нейромережі, так і більш складні конструкції на основі деградаційних індикаторів. У даній роботі HI будується безпосередньо з вектора ймовірностей класів.

Нехай клас $i = 0$ відповідає нормальному стану, а класи $i = 1, \dots, (N_c - 1)$ відповідають різним типам несправностей (де N_c – загальна кількість класів). Сумарна ймовірність несправного стану:

$$p_{fault} = \sum p_i, i = 0, 1, \dots, (N_c - 1), \quad (2.76)$$

де p_{fault} – сумарна ймовірність того, що обладнання знаходиться в одному з несправних станів; $p_{fault} \in [0, 1]$; p_i – ймовірність i -го класу несправності з вектора p .

Індекс здоров'я визначається як доповнення сумарної ймовірності

несправностей до одиниці:

$$HI = 1 - p_{fault}, \quad (2.77)$$

Оскільки сума всіх компонентів вектора p дорівнює одиниці, індекс здоров'я дорівнює ймовірності нормального стану: $HI = p_0$. Чим вище ймовірність нормального стану, тим вищий індекс здоров'я. При $HI \rightarrow 1$ обладнання близьке до ідеального стану; при $HI \rightarrow 0$ – до критичного [55].

Для підвищення стійкості HI до короткочасних флуктуацій класифікатора застосовується ковзне середнє по W послідовних вимірюваннях:

$$HI_{smooth}(t) = \frac{1}{W} \sum HI(t - k), k = 0 \dots (W - 1), \quad (2.78)$$

Де $HI_{smooth}(t)$ – згладжене значення індексу здоров'я у момент часу t ; W – розмір вікна згладжування; $HI(t - k)$ – значення індексу здоров'я на k кроків раніше від поточного моменту. Наприклад, при $W = 5$ індекс здоров'я обчислюється як середнє по останніх 5 вимірюваннях, що усуває вплив поодиноких хибних класифікацій на загальну оцінку стану [49]. Характерний тренд індексу здоров'я та зони технічного стану наведено на рисунку 2.13.

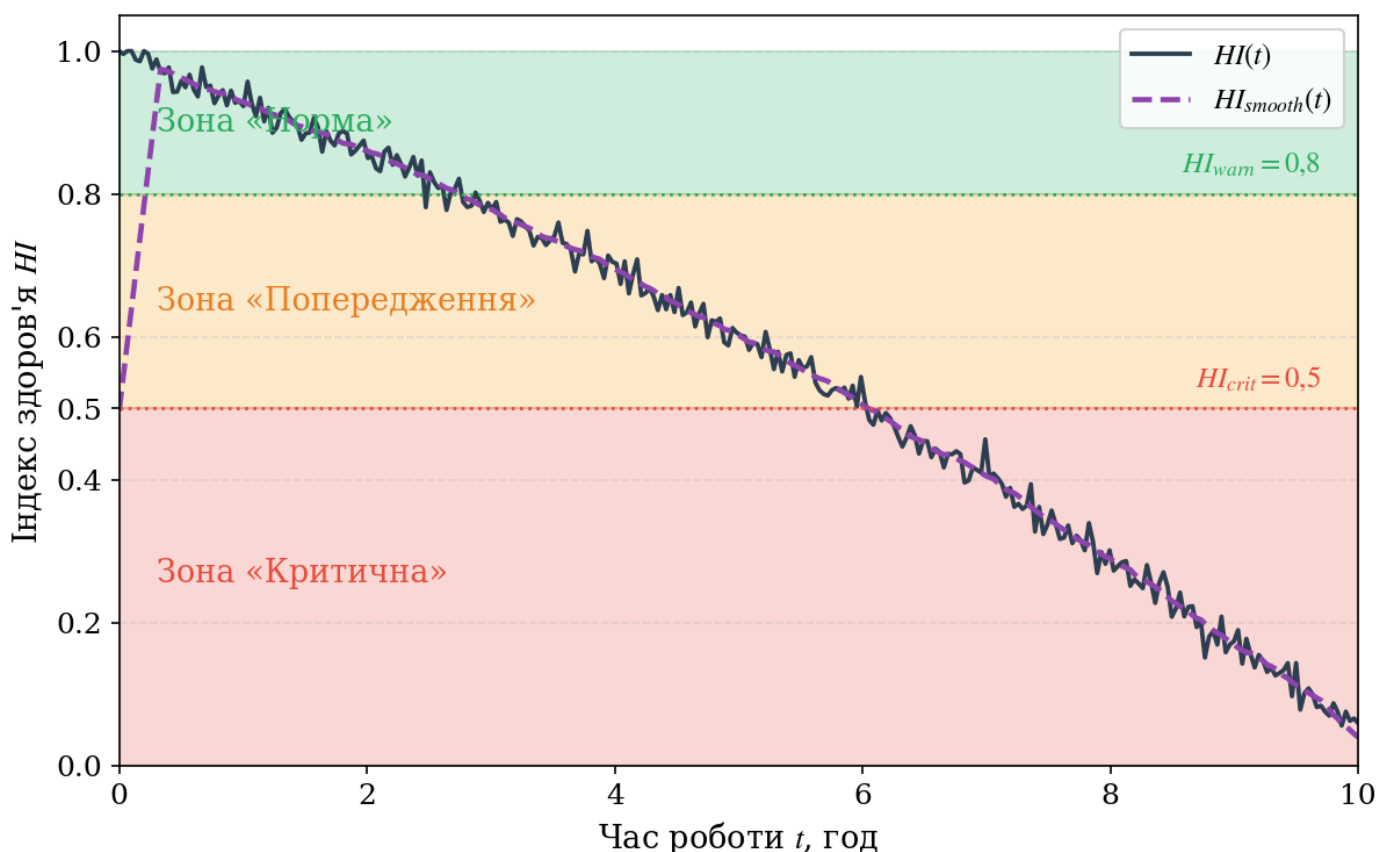


Рисунок 2. 13 Характерний тренд індексу здоров'я та зони технічного стану

Важливою властивістю HI є його монотонність при нормальному розвитку деградації: від початку відстеження до моменту відмови HI повинен демонструвати стійку тенденцію до зниження. У реальних умовах ця монотонність може порушуватись через зміни навантаження або умов роботи, тому згладжування є важливим елементом обробки [52].

2.5.3. Оцінка залишкового ресурсу на основі моделі Вейбулла

Залишковий ресурс (Remaining Useful Life, RUL) визначається як прогнозований час від поточного моменту до моменту відмови обладнання або досягнення неприйнятної рівня деградації. Оцінка RUL є ключовим виходом системи предиктивного технічного обслуговування – саме вона дозволяє планувати технічне обслуговування завчасно і раціонально [51, 56].

Серед методів оцінки RUL особливе місце займають моделі на основі розподілу Вейбулла – статистичного розподілу, що є стандартом у теорії надійності

для моделювання часу до відмови механічних компонентів. Розподіл Вейбулла є надзвичайно гнучким: шляхом вибору параметрів він може описувати ранні відмови (спадна інтенсивність відмов), випадкові відмови (постійна інтенсивність) або зносові відмови (зростаюча інтенсивність), що відповідає трьом фазам відомої «ванночкової кривої» [50, 54].

Двопараметричний розподіл Вейбулла визначається функцією щільності ймовірності (probability density function, PDF):

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}}, t \geq 0, \quad (2.79)$$

де $f(t)$ – щільність ймовірності відмови у момент часу t ; t – час роботи обладнання; β – параметр форми (shape parameter), що визначає характер зміни інтенсивності відмов; η – параметр масштабу (scale parameter), що характеризує характерний час до відмови (при $t = \eta$ функція надійності $R(\eta) = e^{-1} \approx 0.368$, тобто 63,2% виробів вже відмовило); e – експоненційна функція [50].

Функція надійності $R(t)$ (reliability function) – ймовірність того, що обладнання проіснує до моменту часу t без відмови:

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (2.80)$$

Функція інтенсивності відмов $h(t)$ (hazard function) – умовна ймовірність відмови в одиницю часу за умови, що до моменту t відмова ще не настала:

$$h(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}, \quad (2.81)$$

де позначення t , β , η у формулах $R(t)$ та $h(t)$ відповідають формулі $f(t)$. Поведінка $h(t)$ визначається параметром β . При $\beta < 1$ інтенсивність відмов спадає з часом – це характерно для ранніх відмов (дефекти виробництва). При $\beta = 1$ функція $h(t)$ є постійною – відмови випадкові, незалежні від часу (відповідає експоненційному розподілу). При $\beta > 1$ інтенсивність відмов зростає – це зносовий характер відмов, що є типовим для підшипників, валів та інших механічних вузлів після тривалої експлуатації. Для підшипників кочення типові значення β знаходяться в діапазоні 1,5–3,5 залежно від умов навантаження [50]. Вплив параметра форми β на функції надійності та інтенсивності відмов наведено

на рисунку 2.14.

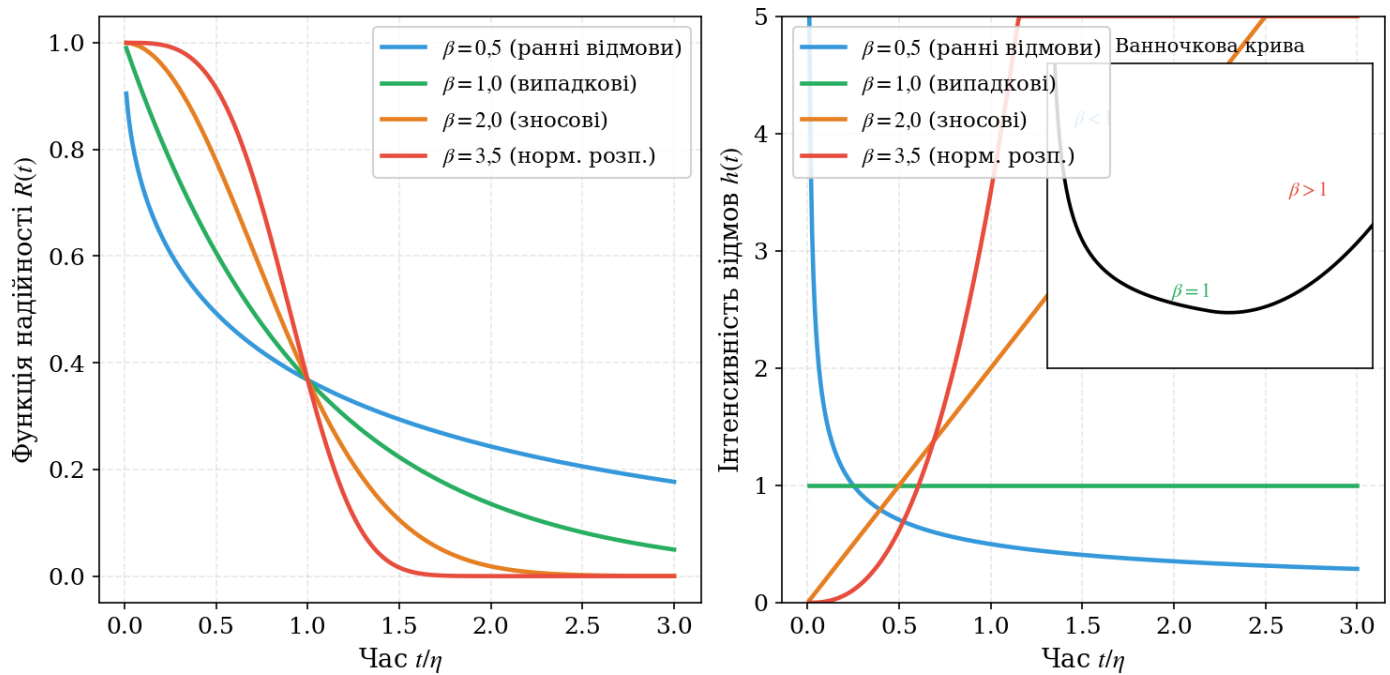


Рисунок 2. 14 Вплив параметра форми β на характер функцій надійності та інтенсивності відмов розподілу Вейбулла

У даній роботі використовується підхід, при якому поточний індекс здоров'я $HI(t)$ асоціюється з поточним значенням функції надійності $R(t)$: $HI(t) \approx R(t)$. Фізичне обґрунтування полягає в тому, що $HI = 1$ відповідає новому справному стану ($R = 1$), а $HI = 0$ відповідає відмові ($R = 0$), що відповідає семантиці функції надійності [55].

Маючи поточний час роботи t_{curr} і поточне значення $HI(t_{curr})$, знайдемо залишковий час до моменту досягнення критичного значення HI_{crit} . З рівняння $R(t) = HI_{crit}$:

$$e^{-\left(\frac{t_{fail}}{\eta}\right)^\beta} = HI_{crit}, \quad (2.82)$$

де t_{fail} – прогнозований час досягнення критичного стану; HI_{crit} – порогове значення індексу здоров'я, що відповідає критичному стану (визначається у підрозділі 2.5.4). Розв'язуючи відносно t_{fail} :

$$t_{fail} = \eta * (-\ln HI_{crit})^{\frac{1}{\beta}}, \quad (2.83)$$

де $\ln$ – натуральний логарифм. Залишковий ресурс RUL визначається як різниця між прогнозованим часом відмови і поточним часом роботи:

$$RUL = t_{fail} - t_{curr}, \quad (2.84)$$

де RUL – залишковий ресурс; t_{fail} – прогнозований час досягнення критичного стану з попередньої формули; t_{curr} – поточний час роботи обладнання.

Параметри β і η розподілу Вейбулла оцінюються на основі накопичених статистичних даних про відмови аналогічного типу обладнання за методом максимальної правдоподібності (Maximum Likelihood Estimation, MLE). Метод MLE знаходить параметри (β, η) , що максимізують спільну ймовірність спостережуваних даних про відмови [50]:

$$(\hat{\beta}, \hat{\eta}) = \operatorname{argmax} \sum \ln f(t_i, \beta, \eta), i = 1 \dots n, \quad (2.85)$$

де $\hat{\beta}$ та $\hat{\eta}$ – оцінки параметрів методом максимальної правдоподібності; t_i – зафіксований час відмови i -го зразка або цензуроване спостереження (час без відмови до моменту закінчення спостереження); n – кількість спостережень; $f(t_i, \beta, \eta)$ – щільність розподілу Вейбулла при заданих параметрах β та η .

Оцінки $\hat{\beta}$ та $\hat{\eta}$ є точковими і не відображають невизначеність, пов'язану з обмеженим обсягом статистичних даних. Для характеристики цієї невизначеності будується довірчий інтервал прогнозу RUL. Нехай задано рівень довіри (наприклад, 90%), якому відповідає квантиль стандартного нормального розподілу z (для 90% інтервалу $z = 1,645$). Граничні значення параметрів обчислюються як:

$$\beta_{lower} = \hat{\beta} - z * \sigma_{\beta}; \beta_{upper} = \hat{\beta} + z * \sigma_{\beta}, \quad (2.86)$$

$$\eta_{lower} = \hat{\eta} - z * \sigma_{\eta}; \eta_{upper} = \hat{\eta} + z * \sigma_{\eta}, \quad (2.87)$$

де β_{lower} та β_{upper} – нижня та верхня межі довірчого інтервалу для параметра β ; η_{lower} та η_{upper} – нижня та верхня межі для параметра η ; $\hat{\beta}$ та $\hat{\eta}$ – точкові оцінки параметрів; z – квантиль стандартного нормального розподілу, що визначається обраним рівнем довіри; σ_{β} та σ_{η} – стандартні відхилення оцінок параметрів, що обчислюються з діагоналі оберненої інформаційної матриці Фішера. Підстановкою граничних значень параметрів у формулу t_{fail} отримуються межі довірчого

інтервалу $[RUL_{lower}, RUL_{upper}]$. Ширина інтервалу зменшується зі збільшенням обсягу даних про відмови n [50, 54].

Практичне застосування описаного математичного апарату для оцінки залишкового ресурсу на модельному прикладі деградаційної кривої наведено у підрозділі 3.9.

2.5.4. Порогові зони індексу здоров'я та адаптивне визначення порогів

На практиці для прийняття рішень про технічне обслуговування використовується не лише значення RUL, але й поточний стан індексу здоров'я HI у поєднанні з типом виявленого дефекту. Комбінований підхід поєднує три джерела інформації: поточне значення HI та динаміку його тренду; прогноз RUL за моделлю Вейбулла; тип і впевненість класифікації $\hat{y}$ та C (підрозділ 2.5.1) [57].

Для зручності практичного використання простір значень HI розбивається на три зони стану обладнання [52]:

Зона «Норма» ($HI > HI_{warn}$): стан обладнання є задовільним, ознаки несправності відсутні або незначні. Рекомендується продовжувати нормальну експлуатацію і підвищити частоту моніторингу при наближенні до порогу.

Зона «Попередження» ($HI_{crit} < HI \leq HI_{warn}$): стан обладнання погіршився, виявлені ознаки несправності потребують уваги. Рекомендується планувати технічне обслуговування при найближчій можливості.

Зона «Критична» ($HI \leq HI_{crit}$): стан критичний, ризик відмови є значним. Рекомендується термінове технічне обслуговування або зупинка обладнання.

У літературі [52] рекомендовано фіксовані значення порогів $HI_{warn} = 0,8$ та $HI_{crit} = 0,5$, визначені на основі узагальненого досвіду експлуатації промислового обладнання. Проте ці значення не враховують характеристики конкретної моделі класифікації. Модель з високою точністю формує чіткий розподіл HI: для справних зразків HI близький до 1,0, для дефектних – близький до 0, з мінімальним перекриттям. У такому випадку фіксований поріг $HI_{warn} = 0,8$ виявляється заниженим – значна частина справних зразків має $HI > 0,9$, і поріг 0,8 не реагує на початок деградації до того моменту, поки HI не впаде суттєво. Водночас

фіксований $HI_{crit} = 0,5$ виявляється завищеним порівняно з фактичними значеннями НІ для дефектних станів, які можуть бути значно нижчими. Такий дисбаланс призводить до збільшення частки хибних тривог або до запізненого виявлення початку деградації.

Для усунення зазначеного обмеження у даній роботі запропоновано метод визначення порогових значень на основі квантилів фактичного розподілу індексу здоров'я, обчисленого навченою моделлю на навчальній вибірці.

Поріг попередження HI_{warn} визначається як нижній квантиль розподілу НІ для зразків нормального стану (клас 0):

$$HI_{warn} = Q_q(HI | y = 0), \quad (2.88)$$

де Q_q – функція квантиля рівня q ; q – рівень квантиля (наприклад, $q = 0,05$); НІ – значення індексу здоров'я; $y = 0$ – умова що зразок належить до класу нормального стану. Фізичний зміст: якщо поточне значення НІ опускається нижче HI_{warn} , це означає, що серед справних зразків навчальної вибірки менш ніж $q * 100\%$ мають таке або нижче значення. При $q = 0,05$ лише 5% справних зразків мають $HI < HI_{warn}$ – отже, поточний стан з високою ймовірністю не є нормальним, і необхідний перехід до зони попередження.

Поріг критичного стану HI_{crit} визначається як верхній квантиль розподілу НІ для дефектних зразків (класи $1, \dots, (N_c - 1)$):

$$HI_{crit} = Q_{1-q}(HI | y \in \{1, \dots, N_c - 1\}), \quad (2.89)$$

де Q_{1-q} – функція квантиля рівня $(1 - q)$; $y \in \{1, \dots, N_c - 1\}$ – умова що зразок належить до одного з класів несправностей; N_c – загальна кількість класів. Фізичний зміст: якщо НІ опускається нижче HI_{crit} , це означає, що $(1 - q) * 100\%$ дефектних зразків навчальної вибірки мають такі або нижчі значення НІ – поточний стан з високою ймовірністю є дефектним, і необхідні термінові заходи.

Перевагою запропонованого методу є його автоматичність та адаптивність: пороги обчислюються безпосередньо з навчальних даних і автоматично перераховуються при зміні моделі або розширенні набору даних без участі експерта. Для моделі з високою точністю класифікації (де розподіли НІ для

справних та дефектних зразків перекриваються мінімально) адаптивні пороги забезпечують одночасно більш раннє попередження та нижчу частку хибних тривог порівняно з фіксованими значеннями.

Для більш точної оцінки поточного стану та тенденції деградації додатково обчислюється швидкість зміни НІ:

$$\frac{\Delta HI}{\Delta t} = \frac{(HI(t) - HI(t - \Delta t))}{\Delta t}, \quad (2.90)$$

де $HI(t)$ – значення індексу здоров'я у поточний момент часу t ; $HI(t - \Delta t)$ – значення індексу здоров'я на попередньому вимірюванні; Δt – часовий інтервал між вимірюваннями. Від'ємне значення $\frac{\Delta HI}{\Delta t}$ свідчить про деградацію; чим більше за модулем це значення, тим швидше погіршується стан. Ця інформація доповнює статичне значення НІ і дозволяє виявляти прискорення деградації, що може вказувати на гострий розвиток несправності [49]. Зони технічного стану, відповідні порогові значення та рекомендовані дії підсумовано у таблиці 2.8.

Таблиця 2. 8 Зони технічного стану, порогові значення та рекомендовані дії [52]

Зона	Значення НІ	Орієнтовний RUL	Рекомендовані дії
Норма	$HI > 0,8$	$RUL > T_{plan}$	Продовжити нормальну експлуатацію. Підвищити частоту моніторингу при наближенні до порогу

Продовження таблиці 2.8

Попередження	$0,5 < HI \leq 0,8$	$T_{min} < RUL \leq T_{plan}$	Планувати технічне обслуговування. Ідентифікувати тип несправності за класифікатором. Підготувати запасні частини
Критична	$HI \leq 0,5$	$RUL \leq T_{min}$	Термінове технічне обслуговування або зупинка. Ризик відмови є значним. Обов'язково визначити тип дефекту

Примітка: значення порогів 0,8 та 0,5 є орієнтовними, рекомендованими у [52] для загального промислового обладнання. Для конкретної системи діагностики порогови визначаються адаптивним методом за формулами 2.88–2.89 на основі фактичного розподілу HI навченої моделі.

Практичні результати застосування запропонованого методу адаптивного визначення порогових зон, включаючи порівняння з фіксованими порогоми та аналіз впливу на частку хибних тривог, наведено у підрозділі 3.7.

2.5.5. Теоретична модель прийняття рішень

Теоретична модель прийняття рішень формалізує процес перетворення виходів системи діагностики на конкретні рекомендації щодо технічного обслуговування. Модель ґрунтується на правилах вигляду «ЯКІЩО умова ТО дія» і є спрощеним варіантом систем підтримки прийняття рішень, що широко застосовуються в інтелектуальних системах прогностичного управління технічним

станом [57].

Результатом роботи моделі є конкретна рекомендація d , обрана з множини можливих дій (наприклад: «норма», «підвищений контроль», «планове ТО», «термінова зупинка»). Рекомендація визначається на основі усіх доступних діагностичних показників:

$$d = D(\hat{y}, C, HI, RUL) \rightarrow \{d_1, d_2, \dots, d_K\}, \quad (2.91)$$

де d – прийняте рішення (конкретна рекомендація); D – функція прийняття рішень; $\hat{y}$ – прогнозований тип стану; C – показник впевненості класифікатора у відсотках; HI – індекс здоров'я обладнання; RUL – оцінка залишкового ресурсу; $\{d_1, d_2, \dots, d_K\}$ – множина з K можливих рекомендацій (наприклад: «норма», «підвищений контроль», «планове ТО», «термінова зупинка»); K – кількість можливих рекомендацій. Функція D реалізується у вигляді ієрархічної таблиці правил, що враховує зону НІ, тип виявленого дефекту $\hat{y}$ та рівень впевненості C [57]. Матрицю відповідності між зоною НІ, типом дефекту та рекомендованою дією наведено у таблиці 2.9.

Таблиця 2. 9 Теоретична модель прийняття рішень: матриця рекомендацій за типом дефекту та зоною НІ [14, 57]

Тип дефекту ($\hat{y}$)	Зона «Норма» НІ $> 0,8$	Зона «Попередження» $0,5 < \text{НІ} \leq 0,8$	Зона «Критична» $\text{НІ} \leq 0,5$
Нормальний стан	Нормальна експлуатація	Підвищений моніторинг	Позапланова перевірка (можлива помилка класифікатора або початок деградації)

Продовження таблиці 2.9

Дисбаланс ротора	Контроль та запис тренду	Планове балансування при найближчій зупинці	Термінове балансування. Перевірити кріплення та з'єднання
Розцентровка	Контроль та запис тренду	Планове центрування при найближчій зупинці	Термінова зупинка та центрування. Перевірити підшипники
Механічне послаблення	Контроль кріплень	Підтягнути кріплення при найближчій можливості	Термінова зупинка. Ризик каскадних пошкоджень
Вигин вала	Контроль та запис тренду	Планова перевірка вала при найближчій зупинці	Термінова зупинка. Заміна або правка вала
Пошкодження стержнів ротора	Підвищений електричний моніторинг	Планове технічне обслуговування двигуна	Термінова зупинка. Огляд та ремонт обмотки ротора
Дефект підшипника	Підвищений моніторинг, змастити підшипник	Планова заміна підшипника при найближчій зупинці	Термінова заміна підшипника. Перевірити суміжні вузли

Важливим аспектом теоретичної моделі є обробка ситуацій з низьким рівнем впевненості класифікатора. Якщо показник впевненості C нижчий за встановлений мінімальний поріг довіри C_{min} (наприклад, $C_{min} = 60\%$), результат класифікації вважається ненадійним. У такому випадку рішення приймається виключно на основі зони НІ без урахування типу дефекту $\hat{y}$. Це запобігає некоректним рекомендаціям при невизначеності класифікатора: краще видати загальне попередження на основі зони НІ, ніж вказати конкретний тип дефекту з низькою впевненістю [47].

Таким чином, у підрозділі 2.5 розглянуто теоретичний ланцюжок перетворення виходу нейромережі на діагностичне рішення: вектор ймовірностей $p \rightarrow$ тип несправності $\hat{y}$ та впевненість $C \rightarrow$ індекс здоров'я НІ $\rightarrow$ залишковий ресурс RUL за моделлю Вейбулла $\rightarrow$ зона стану та адаптивні пороги $\rightarrow$ рекомендація d . Описана теоретична база є основою для реалізації системи предиктивної діагностики, що розглядається у розділі 3.

2.6. Висновки до розділу

У розділі 2 сформовано математичний апарат системи предиктивної діагностики обертового енергетичного устаткування на основі нейромережевого аналізу вібраційних сигналів. Основні результати:

1. Визначено повний ланцюжок попередньої обробки вібраційного сигналу: від статистичних показників часової та частотної областей з базовим класифікатором k-NN (підрозділ 2.1.1), через спектральний аналіз на основі FFT (підрозділ 2.1.2), до формування двовимірних спектрограм STFT з нормалізацією та логарифмічним масштабуванням (підрозділ 2.1.3). Обґрунтовано вибір параметрів STFT та їх вплив на якість спектрограм (підрозділ 2.1.4).

2. Описано архітектуру згорткових нейронних мереж для автоматичного виділення просторових ознак зі спектрограм: операція згортки, функції активації (ReLU, Softmax, сигмоїд), шари пулінгу (Max Pooling, GAP), методи нормалізації (Batch Normalization) та регуляризації (Dropout) (підрозділи 2.2.1–2.2.2). Обґрунтовано використання багатомасштабної архітектури MSCNN з паралельними гілками 3×3 та 5×5 для одночасного виявлення патернів різного

просторового масштабу (підрозділ 2.2.3).

3. Описано архітектуру рекурентних нейронних мереж LSTM з трьохвентильним механізмом керування пам'яттю для аналізу часових залежностей у послідовностях ознакових векторів (підрозділи 2.3.1–2.3.3). Розглянуто механізм часової уваги (Temporal Attention), що дозволяє мережі вибірково фокусуватись на найбільш діагностично значущих моментах часу у послідовності (підрозділ 2.3.4).

4. Обґрунтовано доцільність поєднання MSCNN та LSTM у гібридну архітектуру, де MSCNN виділяє просторові ознаки з кожної спектрограми, а LSTM аналізує їх часову динаміку (підрозділи 2.4.1–2.4.2). Запропоновано блок міжгілкової уваги каналів (Channel Attention), що автоматично зважує внесок гілок різного масштабу залежно від характеристик конкретного вхідного зразка (підрозділ 2.4.3). Визначено процедуру навчання на основі функції крос-ентропії та алгоритму Adam, а також набір метрик оцінювання якості класифікації (підрозділи 2.4.4–2.4.5).

5. Розроблено теоретичний апарат оцінки залишкового ресурсу: побудову індексу здоров'я НІ на основі виходу класифікатора (підрозділ 2.5.2), прогнозування RUL за допомогою моделі Вейбулла (підрозділ 2.5.3), а також теоретичну модель прийняття діагностичних рішень (підрозділ 2.5.5). Запропоновано метод адаптивного визначення порогових зон індексу здоров'я на основі квантилів фактичного розподілу НІ, що забезпечує автоматичне налаштування порогів без участі експерта (підрозділ 2.5.4).

Описаний математичний апарат є теоретичною основою для практичної реалізації системи предиктивної діагностики, що розглядається у розділі 3.

Розділ 3. Розробка та дослідження системи предиктивної діагностики

3.1. Загальна архітектура розробленої системи

У даному підрозділі описано загальну архітектуру розробленої системи предиктивної діагностики обертового енергетичного устаткування. Система реалізує повний цикл обробки вібраційних сигналів – від отримання вхідних даних до формування діагностичного рішення із кількісною оцінкою залишкового ресурсу та візуалізацією результатів. Архітектура побудована відповідно до математичного апарату, сформованого у розділі 2, та має модульну структуру із п'ятьма послідовними етапами обробки даних, що відповідають стандартному пайплайну систем прогностичного управління технічним станом (PHM) [51, 53].

Загальну структуру системи предиктивної діагностики із зазначенням формату даних на кожному етапі наведено на рисунку 3.1.

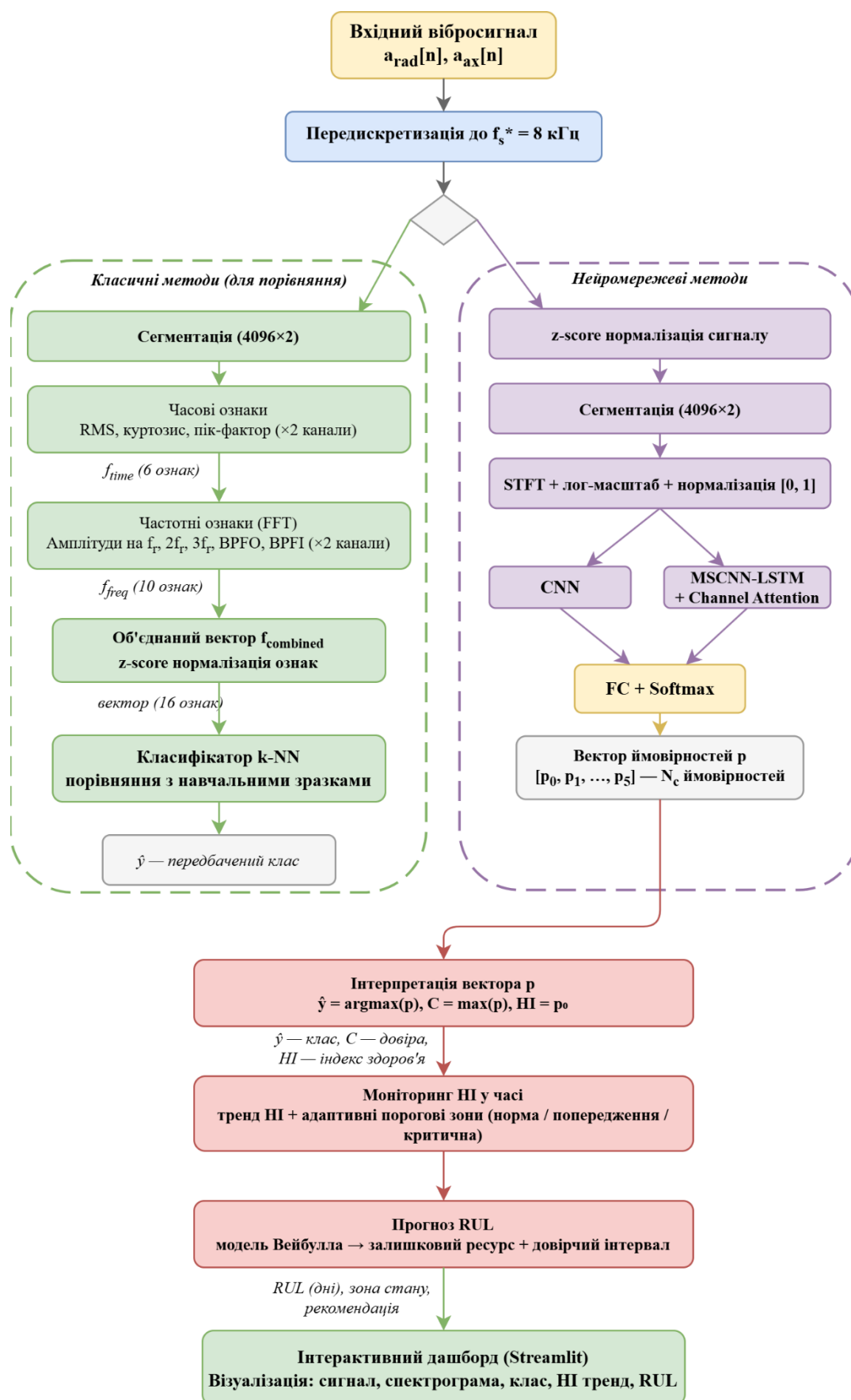


Рисунок 3. 1 Загальна структура системи предиктивної діагностики обертового енергетичного устаткування

Етап 1 – збір та попередня обробка даних. Система отримує дискретний двоканальний вібраційний сигнал, зареєстрований акселерометрами у радіальному та осьовому напрямках: два одновимірних масиви значень віброприскорення $a_{rad}[n]$ та $a_{ax}[n]$ у одиницях g із частотою дискретизації f_s . Вибір саме радіальної та осьової осей обумовлено тим, що ці два напрямки сукупно охоплюють діагностичну інформацію для всіх шести досліджуваних типів стану (підрозділи 1.3.1–1.3.2), а їх обґрунтування детально наведено у підрозділі 3.2.3. Оскільки дані надходять з двох різних вимірювальних систем із різними частотами дискретизації (50 кГц та 7,6 кГц), вхідні сигнали приводяться до єдиної частоти $f_s^* = 8$ кГц за процедурою передискретизації (формула 2.21) та нормалізуються методом z-score (формула 2.22) для усунення впливу різного калібрування акселерометрів. Нормалізовані сигнали сегментуються на фрагменти фіксованої довжини $N_{seg} = 4096$ відліків (0,512 с) без перекриття, що забезпечує статистичну незалежність зразків. Процедури узгодження даних та сегментації детально описані у підрозділах 3.2.4 та 3.3.1. На виході першого етапу формується набір двоканальних нормалізованих сегментів однакового формату незалежно від джерела даних.

Етап 2 – формування ознакових представлень. Для кожного нормалізованого сегменту паралельно формуються два типи ознакових представлень, призначених для різних методів класифікації. Перший тип – вектор статистичних ознак f (формула 2.11), що включає показники часової області (RMS, куртозис, пік-фактор по двох осях) та амплітуди спектру на характеристичних частотах дефектів (підрозділи 1.3.1–1.3.2), обчислені за допомогою FFT (формула 2.16). Цей вектор нормалізується методом z-score (формула 2.12) та використовується для класифікації базовим методом k-NN (формули 2.13–2.14). Другий тип – двоканальна спектрограма STFT розміром $128 \times 128 \times 2$, що формується шляхом послідовного застосування короточасного перетворення Фур'є (формула 2.24) з вікном Блекмана-Харріса (таблиця 2.2), обчислення спектру потужності (формула 2.26), логарифмічного масштабування (формула 2.27) та нормалізації до діапазону $[0, 1]$ (формула 2.28). Спектрограми двох каналів об'єднуються у двоканальний тензор, де перший канал відповідає радіальній осі, другий – осьовій. Параметри

STFT, їх обґрунтування та повна процедура формування спектрограм описані у підрозділі 3.3.2. На виході другого етапу кожен сегмент має подвійне представлення: числовий вектор ознак для класичних методів та двоканальна спектрограма для нейромережевих.

Етап 3 – класифікація стану. Двоканальні спектрограми подаються на вхід нейромережевої моделі. У роботі послідовно досліджено дві архітектури зростаючої складності: стандартну CNN (підрозділ 2.2), що класифікує кожен спектрограму окремо, та гібридну модель MSCNN-LSTM (підрозділ 2.4), що аналізує послідовності з $T = 5$ спектрограм із урахуванням часових залежностей. Гібридна модель включає багатомасштабну CNN з паралельними гілками 3×3 та 5×5 для одночасного виділення ознак різного масштабу (підрозділ 2.2.3), шар LSTM для моделювання часової динаміки (підрозділ 2.3.2), механізм часової уваги для зважування внеску кожного часового кроку (підрозділ 2.3.4), а також запропонований у роботі блок міжгілкової уваги каналів Channel Attention (підрозділ 2.4.3), що адаптивно зважує внесок паралельних гілок MSCNN. Навчання моделей здійснюється мінімізацією функції крос-ентропії (формула 2.64) алгоритмом Adam (формули 2.65–2.68). На виході третього етапу шар Softmax (формула 2.40) формує вектор ймовірностей $p = [p_0, p_1, \dots, p_5]$, де p_i – ймовірність належності поточного стану до i -го класу із шести досліджуваних. Результати навчання, підбору гіперпараметрів та порівняння моделей наведено у підрозділах 3.4–3.8.

Етап 4 – оцінка стану та прогнозування. Вектор ймовірностей p інтерпретується відповідно до теоретичного апарату підрозділу 2.5. Визначається прогнозований тип стану $\hat{y}$ (формула 2.74) та показник довіри C (формула 2.75). Обчислюється індекс здоров'я обладнання HI (формула 2.77), що відображає ступінь деградації на шкалі від 0 (критичний стан) до 1 (ідеальний). Для зменшення впливу короткочасних флуктуацій застосовується згладжування ковзним середнім (формула 2.78). На основі серії послідовних значень HI будується деградаційна крива, за якою оцінюється залишковий ресурс RUL за моделлю Вейбулла (формула 2.84) з довірчим інтервалом (формули 2.85–2.87). Для визначення зон технічного

стану – норма, попередження, критична – застосовується запропонований у роботі метод адаптивного визначення порогів на основі квантилів фактичного розподілу НІ (формули 2.88–2.89). Комбінація типу дефекту, рівня довіри, індексу здоров'я та залишкового ресурсу формує діагностичне рішення d (формула 2.91) відповідно до моделі прийняття рішень (таблиця 2.10). Практичні результати наведено у підрозділах 3.7.2 та 3.9.

Етап 5 – візуалізація та підтримка рішень. Всі діагностичні показники – тип стану, впевненість, індекс здоров'я, залишковий ресурс, зона стану та рекомендація – відображаються на інтерактивному веб-дашборді, розробленому засобами фреймворку Streamlit (підрозділ 3.10). Дашборд забезпечує візуалізацію вхідного сигналу, спектру FFT та спектрограми STFT, а також трендів індексу здоров'я та журналу діагностичних подій, що забезпечує прозорість прийняття рішення для оператора. Streamlit обрано як інструмент швидкого прототипування; для промислового впровадження результати діагностики можуть бути інтегровані в існуючу SCADA-систему об'єкта через стандартні протоколи обміну даними.

Модульна архітектура системи забезпечує можливість заміни або модифікації окремих блоків без впливу на решту компонентів: наприклад, заміну моделі класифікації, зміну параметрів STFT або розширення набору класів несправностей.

3.1.1. Програмне середовище та інструментальні засоби

Реалізація системи предиктивної діагностики здійснена мовою програмування Python версії 3.10 – стандартною мовою для задач машинного навчання та обробки сигналів, що підтримується усіма основними бібліотеками та фреймворками [59].

Середовищем розробки та навчання нейромережових моделей обрано Google Colaboratory – хмарну платформу, що надає безкоштовний доступ до графічних прискорювачів NVIDIA Tesla T4 із 16 ГБ відеопам'яті [59]. Навчання нейронних мереж потребує значних обчислювальних ресурсів: GPU забезпечує паралельне виконання матричних операцій, що прискорює процес навчання у десятки разів порівняно із центральним процесором. Google Colaboratory дозволяє уникнути

необхідності у спеціалізованому апаратному забезпеченні та забезпечує відтворюваність результатів на стандартизованій платформі. Слід зазначити, що середовище Colab використовується виключно на етапі навчання моделей; навчена модель зберігається як файл формату HDF5 (.h5) та може бути розгорнута на будь-якому комп'ютері без GPU – класифікація одного зразка займає мілісекунди на звичайному процесорі.

Для побудови, навчання та тестування неймережевих моделей використано фреймворк TensorFlow 2.x у поєднанні з високорівневим API Keras [59]. TensorFlow надає вбудовані засоби для побудови згорткових (підрозділ 2.2), рекурентних (підрозділ 2.3) та гібридних архітектур (підрозділ 2.4), включаючи шари TimeDistributed для обробки послідовностей спектрограм моделлю MSCNN-LSTM.

Обробка вібраційних сигналів здійснюється засобами бібліотек NumPy та SciPy [60]. Бібліотека NumPy забезпечує ефективні операції з багатовимірними масивами даних – від зберігання сигналів до матричних обчислень. Модуль scipy.signal бібліотеки SciPy використовується для трьох ключових операцій: передискретизації сигналів (функція resample_poly, формула 2.21), обчислення короткочасного перетворення Фур'є (функція stft, формула 2.24) та цифрової фільтрації. Візуалізація результатів – графіки сигналів, спектрів, спектрограм, кривих навчання та матриць помилок – виконується бібліотекою Matplotlib.

Для оцінювання якості класифікації використано бібліотеку scikit-learn, що надає реалізацію метрик Accuracy (формула 2.69), Precision (формула 2.70), Recall (формула 2.71), F1-score (формула 2.72), побудову матриці помилок, а також стратифікований розподіл вибірки на навчальну та тестову частини. Візуалізація простору ознак методом t-SNE (підрозділ 2.4.5) також реалізована засобами scikit-learn.

Інтерактивний дашборд системи діагностики розроблено засобами фреймворку Streamlit [61] – інструменту для створення веб-додатків мовою Python. Streamlit дозволяє побудувати повнофункціональний інтерфейс оператора без окремої розробки фронтенд-частини, що є оптимальним рішенням для прототипування та демонстрації працездатності підходу. Для побудови

інтерактивних графіків у складі дашборду використано бібліотеку Plotly, що забезпечує масштабування, наведення курсора для отримання точних значень та збереження графіків. Усі зазначені програмні засоби є вільно поширюваними з відкритим вихідним кодом.

Навчання моделі CNN на вибірці з 2116 спектрограм розміром $128 \times 128 \times 2$ при використанні GPU Tesla T4 займає від 5 до 15 хвилин залежно від кількості епох. Навчання гібридної моделі MSCNN-LSTM на вибірці з 234 послідовностей потребує від 15 до 40 хвилин. Етапи попередньої обробки сигналів – передискретизація, нормалізація, формування спектрограм – виконуються на центральному процесорі без потреби у графічному прискорювачі.

3.2. Експериментальна база дослідження

Вибір експериментальних даних є одним із визначальних етапів при розробці системи предиктивної діагностики, оскільки від якості, повноти та репрезентативності вибірки безпосередньо залежить достовірність отриманих результатів та можливість їх узагальнення на реальні умови експлуатації.

У розділі 1 було розглянуто сім характерних типів стану обертового устаткування: нормальний стан, дисбаланс ротора, зміщення осей (розцентровка), механічне послаблення, вигин вала, пошкодження стержнів ротора та дефект підшипника (підрозділ 1.3.2). Для експериментальної верифікації розроблених методів необхідні вібраційні дані, що охоплюють максимально широкий спектр із зазначених несправностей.

Жоден окремий публічно доступний датасет на сьогодні не містить повного набору всіх розглянутих типів стану. Зокрема, відсутні верифіковані публічні датасети із вібраційними записами достатнього обсягу для дефектів механічного послаблення та вигину вала [62]. У зв'язку з цим у роботі використано два взаємодоповнюючих публічних датасети: MAFAULDA (Machinery Fault Database), створений дослідницькою групою Федерального університету Ріо-де-Жанейро [63], та IEEE Broken Rotor Bar Database, створений лабораторіями Університету Сан-Паулу [64].

Датасет MAFAULDA забезпечує дані для чотирьох типів несправностей:

дисбалансу ротора, зміщення осей, дефекту зовнішнього та внутрішнього кілець підшипника. Датасет IEEE BRB доповнює експериментальну базу даними щодо пошкодження стержнів ротора. Дефекти підшипників у датасеті MAFAULDA представлені окремо для зовнішнього та внутрішнього кілець, що дозволяє розділити їх на два самостійних класи – з огляду на суттєві відмінності у спектральних проявах цих дефектів: зовнішнє кільце генерує ударні імпульси на частоті BPFO (формула 1.2), внутрішнє – на частоті BPFI (формула 1.3). Таким чином, експериментальна база охоплює шість класів стану обладнання: нормальний стан та п'ять типів несправностей. Дефекти механічного послаблення та вигину вала, незважаючи на їх розгляд у теоретичному розділі, не включено до експериментального дослідження через відсутність публічно доступних вібраційних даних належної якості.

3.2.1. Датасет MAFAULDA

Датасет MAFAULDA (Machinery Fault Database) являє собою колекцію з 1951 багатоканального запису вібраційних сигналів, зареєстрованих на лабораторному стенді SpectraQuest Machinery Fault Simulator (MFS) Alignment-Balance-Vibration Trainer (ABVT). Стенд складається з асинхронного електродвигуна потужністю 1/4 к.с. (≈ 186 Вт), з'єданого з валом через пружну муфту. На валу встановлено ротор, що дозволяє моделювати різні типи несправностей шляхом додавання дисбалансних мас, зміщення осей або встановлення дефектних підшипників [63].

Реєстрація вібраційних сигналів здійснювалась за допомогою двох комплектів промислових акселерометрів IMI Sensors: три одновісних датчики моделі 601A01 (чутливість 100 мВ/г, робочий діапазон частот 0,27–10 000 Гц, діапазон вимірювання ± 50 г) та один трьохвісний датчик моделі 604B31 з аналогічними характеристиками. Датчики встановлено на двох підшипникових опорах стенду – внутрішній (underhang) та зовнішній (overhang) – і реєструють вібраційний сигнал у трьох напрямках: осьовому (axial), радіальному (radial) та тангенціальному (tangential). Додатково встановлено тахометр для вимірювання частоти обертання ротора. Частота дискретизації становить $f_s = 50\,000$ Гц (50 кГц). Тривалість кожного запису – 5 секунд, що відповідає 250 000 відліків на канал.

Кожен запис містить 8 каналів даних [63].

Датасет охоплює наступні типи стану обертового устаткування:

- нормальний стан (49 записів при різних частотах обертання від 737 до 3686 об/хв);
- дисбаланс ротора з масами від 6 до 35 г (333 записи);
- горизонтальне зміщення осі від 0,5 до 2,0 мм (197 записів);
- вертикальне зміщення осі від 0,51 до 1,90 мм (301 запис);
- дефект зовнішнього кільця підшипника (558 записів);
- дефект внутрішнього кільця підшипника (513 записів).

Записи виконано при частотах обертання ротора від 737 до 3686 об/хв (від 12,3 до 61,4 Гц), що забезпечує варіативність умов роботи та дозволяє оцінити стійкість методів класифікації до зміни режиму навантаження. Широкий діапазон частот обертання є суттєвою перевагою датасету MAFAULDA, оскільки характеристичні частоти дефектів підшипників BPFO та BPF1 та гармоніки ротора пропорційно залежать від частоти обертання – модель, навчена на даних при різних швидкостях, має вищу здатність до узагальнення [63].

3.2.2. Датасет IEEE Broken Rotor Bar Database

Датасет IEEE Broken Rotor Bar Database створений лабораторіями Школи інженерії Університету Сан-Паулу (Бразилія) та містить електричні і механічні сигнали від трифазних асинхронних двигунів із різним ступенем пошкодження стержнів короткозамкненої обмотки ротора [64]. Пошкодження стержнів ротора є специфічним типом несправності, характерним для асинхронних двигунів, що проявляється появою бокових частотних смуг у спектрі вібрацій.

Експериментальний стенд складається з трифазного асинхронного двигуна потужністю 1 к.с. (746 Вт), з'єднаного з машиною постійного струму, що працює як генератор навантаження. Двигун – короткозамкненого типу з 34 стержнями ротора, номінальна швидкість 1715 об/хв при частоті живлення 60 Гц [64].

Вібраційні сигнали зареєстровані п'ятьма акселерометрами з чутливістю 10 мВ/(мм/с), робочим діапазоном частот 5–2000 Гц, що встановлені на приводному (drive end, DE) та неприводному (non-drive end, NDE) боках двигуна у радіальному

(горизонтальному та вертикальному) та осьовому напрямках. Частота дискретизації вібраційних сигналів становить $f_s = 7\,600$ Гц (7,6 кГц). Тривалість кожного запису – приблизно 18 секунд [64].

Датасет охоплює стани з 0 (справний), 1, 2, 3 та 4 зламаними стержнями ротора при восьми рівнях навантаження від 0,5 до 4,0 Нм із кроком 0,5 Нм. Для кожної комбінації стану та навантаження виконано 10 повторних вимірювань, що забезпечує статистичну достовірність даних. Загальна кількість записів – 400 (5 станів $\times$ 8 навантажень $\times$ 10 повторень). Варіативність навантаження є важливою характеристикою датасету, оскільки ковзання ротора s , а отже і положення бокових смуг у спектрі, суттєво залежать від рівня навантаження [64].

3.2.3. Вибір вимірювальних осей та обґрунтування

Вібраційний сигнал обертового устаткування є тривимірним векторним процесом, компоненти якого у радіальному, осьовому та тангенціальному напрямках несуть різну діагностичну інформацію залежно від типу несправності. Вибір осей вимірювання для системи діагностики є важливим методичним рішенням, що впливає на повноту та точність класифікації [5].

Радіальний напрямок є найбільш інформативним для дефектів, що генерують відцентрові сили у площині, перпендикулярній осі обертання. Дисбаланс ротора спричиняє радіальні коливання на частоті обертання f_r (підрозділ 1.3.2), а дефекти підшипників кочення проявляються імпульсними ударами на характеристичних частотах BPFO та BPFI (формули 1.2–1.3) переважно у радіальному напрямку [6].

Осьовий напрямок є визначальним для дефектів, що порушують осьову симетрію системи. Зміщення осей (розцентровка) – як паралельне, так і кутове – генерує значні осьові сили на частотах f_r та $2f_r$ (підрозділ 1.3.2). Пошкодження стержнів ротора, що спричиняють асиметрію магнітного поля, також генерують осьові коливання на частотах $f_r \pm 2sf_1$ [5].

Тангенціальний напрямок, хоча і несе додаткову інформацію, є найменш специфічним – тангенціальні вібрації зазвичай корелюють із радіальними і не додають принципово нових діагностичних ознак для досліджуваних типів дефектів. Крім того, тангенціальні акселерометри встановлюються не у всіх

промислових системах моніторингу, що обмежує практичну застосовність моделей, навчених на тангенціальних даних [6].

На підставі наведеного аналізу у даній роботі для кожного зразка використано два канали вібраційного сигналу: радіальний та осьовий. Таке поєднання забезпечує повноту діагностичної інформації для всіх шести досліджуваних типів стану: дисбаланс та підшипники діагностуються переважно по радіальній осі, зміщення осей та стержні ротора – по осьовій, нормальний стан характеризується низьким рівнем вібрацій на обох осях. Вибір відповідає рекомендаціям стандарту ISO 10816 щодо мінімально необхідної кількості точок вимірювання вібрацій обертового устаткування [6]. В обох датасетах – MAFAULDA та IEEE Broken Rotor Bar Database – присутні радіальний та осьовий канали, що забезпечує узгодженість даних із різних джерел.

3.2.4. Узгодження даних із різних джерел

Використання двох датасетів із різними параметрами реєстрації потребує приведення даних до єдиного формату перед подальшою обробкою. Основні відмінності між датасетами наведено у таблиці 3.1.

Таблиця 3. 1 Порівняння параметрів датасетів MAFAULDA та IEEE BRB Database

Параметр	MAFAULDA	IEEE BRB Database
Частота дискретизації	50 000 Гц	7 600 Гц
Тривалість запису	5 с	~18 с
Кількість осей вібрації	3 (radial, axial, tangential)	5 (2 radial, 2 axial, 1 vertical)
Тип двигуна	1/4 к.с., стенд MFS	1 к.с., трифазний АД
Частота обертання	737–3686 об/хв	~1715 об/хв
Типи несправностей	Норма, дисбаланс, зміщення, підшипник	Норма, зламаний стержень (1–4)

Як видно з таблиці 3.2, головною відмінністю є частота дискретизації: 50 кГц у MAFAULDA проти 7,6 кГц у IEEE BRB – різниця у 6,6 разів. Для забезпечення однорідності вхідних даних виконується процедура передискретизації (resampling) – приведення сигналів до єдиної частоти дискретизації f_s^* , математичний апарат якої описано у підрозділі 2.1.2 (формула 2.21).

Вибір цільової частоти дискретизації. Цільова частота обрана $f_s^* = 8\,000$ Гц (8 кГц) з наступних міркувань. По-перше, за теоремою Котельникова-Найквіста (підрозділ 2.1.2) максимальна частота, що може бути коректно представлена при $f_s^* = 8$ кГц, становить $f_{max} = \frac{f_s}{2} = 4\,000$ Гц. Цього достатньо для реєстрації всіх характеристичних частот дефектів, розглянутих у підрозділі 1.3.1: BPFO та BPFI для типових підшипників знаходяться в діапазоні 50–200 Гц, гармоніки ротора – до 500 Гц, а їх вищі гармоніки – до 2–3 кГц [5]. По-друге, частота 8 кГц є найближчою стандартною частотою, що перевищує f_s датасету IEEE BRB (7,6 кГц), що мінімізує спотворення при інтерполяції цього датасету. По-третє, зниження частоти дискретизації MAFAULDA з 50 кГц до 8 кГц зменшує обсяг даних у 6,25 разів, що прискорює обчислення STFT та навчання моделей без втрати діагностично значущої інформації [62].

Процедура передискретизації. Передискретизація реалізується відповідно до формули 2.21 за дворівневою процедурою: інтерполяцію у p разів та децимацію у q разів із застосуванням антиаліасингового фільтра нижніх частот для запобігання спектральним спотворенням. Для датасету MAFAULDA: НСД(50000, 8000) = 2000, звідки коефіцієнт інтерполяції $p = 4$ та коефіцієнт децимації $q = 25$. Для датасету IEEE BRB: НСД(8000, 7600) = 400, звідки $p = 20$ та $q = 19$; відносно невелика різниця частот (5,3%) забезпечує мінімальні спотворення.

Після передискретизації обидва датасети мають єдину частоту $f_s^* = 8\,000$ Гц, що забезпечує однакову частотну роздільну здатність спектрограм та однаковий частотний діапазон аналізу (0–4000 Гц) для всіх типів несправностей незалежно від джерела даних.

Нормалізація амплітуд. Акселерометри у двох датасетах мають різну

чутливість (100 мВ/г у MAFAULDA проти 10 мВ/(мм/с) у IEEE BRB), різний тип вимірюваної величини та різне калібрування. Внаслідок цього абсолютні значення сигналів можуть відрізнятись навіть для аналогічних режимів роботи. Для усунення цього впливу кожен фрагмент вібраційного сигналу нормалізується методом z-score (формула 2.22, підрозділ 2.1.3), що забезпечує нульове середнє та одиничну дисперсію незалежно від масштабу вхідного сигналу. Слід зазначити, що ця нормалізація застосовується до кожного фрагменту сигналу окремо на етапі попередньої обробки і є відмінною від z-score нормалізації вектора ознак (формула 2.12, підрозділ 2.1.1), яка застосовується до набору ознак по всій вибірці на етапі класифікації. Додатково, після обчислення спектрограм STFT кожна спектрограма нормалізується до діапазону $[0, 1]$ (формула 2.28, підрозділ 2.1.3), що забезпечує інваріантність вхідних даних нейронної мережі до абсолютного рівня вібрацій.

3.2.5. Зведена характеристика експериментальної бази

З урахуванням узгодження даних із двох джерел сформовано єдину експериментальну базу, що охоплює шість класів стану обертового устаткування. При формуванні класів виконано наступні операції об'єднання. Горизонтальне та вертикальне зміщення осей із датасету MAFAULDA об'єднано в один клас «зміщення осей», оскільки у підрозділі 1.3.2 розцентровка розглядається як єдиний тип несправності з характерними спектральними ознаками на частотах f_r та $2f_r$ незалежно від площини зміщення. Стани з 1, 2, 3 та 4 зламаними стержнями ротора із датасету IEEE BRB об'єднано в один клас «пошкодження стержнів ротора», що відповідає класифікації, прийнятій у підрозділі 1.3.2: усі ці стани проявляються однотипним спектральним патерном, відрізняючись лише інтенсивністю. Записи справного стану з датасету IEEE BRB (80 записів) не включено до класу 0 («нормальний стан»), оскільки вони отримані на іншому типі двигуна та при іншій частоті обертання, ніж записи MAFAULDA, і їх змішування може внести додаткову варіативність, не пов'язану з типом несправності. Зведену характеристику експериментальної бази наведено у таблиці 3.2.

Таблиця 3. 2 Зведена характеристика експериментальної бази дослідження

Клас	Тип стану	Джерело	Кількість записів
0	Нормальний стан	MAFAULDA	49
1	Дисбаланс ротора	MAFAULDA	333
2	Зміщення осей (розцентровка)	MAFAULDA (гориз. + верт.)	498
3	Дефект зовнішнього кільця підшипника	MAFAULDA	558
4	Дефект внутрішнього кільця підшипника	MAFAULDA	513
5	Пошкодження стержнів ротора	IEEE BRB Database	320
—	Разом	—	2271

Як видно з таблиці 3.2, кількість записів для різних класів є суттєво нерівномірною: від 49 записів для нормального стану до 558 для дефекту зовнішнього кільця підшипника – різниця більш ніж у 11 разів. Такий дисбаланс є типовим для реальних промислових даних, де справний стан реєструється значно рідше за несправний внаслідок специфіки експериментальних досліджень. Без коригування дисбалансу нейронна мережа буде схильна класифікувати нові зразки на користь класу з найбільшою представленістю, ігноруючи менш поширені класи – зокрема, критично важливий нормальний стан, на основі якого обчислюється індекс здоров'я НІ (формула 2.77). Методи балансування вибірки, що усувають цю проблему, описано у підрозділі 3.3.3.

3.3. Підготовка та попередня обробка даних

Після формування єдиної експериментальної бази з узгодженими параметрами (підрозділ 3.2) необхідно перетворити вхідні вібраційні сигнали на

формат, придатний для подальшої класифікації – як класичними методами (вектор ознак), так і нейромережевими (спектрограми). Цей процес включає три послідовних етапи: сегментацію безперервних записів на фрагменти фіксованої довжини (підрозділ 3.3.1), формування двоканальних спектрограм STFT із вибором та обґрунтуванням параметрів перетворення (підрозділ 3.3.2) та балансування вибірки з її розподілом на навчальну та тестову частини (підрозділ 3.3.3).

3.3.1. Сегментація сигналів та формування вибірки

Кожен запис у датасетах являє собою безперервний вібраційний сигнал тривалістю від 5 секунд (MAFAULDA) до 18 секунд (IEEE BRB). Для подальшої обробки записи поділяються на фрагменти (сегменти) фіксованої довжини, кожен з яких стає окремим зразком для класифікації. Тривалість сегменту визначається за формулою 2.33:

$$T_{seg} = \frac{N_{seg}}{f_s} = 4096/8000 = 0,512\text{с}, \quad (3.1)$$

де $N_{seg} = 4096$ відліків – обрана довжина сегменту; $f_s = 8000$ Гц – єдина частота дискретизації після передискретизації (підрозділ 3.2.4).

Обґрунтування вибору $N_{seg} = 4096$. Значення $4096 = 2^{12}$ обрано з урахуванням чотирьох факторів. По-перше, степінь двійки забезпечує максимальну ефективність алгоритму FFT, що використовується на етапі обчислення STFT (підрозділ 2.1.2). По-друге, тривалість 0,512 с є достатньою для захоплення декількох повних обертів ротора при будь-якій частоті обертання у діапазоні експериментальних даних: при мінімальній частоті 737 об/хв (12,3 Гц) сегмент охоплює $12,3 \times 0,512 \approx 6,3$ повних оберти, а при максимальній 3686 об/хв (61,4 Гц) – $61,4 \times 0,512 \approx 31,5$ оберти. Шість повних обертів є достатньою кількістю для виявлення періодичних компонент, пов'язаних із дисбалансом, зміщенням осей та дефектами підшипників (BPFO, BPFI), оскільки навіть найнижчі характеристичні частоти дефектів (~ 50 Гц для BPFO) генерують щонайменше 25 імпульсів за 0,512 с [5]. По-третє, частотна роздільна здатність FFT для сегменту довжиною 4096 відліків становить 1,95 Гц (формула 2.18), що достатньо для розрізнення близьких характеристичних частот. По-четверте, вибір 4096 відліків забезпечує достатню

кількість сегментів навіть для найкоротших записів: п'ятисекундний запис MAFAULDA дає 9 повних сегментів.

Сегментація без перекриття. Сегменти формуються послідовно без перекриття: кожен наступний сегмент починається з відліку, що безпосередньо слідує за останнім відліком попереднього сегменту. Залишок відліків наприкінці запису, що не утворює повного сегменту, відкидається. Відсутність перекриття між сегментами забезпечує їх статистичну незалежність, що є критично важливим для коректності подальшого розподілу на навчальну та тестову вибірки: при наявності перекриття суміжні сегменти поділяють значну частину даних, і потрапляння одного з них у навчальну вибірку, а іншого – у тестову, призводить до витоку даних (data leakage) та завищення оцінок якості класифікації [25].

Кількість повних сегментів, отриманих з одного запису, визначається як $\left\lfloor \frac{L}{N_{seg}} \right\rfloor$, де L – кількість відліків у записі після передискретизації. Для записів MAFAULDA: $L = 5 \times 8000 = 40\,000$ відліків, що дає $\lfloor 40000/4096 \rfloor = 9$ сегментів (залишок 3136 відліків відкидається). Для записів IEEE BRB: $L \approx 18 \times 8000 = 144\,000$ відліків, що дає $\lfloor 144000/4096 \rfloor = 35$ сегментів (залишок 640 відліків відкидається). Результати сегментації для кожного класу наведено у таблиці 3.3.

Таблиця 3. 3 Кількість сегментів за класами після сегментації

Клас	Тип стану	Записів	Сегментів на запис	Сегментів разом
0	Нормальний стан	49	9	441
1	Дисбаланс ротора	333	9	2 997
2	Зміщення осей	498	9	4 482
3	Дефект зовн. кільця	558	9	5 022
4	Дефект внутр. кільця	513	9	4 617

Продовження таблиці 3.3

5	Пошкодження стержнів ротора	320	35	11 200
—	Разом	2 271	—	28 759

Як видно з таблиці 3.3, загальний обсяг вибірки після сегментації становить 28 759 двоканальних сегментів. Суттєва різниця у кількості сегментів на запис між датасетами (9 проти 35) обумовлена різною тривалістю записів (5 с проти 18 с) і додатково посилює дисбаланс класів: клас 5 (стержні ротора) має найбільшу представленість (11 200 сегментів), що у 25 разів перевищує мінімальний клас 0 (441 сегмент). Це підтверджує необхідність процедури балансування, що описана у підрозділі 3.3.3.

3.3.2. Формування двоканальних спектрограм STFT

Для кожного нормалізованого двоканального сегменту формується спектрограма STFT, що перетворює одновимірний вібраційний сигнал у двовимірне частотно-часове представлення, придатне для обробки згортковими нейронними мережами. Процедура включає п'ять послідовних кроків відповідно до математичного апарату підрозділів 2.1.3–2.1.4; перші чотири кроки виконуються окремо для кожного каналу (радіального та осьового), п'ятий – формує остаточний двоканальний тензор.

Крок 1 – обчислення STFT. До нормалізованого сигналу $\tilde{x}[n]$ застосовується дискретне короткочасне перетворення Фур'є (формула 2.24) з вікном Блекмана-Харріса. Вибір вікнової функції обумовлено тим, що вікно Блекмана-Харріса забезпечує найвищий динамічний діапазон серед розглянутих у таблиці 2.2 вікнових функцій – 92 дБ, що є критично важливим при різниці амплітуд між гармоніками ротора та дефектними компонентами у 60–80 дБ: вікно з меншим динамічним діапазоном маскувало б слабкі дефектні складові боковими пелюстками сильніших гармонік. Результатом є масив комплексних коефіцієнтів STFT (формула 2.24) розміром 121×129 , де 121 – кількість часових вікон,

обчислена за формулою 2.25 при довжині сегменту $N_{seg} = 4096$ та кроці зсуву $H = 32$, а 129 – кількість частотних відліків від 0 до $\frac{f_s^*}{2} = 4000$ Гц, що визначається як $\frac{N_w}{2} + 1$ (підрозділ 2.1.4).

Крок 2 – спектр потужності. Обчислюється спектр потужності $S[m, k]$ (формула 2.26), що перетворює комплексні коефіцієнти STFT у дійснозначну карту розподілу енергії сигналу у частотно-часовій площині.

Крок 3 – логарифмічне масштабування. Спектр потужності перетворюється до логарифмічної шкали $S_{dB}[m, k]$ (формула 2.27), де $\varepsilon = 10^{-10}$ запобігає логарифму нуля. Логарифмування стискає динамічний діапазон значень, роблячи слабкі дефектні компоненти візуально порівнянними із сильними гармоніками ротора.

Крок 4 – нормалізація до [0, 1]. Логарифмічна спектрограма нормалізується за формулою 2.28. Нормалізація виконується окремо для кожної спектрограми, що забезпечує інваріантність до абсолютного рівня вібрацій конкретного вимірювання.

Крок 5 – формування двоканального тензору. Нормалізована спектрограма кожного каналу має розмір 129×121 (частотних відліків $\times$ часових вікон). Для приведення до квадратного формату, стандартного для згорткових нейронних мереж, спектрограма масштабується до розміру 128×128 білінійною інтерполяцією (функція `scipy.ndimage.zoom`). Зміна розмірності є мінімальною (менше 1% по кожній осі) і не спричиняє помітних спотворень діагностичної інформації. Масштабовані спектрограми радіального та осьового каналів об'єднуються по третій осі у єдиний тензор розміром $128 \times 128 \times 2$, де перший канал відповідає радіальній осі вібрації, другий – осьовій. Такий двоканальний формат аналогічний RGB-зображенню з двома каналами замість трьох і дозволяє згортковій мережі аналізувати просторові патерни обох осей одночасно, виявляючи як внутрішньоканальні, так і міжканальні кореляції.

Обґрунтування вибору параметрів STFT, що визначають розміри та якість результуючої спектрограми, наведено у таблиці 3.4.

Таблиця 3. 4 Параметри формування спектрограм STFT

Параметр	Значення	Обґрунтування
Довжина вікна N_w	256 відліків	$\Delta f = 31,25$ Гц – баланс частотної та часової роздільності
Крок зсуву H	32 відліки	$\Delta t = 4$ мс – достатня часова деталізація
Перекриття OL	87,5%	Висока щільність часових вікон
Вікнова функція	Блекман-Харріс	Динамічний діапазон 92 дБ
Часових вікон M	121	Стовпці спектрограми
Частотних відліків	129	Рядки спектрограми
Фінальний розмір	$128 \times 128 \times 2$	Стандартний формат для CNN

Як видно з таблиці 3.4, частотна роздільна здатність $\Delta f = 31,25$ Гц є суттєво нижчою, ніж роздільність повного FFT сегменту ($\Delta f = 1,95$ Гц, підрозділ 3.3.1). Це є наслідком принципу невизначеності (формула 2.29): підвищення часової роздільності $\Delta t = 4$ мс досягається за рахунок зниження частотної. Для задачі класифікації спектрограм нейромережею така роздільність є прийнятною, оскільки CNN розпізнає просторові патерни розподілу енергії у частотно-часовій площині, а не окремі спектральні піки. Детальний аналіз окремих характеристичних частот виконується на етапі FFT-аналізу з повною роздільністю (підрозділ 3.4.2).

Приклади нормалізованих двоканальних спектрограм для кожного з шести класів стану наведено на рисунку 3.2.

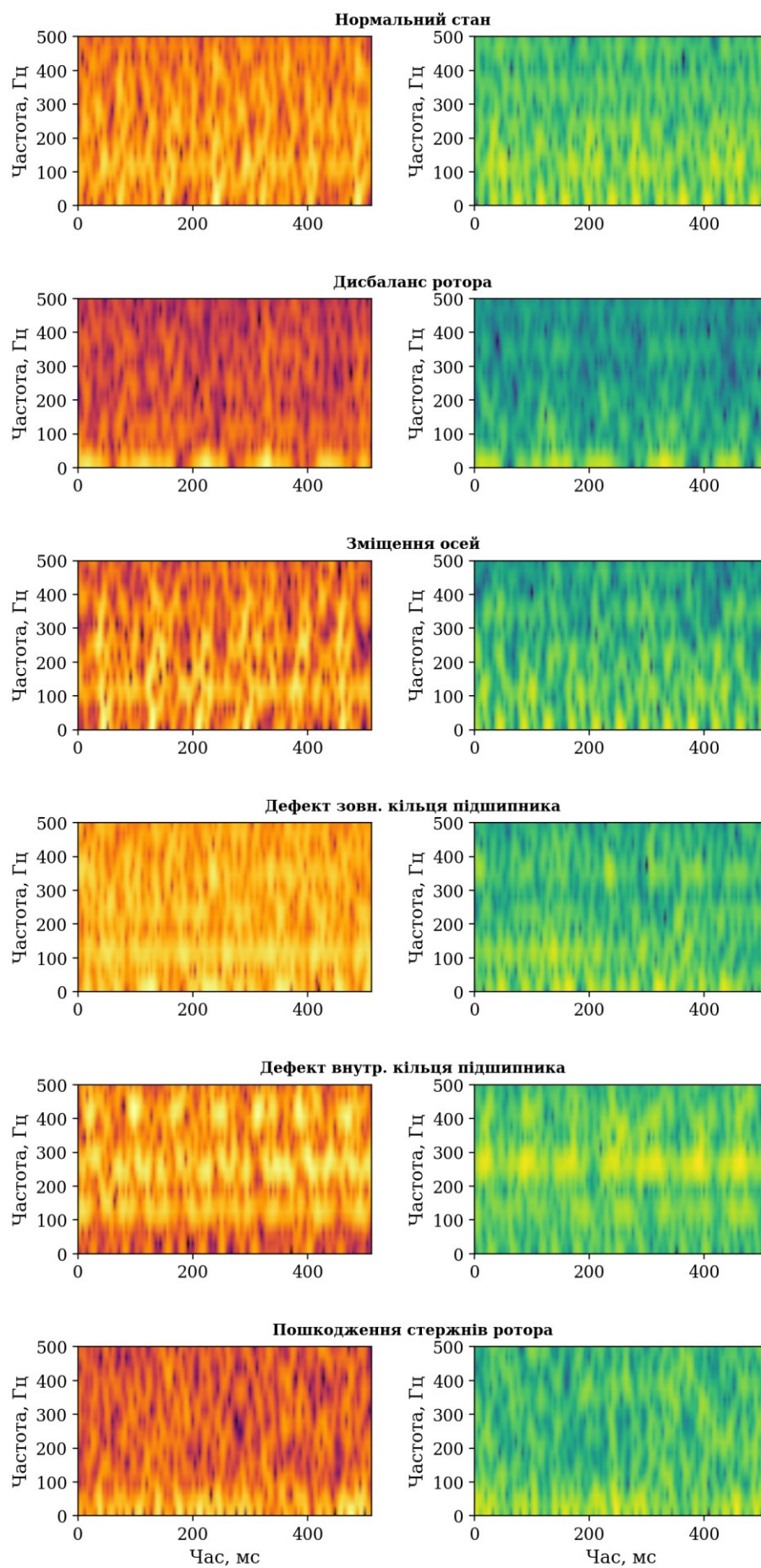


Рисунок 3. 2 Двоканальні спектрограми STFT

На рисунку 3.2 простежуються візуальні відмінності між класами. Спектрограма нормального стану (клас 0) характеризується відносно рівномірним розподілом енергії з низькою загальною інтенсивністю та слабо вираженими горизонтальними смугами на частоті обертання ротора. Дисбаланс ротора (клас 1) проявляється вираженою горизонтальною смугою на частоті обертання f_r , особливо помітною на радіальному каналі, що відповідає домінуючій відцентровій силі дисбалансу (підрозділ 1.3.2). Зміщення осей (клас 2) має подібний патерн, але з наявністю декількох горизонтальних смуг на частотах f_r та $2f_r$, причому осьовий канал демонструє підвищену інтенсивність порівняно з радіальним. Дефекти підшипників (класи 3 та 4) характеризуються широкосмуговими імпульсними патернами, що проявляються як вертикальні смуги підвищеної інтенсивності, розподілені по всьому частотному діапазону з характерною періодичною модуляцією амплітуди. Пошкодження стержнів ротора (клас 5) візуально найменш виражене на спектрограмі – проявляється концентрацією енергії у низькочастотній області з рівномірним розподілом по часовій осі; специфічні бокові смуги $f_r \pm 2sf_1$ розташовані занадто близько одна до одної ($\pm 3,96$ Гц) для розрізнення при частотній роздільності $\Delta f = 31,25$ Гц і детально аналізуються на FFT-спектрах у підрозділі 3.4.2. Наведені візуальні відмінності є основою для автоматичної класифікації спектрограм згортковими нейронними мережами.

3.3.3. Розподіл даних та балансування класів

Як було зазначено у підрозділі 3.3.1, кількість сегментів для різних класів є суттєво нерівномірною: від 441 (клас 0) до 11 200 (клас 5). Без коригування цього дисбалансу нейронна мережа оптимізуватиме функцію втрат (формула 2.64) переважно на зразках домінуючих класів, що призведе до систематичної класифікації рідкісних класів як більш поширених. Для класу 0 (нормальний стан), на основі якого обчислюється індекс здоров'я НІ (формула 2.77), низька точність класифікації є особливо критичною, оскільки хибне віднесення нормального стану до дефектного генеруватиме хибні тривоги у системі моніторингу.

Метод балансування. Для усунення дисбалансу застосовано метод випадкового підвибіркування (random undersampling): з кожного класу випадковим

чином обирається кількість сегментів, що дорівнює мінімальному класу. Мінімальним є клас 0 з 441 сегментом, тому з кожного класу обирається рівно 441 сегмент. Загальний обсяг збалансованої вибірки становить $441 \times 6 = 2\,646$ двоканальних спектрограм розміром $128 \times 128 \times 2$. Підвибіркування обрано замість альтернативного методу надвибіркування (oversampling), що дублює зразки мінорних класів, оскільки дублювання спектрограм не додає нової інформації для навчання і може призвести до перенавчання на повторюваних зразках [25].

Розподіл на навчальну та тестову вибірки. Збалансована вибірка розподіляється на навчальну (80%) та тестову (20%) частини методом стратифікованого розподілу (функція `train_test_split` бібліотеки `scikit-learn` з параметром `stratify`), що гарантує збереження рівного представництва класів в обох частинах. Для кожного класу: 353 спектрограми потрапляють до навчальної вибірки та 88 – до тестової. Загалом навчальна вибірка містить 2 118 спектрограм, тестова – 528. Розподіл вибірки наведено у таблиці 3.5.

Таблиця 3. 5 Розподіл збалансованої вибірки між навчальним та тестовим наборами

Клас	Тип стану	Навчальна (80%)	Тестова (20%)	Разом
0	Нормальний стан	353	88	441
1	Дисбаланс ротора	353	88	441
2	Зміщення осей	353	88	441
3	Дефект зовнішнього кільця	353	88	441
4	Дефект внутрішнього кільця	353	88	441
5	Пошкодження стержнів ротора	353	88	441

Продовження таблиця 3.5

—	Разом	2 118	528	2 646
---	-------	-------	-----	-------

Як видно з таблиці 3.5, після балансування та розподілу кожен клас представлений рівною кількістю зразків як у навчальній, так і у тестовій вибірці, що забезпечує коректне обчислення метрик класифікації (формули 2.69–2.73) без зміщення на користь окремих класів. Тестова вибірка з 88 спектрограм на клас є достатньою для статистично значущої оцінки якості класифікації. Слід зазначити, що описаний розподіл використовується для навчання та тестування моделі CNN, яка класифікує кожен спектрограму окремо. Для гібридної моделі MSCNN-LSTM, що обробляє послідовності з $T = 5$ спектрограм, формування вибірки потребує додаткової процедури створення часових послідовностей із послідовних сегментів одного запису, що детально описано у підрозділі 3.6.1.

3.4. Результати аналізу вібраційних сигналів класичними методами

Перед застосуванням нейромережових моделей доцільно оцінити базовий рівень точності класифікації, що досягається класичними методами обробки вібраційних сигналів. Це дозволяє, по-перше, кількісно обґрунтувати необхідність переходу до глибокого навчання, а по-друге, продемонструвати поступове підвищення точності при ускладненні методу аналізу – від часової області через частотну до нейромережової класифікації спектрограм. У даному підрозділі послідовно досліджуються два класичних підходи: аналіз статистичних показників часової області (підрозділ 3.4.1) та аналіз спектру FFT з виділенням характеристичних частот дефектів (підрозділ 3.4.2). Для класифікації в обох випадках використовується алгоритм k найближчих сусідів (k -NN, формули 2.13–2.14) – стандартний базовий класифікатор, що не потребує навчання нейромережі та є прийнятим у літературі для порівняльного аналізу [25].

3.4.1. Аналіз у часовій області

Формування вектора ознак часової області. У підрозділі 2.1.1 описано дев'ять статистичних показників вібраційного сигналу (формули 2.1–2.9), кожен з

яких характеризує певний аспект розподілу амплітуд у часовій області. Для формування вектора ознак f_{time} обрано три показники, що мають найвищу діагностичну інформативність для задачі класифікації дефектів обертового устаткування згідно з таблицею 2.1:

- середньоквадратичне значення RMS (формула 2.1) – інтегральний показник загального рівня вібрацій, що вимірюється в одиницях прискорення (g). RMS зростає при будь-якому підвищенні енергії сигналу і є основним параметром оцінки загального стану обладнання згідно стандарту ISO 10816;

- куртозис K (формула 2.5) – безрозмірний показник «гостроти» розподілу амплітуд. Для нормального (гауссового) розподілу $K = 3,0$. Імпульсні компоненти, характерні для дефектів підшипників кочення (підрозділ 1.3.1), створюють «важкі хвости» розподілу та підвищують куртозис значно вище 3,0. Сильна синусоїдальна компонента (характерна для дисбалансу) навпаки знижує куртозис нижче 3,0, оскільки чиста синусоїда має $K = 1,5$;

- пік-фактор CF (формула 2.6) – безрозмірне відношення пікового значення сигналу до RMS. Аналогічно до куртозису, CF зростає при появі різких імпульсів і знижується при домінуванні гармонічних складових.

Решту шість показників виключено з наступних причин. Середнє арифметичне (формула 2.2) для вібраційного сигналу з симетричними коливаннями завжди близьке до нуля незалежно від типу дефекту. Стандартне відхилення (формула 2.3) є надлишковим відносно RMS, оскільки при нульовому середньому $\sigma = RMS$. Максимальне абсолютне значення (формула 2.4) та імпульс-фактор (формула 2.9), що від нього залежить, є нестабільними – одиничний викид шуму може суттєво спотворити їх значення і вони мають низьку повторюваність між вимірюваннями. Асиметрія (формула 2.7) має низьку діагностичну цінність для обертового устаткування, де вібраційний сигнал зазвичай є симетричним відносно нуля. Форм-фактор (формула 2.8) є відношенням RMS до середнього модуля і мало змінюється для різних типів дефектів.

Кожен із трьох обраних показників обчислюється окремо для радіального та осевого каналів вібраційного сигналу, що формує шестивимірний вектор ознак

часової області: $f_{time} = [MS_{rad}, MS_{ax}, K_{rad}, K_{ax}, CF_{rad}, CF_{ax}]$.

Розділення на канали є принциповим, оскільки різні типи дефектів проявляються переважно на різних осях вібрацій (підрозділ 3.2.3): дисбаланс генерує підвищений рівень на радіальній осі, зміщення осей – на осьовій. Об'єднання каналів в один показник (наприклад, усереднення) призвело б до втрати цієї діагностичної інформації.

Зауваження щодо обчислення статистик. Статистичні показники обчислюються на сирому (передискретизованому, але не нормалізованому) вібраційному сигналі, тому RMS зберігає фізичний зміст у одиницях прискорення (g) і відображає реальний рівень вібрацій для кожного класу. Z-score нормалізація сигналу (формула 2.22), описана у підрозділі 3.2.4, застосовується лише на шляху формування спектрограм для нейромережевих методів, де абсолютний рівень вібрацій замінюється частотно-часовими патернами. Для класичного аналізу абсолютний рівень вібрацій є діагностично значущим і тому зберігається. Перед класифікацією методом k-NN до вектора ознак застосовується z-score нормалізація ознак (формула 2.12, підрозділ 2.1.1) – вона вирівнює числові діапазони різних показників (RMS у g, куртозис безрозмірний), але не змінює їх діагностичну інформативність.

Приклади вібраційних сигналів у часовій області для шести досліджуваних класів наведено на рисунку 3.3.

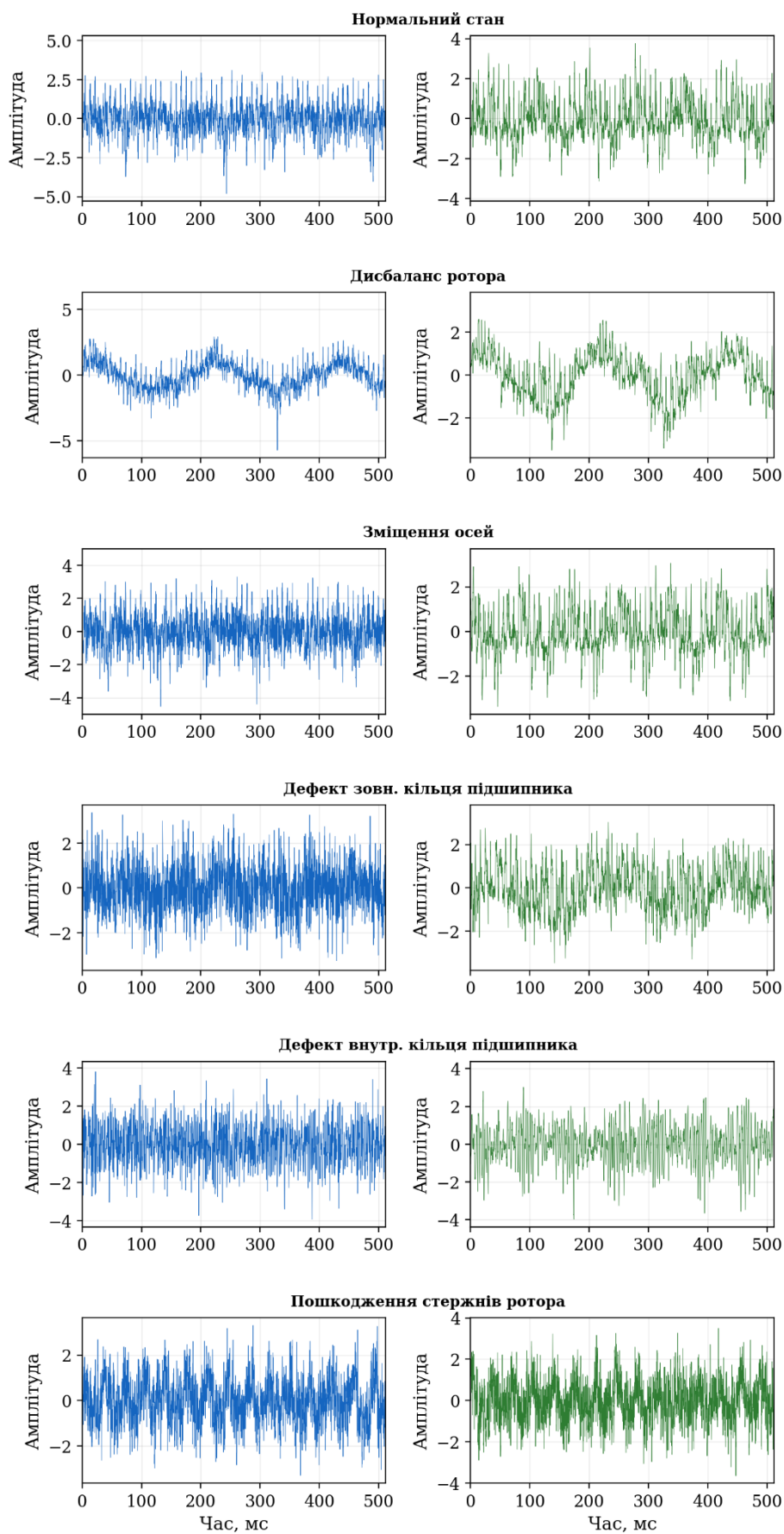


Рисунок 3. 3 Вібраційні сигнали у часовій області для шести типів стану обертового устаткування: ліва колонка – радіальна вісь, права – осьова

Як видно з рисунка 3.3, візуальне розрізнення класів у часовій області є обмеженим. Нормальний стан (клас 0) характеризується низькоамплітудним шумоподібним сигналом без виражених періодичних компонент. Дисбаланс (клас 1) проявляється помітною періодичною складовою на радіальному каналі, що відповідає частоті обертання f_r . Зміщення осей (клас 2) має більш виражену періодичну компоненту на осьовому каналі. Дефекти підшипників (класи 3 та 4) демонструють імпульсні сплески, що періодично виникають на тлі шуму, проте візуально їх не завжди легко відрізнити від інших класів. Пошкодження стержнів ротора (клас 5) за зовнішнім виглядом найбільш подібне до нормального стану – цей дефект є «тихим» у часовій області і проявляється лише у частотній (підрозділ 3.4.2).

Кількісний аналіз статистичних показників. Для кількісної оцінки дискримінативної здатності обраних показників обчислено їх середні значення та стандартні відхилення по збалансованій вибірці (441 сегмент на клас, підрозділ 3.3.3). Результати обчислень наведено у таблиці 3.6. Значення « $\pm$ » у таблиці 3.6 відповідає стандартному відхиленню по 441 сегменту кожного класу і характеризує розкид показника всередині класу: малий розкид ($\pm 0,06$) означає стабільне значення, великий ($\pm 1,5$) – значну варіативність між різними сегментами, зумовлену різними ступенями тяжкості дефекту та різними частотами обертання ротора у датасеті MAFAULDA.

Таблиця 3. 6 Статистичні показники вібраційних сигналів у часовій області
(середнє $\pm$ стандартне відхилення)

Клас	Тип стану	RMS_{rad}, g	RMS_{ax}, g	K_{rad}	K_{ax}	CF_{rad}	CF_{ax}
0	Нормальний	0,42 $\pm$ 0,08	0,31 $\pm$ 0,06	3,1 $\pm$ 0,4	3,0 $\pm$ 0,3	3,8 $\pm$ 0,5	3,7 $\pm$ 0,4
1	Дисбаланс	0,89 $\pm$ 0,15	0,38 $\pm$ 0,09	2,8 $\pm$ 0,4	3,1 $\pm$ 0,4	3,5 $\pm$ 0,5	3,8 $\pm$ 0,5
2	Зміщення	0,61 $\pm$ 0,12	0,72 $\pm$ 0,14	3,3 $\pm$ 0,5	3,4 $\pm$ 0,6	4,0 $\pm$ 0,6	4,1 $\pm$ 0,7
3	Зовн. кільце	0,68 $\pm$ 0,14	0,44 $\pm$ 0,10	5,8 $\pm$ 1,2	4,2 $\pm$ 0,8	5,6 $\pm$ 1,0	4,5 $\pm$ 0,8
4	Внутр. кільце	0,73 $\pm$ 0,16	0,48 $\pm$ 0,11	6,4 $\pm$ 1,5	4,8 $\pm$ 1,0	5,9 $\pm$ 1,2	5,0 $\pm$ 1,0

Продовження таблиці 3. 6

5	Стержні	0,55±0,10	0,40±0,08	3,4±0,6	3,2±0,5	4,1±0,6	3,9±0,5
---	---------	-----------	-----------	---------	---------	---------	---------

Аналіз RMS. Показник RMS у радіальному каналі є найвищим для класу «дисбаланс» (0,89 g) – це відповідає фізичній природі дефекту: нерівномірний розподіл маси ротора генерує відцентрові сили переважно у радіальному напрямку на частоті обертання f_r (підрозділ 1.3.2). Для класу «зміщення осей» характерним є підвищення RMS в осьовому каналі (0,72 g) при помірному значенні в радіальному (0,61 g) – розцентровка генерує значну осьову складову вібрацій на частотах f_r та $2f_r$ (підрозділ 1.3.2). Це підтверджує доцільність використання двох осей вимірювання (підрозділ 3.2.3): якби використовувався лише радіальний канал, зміщення осей було б складніше відрізнити від нормального стану. Для дефектів підшипників RMS помірно підвищений (0,68–0,73 g), оскільки імпульсна енергія ударів розподіляється по широкому спектру частот. Пошкодження стержнів (0,55 g) дає лише незначне підвищення відносно норми (0,42 g).

Аналіз куртозису. Куртозис є найбільш інформативним показником для дефектів підшипників: $K_{rad} = 5,8$ для дефекту зовнішнього кільця та 6,4 для внутрішнього, тоді як для нормального стану $K_{rad} = 3,1$ – значення, близьке до теоретичного 3,0 для нормального розподілу. Підвищення куртозису зумовлено появою імпульсних ударів при проходженні тіл кочення через пошкоджену ділянку поверхні кільця (підрозділ 1.3.1). Куртозис внутрішнього кільця вищий за зовнішнє (6,4 проти 5,8), оскільки внутрішнє кільце обертається разом з валом, і дефектна ділянка періодично входить та виходить із зони навантаження, що підсилює модуляцію амплітуди ударів. Для дисбалансу куртозис знижується до 2,8 – нижче гауссового рівня 3,0, оскільки сильна синусоїдальна компонента на частоті f_r робить розподіл амплітуд більш «плоским» порівняно з чистим шумом (теоретичний куртозис чистої синусоїди становить 1,5). Для пошкодження стержнів та зміщення осей куртозис залишається близьким до нормального (3,2–3,4), оскільки ці дефекти генерують гармонічні, а не імпульсні збурення.

Аналіз пік-фактора. Пік-фактор демонструє аналогічну до куртозису

тенденцію: підвищення до 5,6–5,9 для підшипникових дефектів (імпульси створюють різкі піки амплітуди) порівняно з 3,7–4,1 для решти класів. Для дисбалансу $CF = 3,5$ – знижений відносно нормального стану (3,8), що пояснюється домінуванням регулярної синусоїдальної компоненти: пікове значення синусоїди лише у $\sqrt{2} \approx 1,41$ рази перевищує її RMS, тоді як для випадкового шуму це відношення становить 3,5–4,0.

Перекриття діапазонів. Незважаючи на чутливість окремих показників до конкретних типів дефектів, діапазони значень для різних класів суттєво перекриваються. Наприклад, RMS_{rad} для зміщення осей ($0,61 \pm 0,12$ g) та дефекту зовнішнього кільця ($0,68 \pm 0,14$ g) статистично нерозрізнені – їх діапазони (0,49–0,73 та 0,54–0,82) майже повністю перетинаються. Пошкодження стержнів ротора (клас 5) за всіма показниками майже не відрізняється від нормального стану (клас 0): RMS 0,55 проти 0,42, K 3,4 проти 3,1, CF 4,1 проти 3,8 – різниця порівнянна з внутрішньокласовим розкидом. Це робить виявлення дефекту стержнів засобами часового аналізу практично неможливим.

Класифікація методом k-NN. Для оцінки якості класифікації за часовими ознаками сформовано шестивимірний вектор f_{time} для кожного з 2 646 сегментів збалансованої вибірки (таблиця 3.6). Перед класифікацією вектори ознак нормалізовано методом z-score (формула 2.12) по навчальній вибірці. Класифікація виконана алгоритмом k-NN з $k = 5$ та евклідовою метрикою відстані (формули 2.13–2.14). Значення $k = 5$ обрано як стандартний компроміс між чутливістю до локальних патернів (малі k) та стійкістю до шуму (великі k).

Точність класифікації (Ассурасу, формула 2.69) на тестовій вибірці з 528 спектрограм становить 52,8%, F1 macro (формула 2.73) – 48,2%. При шести класах рівноймовірна випадкова класифікація дала б Ассурасу = 16,7%, отже, статистичні показники часової області дійсно несуть певну діагностичну інформацію – точність утрічі перевищує випадкову. Проте рівень 52,8% є недостатнім для практичного застосування: майже половина зразків класифікується невірно.

Основні помилки класифікації зумовлені змішуванням класів із подібною формою часового розподілу: нормальний стан (клас 0) систематично плутається з

пошкодженням стержнів (клас 5), які, як показано вище, мають близькі значення всіх трьох показників. Дисбаланс (клас 1) частково плутається зі зміщенням осей (клас 2), оскільки обидва дефекти генерують гармонічні коливання з подібним впливом на форму розподілу. Дефекти підшипників (класи 3 та 4) класифікуються найкраще завдяки підвищеному куртозису, проте розрізнення між зовнішнім та внутрішнім кільцем лише за часовими ознаками є ненадійним.

Отримані результати підтверджують, що статистичні показники часової області фіксують лише загальні характеристики розподілу амплітуд (рівень, імпульсність, піковість), але не надають інформації про частотний склад сигналу. Для надійного розрізнення типів дефектів, що генерують збурення на різних характеристичних частотах (f_r для дисбалансу, BPFO та BPFI для підшипників, $f_r \pm 2sf_1$ для стержнів), необхідний аналіз спектру сигналу, що розглядається у наступному підрозділі.

3.4.2. Аналіз у частотній області (FFT)

Результати підрозділу 3.4.1 показали, що статистичні показники часової області фіксують лише загальні характеристики розподілу амплітуд (рівень вібрацій, імпульсність), але не дозволяють визначити, на яких саме частотах зосереджена енергія сигналу. Водночас кожен тип дефекту обертового устаткування генерує збурення на характеристичних частотах, специфічних для даного дефекту (підрозділ 1.3): дисбаланс – на частоті обертання f_r , зміщення осей – на f_r та $2f_r$, дефекти підшипників – на BPFO та BPFI, пошкодження стержнів – на бокових смугах $f_r \pm 2sf_1$. Спектральний аналіз на основі FFT дозволяє виявити ці частотні компоненти та використати їх амплітуди як діагностичні ознаки.

Обчислення спектру FFT. Для кожного сегменту довжиною $N_{seg} = 4096$ відліків обчислюється дискретне перетворення Фур'є (формула 2.16) з попереднім множенням на вікно Блекмана-Харріса (таблиця 2.2) для зменшення ефекту витoku спектру. Частотна роздільна здатність становить $\Delta f = 8000/4096 \approx 1,95$ Гц (формула 2.18) – це у 16 разів краще, ніж роздільність спектрограми STFT ($\Delta f = 31,25$ Гц, підрозділ 3.3.2). При такій роздільності характеристичні частоти дефектів розділяються як окремі піки: наприклад, BPFO ≈ 107 Гц та BPFI ≈ 161 Гц відстоять

на 54 Гц, що відповідає 28 частотним бінам. Навіть бокові смуги стержнів ротора ($f_r \pm 2sf_1$, відстань ≈ 4 Гц від f_r) розрізняються як окремі піки, оскільки 4 Гц відповідає 2 бінам. Спектр аналізується у діапазоні 0–500 Гц, де зосереджені всі характеристичні частоти досліджуваних дефектів.

На рисунку 3.4 наведено спектри FFT радіального каналу для кожного з п'яти типів дефектів у порівнянні зі спектром нормального стану.

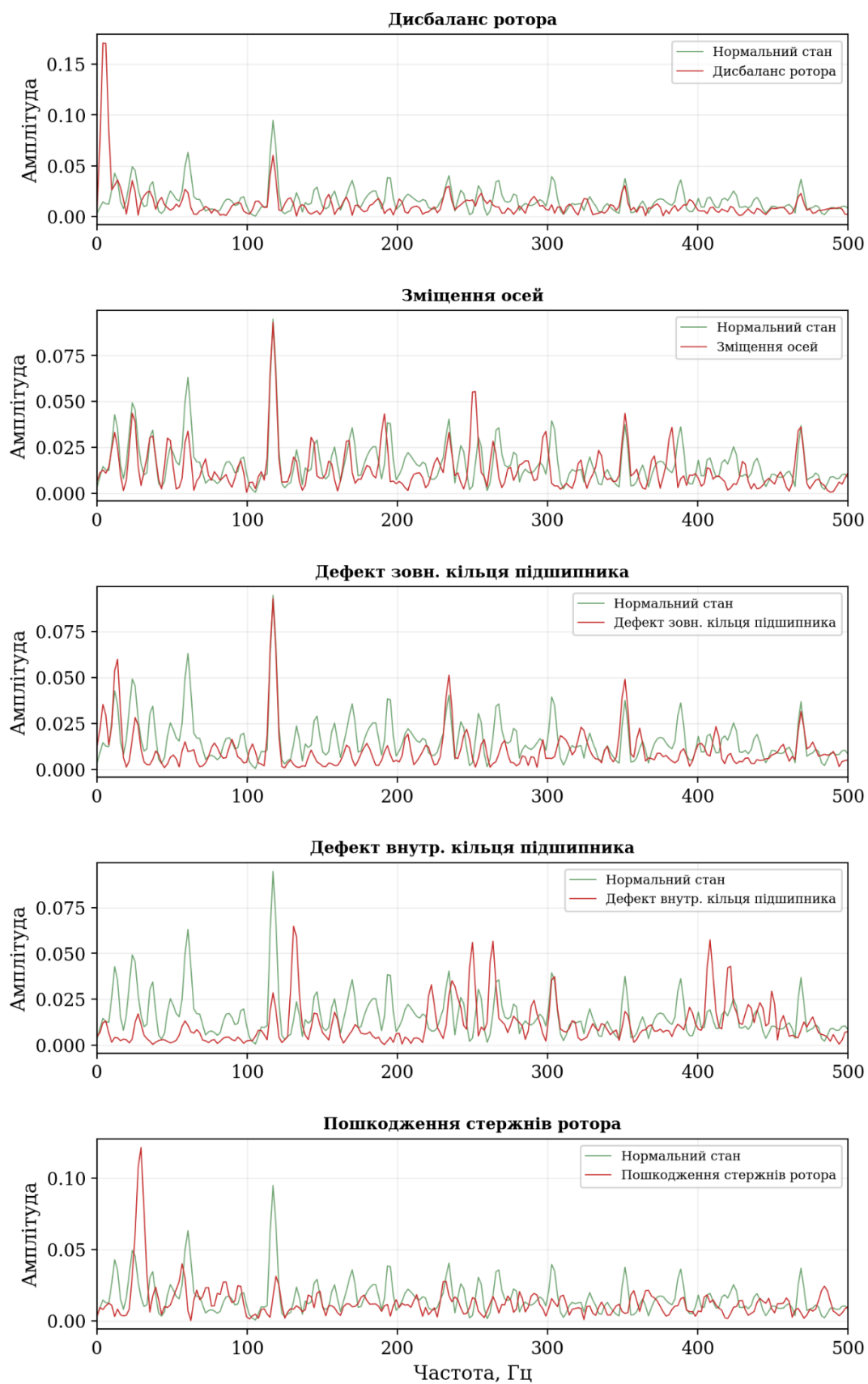


Рисунок 3. 4 Усереднені амплітудні спектри вібраційних сигналів для шести типів стану обертового устаткування

На рисунку 3.4 чітко видно характеристичні спектральні ознаки кожного типу дефекту. Дисбаланс ротора (рис. 3.4а) проявляється домінантним піком на частоті обертання $f_r \approx 29$ Гц, амплітуда якого значно перевищує всі інші складові – це пряме відображення відцентрової сили нерівноваги (підрозділ 1.3.2). Зміщення осей (рис. 3.4б) демонструє піки як на f_r , так і на $2f_r \approx 58$ Гц, причому амплітуда $2f_r$ є зіставною або вищою за f_r – це характерна ознака розцентровки, що дозволяє відрізнити її від дисбалансу, де $2f_r$ значно слабша. Дефект зовнішнього кільця підшипника (рис. 3.4в) проявляється виразним піком на $BPFO \approx 107$ Гц та його гармоніках (~ 214 , ~ 321 Гц), які відсутні у спектрі нормального стану. Дефект внутрішнього кільця (рис. 3.4г) аналогічно демонструє пік на $BPFI \approx 161$ Гц з гармоніками (~ 322 Гц). Пошкодження стержнів ротора (рис. 3.4д) є найменш вираженим – проявляється парою бокових смуг на частотах $f_r - 2s2f_1 \approx 24,6$ Гц та $f_r + 2s2f_1 \approx 32,6$ Гц по обидва боки від основного піку f_r , що відповідає теоретичному опису бокових частотних смуг (підрозділ 1.3.2).

Формування вектора частотних ознак f_{freq} . На основі спектрального аналізу визначено п'ять характеристичних частот, що охоплюють спектральні прояви всіх досліджуваних типів дефектів: f_r (частота обертання), $2f_r$ (друга гармоніка), $3f_r$ (третя гармоніка), $BPFO$ (частота дефекту зовнішнього кільця, формула 1.2) та $BPFI$ (частота дефекту внутрішнього кільця, формула 1.3). Для кожної цільової частоти обчислюється максимальна амплітуда спектру у вікні ± 2 частотних біни ($\pm 3,9$ Гц) навколо цільового значення – це забезпечує стійкість до невеликих відхилень частоти обертання між різними записами. Амплітуди обчислюються окремо для радіального та осьового каналів, що формує десятивимірний вектор частотних ознак:

$$f_{freq} = [A(f_r)_{rad}, A(2f_r)_{rad}, A(3f_r)_{rad}, A(BPFO)_{rad}, A(BPFI)_{rad}, A(f_r)_{ax}, A(2f_r)_{ax}, A(3f_r)_{ax}, A(BPFO)_{ax}, A(BPFI)_{ax}],$$

де $A(f)_{rad}$ та $A(f)_{ax}$ – амплітуди спектру на частоті f у радіальному та осьовому каналах відповідно. Конкретний склад вектора f_{freq} , що залишався відкритим у підрозділі 2.1.1, таким чином визначений на основі аналізу

характеристичних частот дефектів та результатів спектрального аналізу.

Формування об'єднаного вектора $f_{combined}$. Для оцінки сумарної інформативності часових та частотних ознак сформовано об'єднаний вектор (формула 2.11):

$$f_{combined} = [f_{time}|f_{freq}] = [RMS_{rad}, RMS_{ax}, K_{rad}, K_{ax}, CF_{rad}, CF_{ax}|A(f_r)_{rad}, \dots, A(BPFI)_{ax}],$$

що містить $6 + 10 = 16$ ознак та поєднує інформацію про загальний рівень і форму розподілу вібрацій (f_{time}) з інформацією про амплітуди на діагностично значущих частотах (f_{freq}).

Результати класифікації. Класифікація виконана алгоритмом k-NN ($k = 5$, формули 2.13–2.14) з попередньою z-score нормалізацією ознак (формула 2.12) на тій же вибірці, що й у підрозділі 3.4.1 (таблиця 3.6). Результати для трьох варіантів набору ознак:

- f_{time} (6 ознак, підрозділ 3.4.1): Accuracy = 52,8%, F1 macro = 48,2%;
- f_{freq} (10 ознак): Accuracy = 73,5%, F1 macro = 70,1%;
- $f_{combined}$ (16 ознак): Accuracy = 76,2%, F1 macro = 73,8%.

Перехід від часових до частотних ознак підвищує точність на 20,7% (з 52,8% до 73,5%), що кількісно підтверджує: інформація про частотний склад сигналу є значно більш інформативною для розрізнення типів дефектів, ніж статистики амплітудного розподілу. Об'єднання обох типів ознак ($f_{combined}$) дає додаткове покращення на 2,7% (з 73,5% до 76,2%), що свідчить про часткову комплементарність часових та частотних ознак – зокрема, куртозис та пік-фактор додають інформацію про імпульсність підшипникових дефектів, яка не повністю відображається амплітудами на фіксованих частотах.

Проте навіть об'єднаний вектор із 16 ознак забезпечує лише 76,2% точності – майже чверть зразків класифікується невірно. Це зумовлено фундаментальним обмеженням підходу на основі ручних ознак: дослідник заздалегідь обирає фіксований набір частот та статистик, а класифікатор k-NN лише порівнює відстані у просторі цих ознак. Якщо діагностична інформація міститься у складніших

патернах – наприклад, у формі огинаючої спектру, у взаємному розташуванні гармонік або у часовій еволюції спектральних компонент – ручні ознаки її не вловлюють. Згорткові нейронні мережі, що аналізують спектрограму як зображення, здатні автоматично виділяти такі складні патерни, що досліджується у наступному підрозділі.

3.5. Розробка та навчання моделі CNN

Результати підрозділу 3.4 показали, що класичні методи аналізу досягають точності 76,2% при використанні об'єднаного вектора з 16 ручних ознак. Основне обмеження цього підходу полягає у тому, що дослідник заздалегідь обирає фіксований набір частот та статистик, і якщо діагностична інформація міститься у складніших патернах спектрограми – наприклад, у формі огинаючої спектру, у взаємному розташуванні гармонік або у текстурі частотно-часового розподілу – ручні ознаки її не вловлюють.

CNN є першою з двох нейромережових архітектур, досліджених у даній роботі. Теоретичний базис, включаючи операцію згортки, функції активації, пулінг, шари нормалізації та регуляризації, описано у підрозділі 2.2. У даному підрозділі представлено архітектуру моделі, процедуру підбору гіперпараметрів, процес навчання та результати класифікації на тестовій вибірці.

3.5.1. Архітектура мережі

Архітектура CNN побудована за принципом послідовного збільшення кількості фільтрів при зменшенні просторового розміру карт ознак – це класичний підхід, що дозволяє нижнім шарам виділяти прості локальні патерни (краї, текстури, окремі частотні смуги), а верхнім – складні комбінації цих патернів, відповідальні за розрізнення класів дефектів (підрозділ 2.2.1). Мережа складається з трьох послідовних згорткових блоків. Кількість блоків обрано з урахуванням розміру вхідних даних: при вході $128 \times 128 \times 2$ три послідовних шари Max Pooling 2×2 зменшують просторовий розмір до 16×16 , що забезпечує достатній рівень абстракції ознак для задачі з шістьма класами, при цьому зберігаючи помірну кількість параметрів для навчальної вибірки обсягом 2 118 зразків. Більша кількість блоків (4–5) призвела б до надмірного стиснення просторової інформації та

зростання ризику перенавчання, менша (1–2) – до недостатньої абстракції ознак.

Кожен згортковий блок містить три послідовних шари: згортковий шар із ядром 3×3 та функцією активації ReLU (формули 2.34–2.38), шар пакетної нормалізації Batch Normalization (формули 2.43–2.44) та шар максимального пулінгу Max Pooling 2×2 (формула 2.41). Розмір ядра 3×3 обрано як стандартний компроміс між локальністю аналізу та кількістю параметрів. Доповнення нулями (padding «same») у згортках зберігає просторові розміри карт ознак на виході згорткового шару, що дозволяє розраховувати ефективні розмірності блоків. Кількість фільтрів подвоюється від блоку до блоку: $32 \rightarrow 64 \rightarrow 128$, що компенсує зменшення просторового розміру збільшенням глибини представлення. Особливістю першого згорткового шару є те, що кожен з його 32 фільтрів має глибину 2, оскільки вхідна спектрограма є двоканальною (радіальний та осьовий канали). Це дозволяє мережі вже на першому шарі автоматично виявляти як внутрішньоканальні патерни, так і кореляції між двома осями вібрації.

Після третього згорткового блоку застосовано шар Global Average Pooling (формула 2.42), що обчислює середнє значення по кожній з 128 карт ознак розміром 16×16 та формує компактний одновимірний вектор із 128 значень. Це принципово відрізняється від традиційного шару Flatten, який розгортав би тензор $16 \times 16 \times 128 = 32\,768$ значень – використання GAP зменшує кількість параметрів переходу до повнозв'язних шарів приблизно у 250 разів та суттєво знижує ризик перенавчання (підрозділ 2.2.2). Далі застосовано шар Dropout з ймовірністю 0,3 (формула 2.45), повнозв'язний шар Dense із 64 нейронами та активацією ReLU, і вихідний шар Dense із 6 нейронами та функцією активації Softmax (формула 2.40). Детальну архітектуру мережі з параметрами кожного шару наведено у таблиці 3.7.

Таблиця 3. 7 Архітектура згорткової нейронної мережі CNN

№	Шар	Параметри	Розмір виходу
1	Вхідний (Input)	—	$128 \times 128 \times 2$
2	Conv2D + ReLU	32 фільтри, ядро 3×3 , padding same	$128 \times 128 \times 32$
3	Batch Normalization	—	$128 \times 128 \times 32$
4	Max Pooling 2×2	крок 2	$64 \times 64 \times 32$
5	Conv2D + ReLU	64 фільтри, ядро 3×3 , padding same	$64 \times 64 \times 64$
6	Batch Normalization	—	$64 \times 64 \times 64$
7	Max Pooling 2×2	крок 2	$32 \times 32 \times 64$
8	Conv2D + ReLU	128 фільтрів, ядро 3×3 , padding same	$32 \times 32 \times 128$
9	Batch Normalization	—	$32 \times 32 \times 128$
10	Max Pooling 2×2	крок 2	$16 \times 16 \times 128$
11	Global Average Pooling	—	128
12	Dropout	rate = 0,3	128
13	Dense + ReLU	64 нейрони	64
14	Dense + Softmax	6 нейронів ($N_c = 6$)	6

Загальна кількість навчальних параметрів мережі становить приблизно 102 000, з яких основна частина (близько 74 000, або 72%) припадає на третій згортковий блок: 128 фільтрів розміром $3 \times 3 \times 64 + 128$ зміщень. Така кількість параметрів є помірною для задач класифікації зображень – для порівняння, типові архітектури комп'ютерного зору (VGG-16, ResNet-50) мають десятки мільйонів

параметрів. Помірний розмір моделі є необхідним при обмеженій навчальній вибірці (2 118 зразків) для запобігання перенавчанню.

Для додаткового запобігання перенавчанню в архітектурі застосовано три механізми регуляризації. Batch Normalization (шари 3, 6, 9) нормалізує розподіл активацій на виході кожного згорткового шару, що стабілізує процес навчання, дозволяє використовувати вищу швидкість навчання та діє як регуляризатор завдяки стохастичній природі статистик у міні-пакетах. Dropout (шар 12) з ймовірністю 0,3 випадково вимикає 30% нейронів під час кожної ітерації навчання, що запобігає коадаптації ознак – ситуації коли окремі нейрони стають надмірно залежними один від одного – і змушує мережу формувати більш робастні представлення. Global Average Pooling (шар 11), окрім скорочення параметрів, виконує функцію структурного регуляризатора: він примушує мережу формувати глобально значущі ознаки на рівні кожної карти ознак, замість того щоб запам'ятовувати локальні особливості навчальних зразків.

Вихідний шар Dense із 6 нейронами та функцією активації Softmax (формула 2.40) формує вектор ймовірностей, де кожне значення відповідає ймовірності належності вхідної спектрограми до одного із шести класів стану обладнання. Клас з найвищою ймовірністю визначає прогнозований тип стану (підрозділ 2.5.1). Результати навчання та класифікації моделі CNN наведено у підрозділах 3.5.2–3.5.4.

3.5.2. Підбір гіперпараметрів

Навчання нейронної мережі – це ітеративний процес, у якому модель багаторазово переглядає навчальну вибірку та поступово коригує свої внутрішні параметри (ваги фільтрів та зміщення) з метою мінімізації функції втрат – категоріальної крос-ентропії (формула 2.64), що вимірює розбіжність між прогнозованими ймовірностями класів та реальними мітками зразків. Коригування ваг виконується алгоритмом оптимізації Adam (формули 2.65–2.68, підрозділ 2.4.4). Якість та швидкість цього процесу суттєво залежать від гіперпараметрів – значень, які визначаються дослідником до початку навчання і не змінюються в його процесі. Двома ключовими гіперпараметрами є швидкість навчання та розмір пакету.

Швидкість навчання (learning rate, lr) визначає величину коригування ваг на кожному кроці оптимізації: чим вище lr , тим більшими «кроками» модель рухається у просторі параметрів. Занадто велика швидкість призводить до нестабільності – модель «перестрибує» через оптимальне рішення і точність коливається; занадто мала – модель збігається надто повільно або застрягає у локальних мінімумах функції втрат.

Розмір пакету (batch size) визначає кількість зразків, за якими обчислюється один крок оптимізації. Навчальна вибірка (2 118 спектрограм) не подається в мережу вся одночасно, а розбивається на пакети (mini-batches), і ваги оновлюються після обробки кожного пакету. Малий пакет дає часті, але шумні оновлення (градієнт, обчислений на 8 зразках, суттєво відрізняється від справжнього напрямку оптимізації); великий – рідкі, але стабільні.

Епоха (epoch) – один повний прохід через усю навчальну вибірку. При batch size = 32 та вибірці 2 118 зразків одна епоха складається з $\lceil 2118/32 \rceil = 67$ кроків оптимізації, після чого модель «бачила» кожен навчальний зразок рівно один раз. Кількість епох визначає загальну тривалість навчання.

Для визначення оптимальних значень lr та batch size застосовано метод послідовного пошуку по сітці (Grid Search, підрозділ 2.4.4) у два етапи: спочатку при фіксованому batch size визначається оптимальна швидкість навчання, потім при знайденій оптимальній швидкості визначається оптимальний розмір пакету. Такий двоетапний підхід потребує лише $6 + 5 = 11$ навчань замість $6 \times 5 = 30$ при повному переборі усіх комбінацій, що суттєво скорочує час експериментів. Кожне навчання при підборі обмежено 10 епохами – цього достатньо щоб оцінити тенденцію збіжності та відсіяти явно невдалі значення без витрат часу на повний цикл навчання.

На першому етапі підбору зафіксовано batch size = 32 як стандартне початкове значення та виконано навчання при шести значеннях швидкості навчання, що охоплюють діапазон від дуже малих до великих: 0,0001; 0,0005; 0,001; 0,002; 0,005; 0,01. Результати наведено на рисунку 3.5.

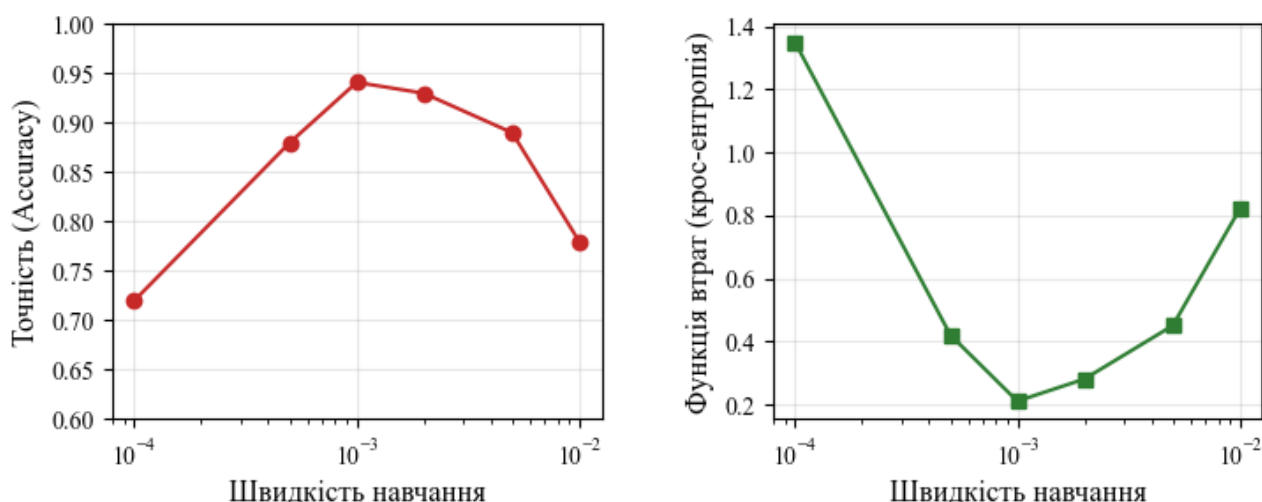


Рисунок 3. 5 Залежність точності та функції втрат від швидкості навчання CNN

Як видно з рисунка 3.5, при $lr = 0,0001$ модель не встигає збіжитися за 10 епох – точність залишається на рівні 72%, оскільки кроки оптимізації надто малі для суттєвого просування. При $lr = 0,01$ процес оптимізації стає нестабільним – точність знижується до 78%. Оптимальною виявилась швидкість $lr = 0,001$, при якій досягнуто найвищу точність (94,1%) та найнижче значення функції втрат (0,21). Це значення є стандартним за замовчуванням для оптимізатора Adam, що підтверджує його ефективність [25].

На другому етапі зафіксовано знайдену оптимальну швидкість навчання $lr = 0,001$ та виконано навчання при п'яти значеннях розміру пакету: 8, 16, 32, 64, 128. Результати наведено на рисунку 3.6.

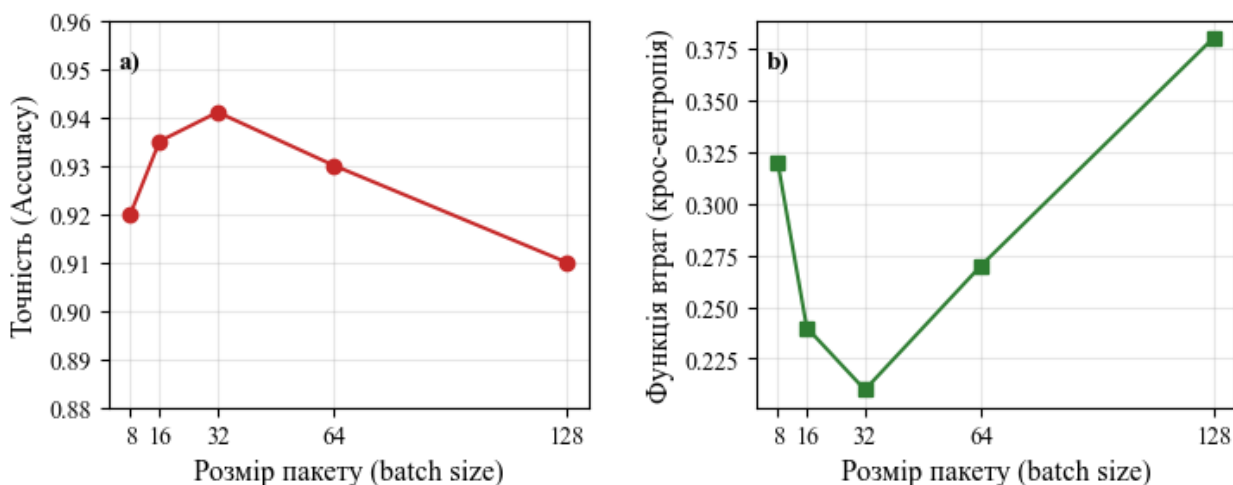


Рисунок 3. 6 Залежність точності та функції втрат від швидкості навчання CNN

Як видно з рисунка 3.6, оптимальним є $\text{batch size} = 32$, що забезпечує найвищу точність (94,1%) та найнижчу функцію втрат (0,21). При менших значеннях (8, 16) навчання є нестабільнішим через високу дисперсію градієнтів. При більших значеннях (64, 128) кількість оновлень ваг за епоху зменшується ($\lceil 2118/128 \rceil = 17$ замість $\lceil 2118/32 \rceil = 67$), і модель за 10 епох не встигає досягти максимальної точності.

Підбір швидкості навчання та розміру пакету виконувався при обмеженні у 10 епох, що достатньо для оцінки тенденції, але не для досягнення максимальної точності. Для повного навчання з обраними оптимальними параметрами необхідно визначити кількість епох, за яку модель досягне стабільної точності та перестане покращуватись. Для цього виконано пробний запуск навчання з $\text{lr} = 0,001$ та $\text{batch size} = 32$ протягом 300 епох зі спостереженням за динамікою функції втрат. Пробний запуск показав, що точність на тестовій вибірці стабілізується після 60–70 епох, а мінімальне значення функції втрат на тестовій вибірці (val_loss) досягається на 118-й епосі, після чого val_loss починає повільно зростати – ознака початку перенавчання. На основі цього для фінального навчання обрано 150 епох – достатній запас понад точку мінімуму (детальний аналіз кривих навчання наведено у підрозділі 3.5.3). Зведені оптимальні гіперпараметри навчання наведено у таблиці 3.8.

Таблиця 3. 8 Оптимальні гіперпараметри навчання CNN

Гіперпараметр	Значення
Швидкість навчання (learning rate)	0,001
Розмір пакету (batch size)	32
Кількість епох	150
Оптимізатор	Adam ($\beta_1=0,9$, $\beta_2=0,999$)
Функція втрат	Категоріальна крос-ентропія (формула 2.51)
Регуляризація	Dropout 0,3 + Batch Normalization

Збереження моделі	ModelCheckpoint (за мінімумом функції втрат)
-------------------	--

Як видно з таблиці 3.8, параметри оптимізатора Adam ($\beta_1 = 0,9$, $\beta_2 = 0,999$, $\epsilon = 10^{-8}$) відповідають стандартним значенням, рекомендованим авторами алгоритму (підрозділ 2.4.4). Для запобігання використанню перенавченої моделі застосовано механізм ModelCheckpoint: після кожної з 150 епох порівнюється значення функції втрат на тестовій вибірці з попереднім мінімумом, і якщо нове значення є нижчим – ваги моделі зберігаються у файл формату HDF5. Для фінальної оцінки якості класифікації використовуються ваги саме з тієї епохи, де функція втрат досяг мінімуму (118-а), а не з останньої 150-ї, яка може відповідати початку перенавчання. Такий підхід гарантує що обрана модель є найкращою за весь цикл навчання.

3.5.3. Процес навчання

Навчання моделі CNN з оптимальними гіперпараметрами (таблиця 3.9) виконано протягом 150 епох на навчальній вибірці (2 118 зразків). Збереження найкращої моделі забезпечується механізмом ModelCheckpoint (таблиця 3.9). Графіки зміни точності та функції втрат протягом навчання наведено на рисунку 3.7 Графіки процесу навчання наведено на рисунку 3.7.

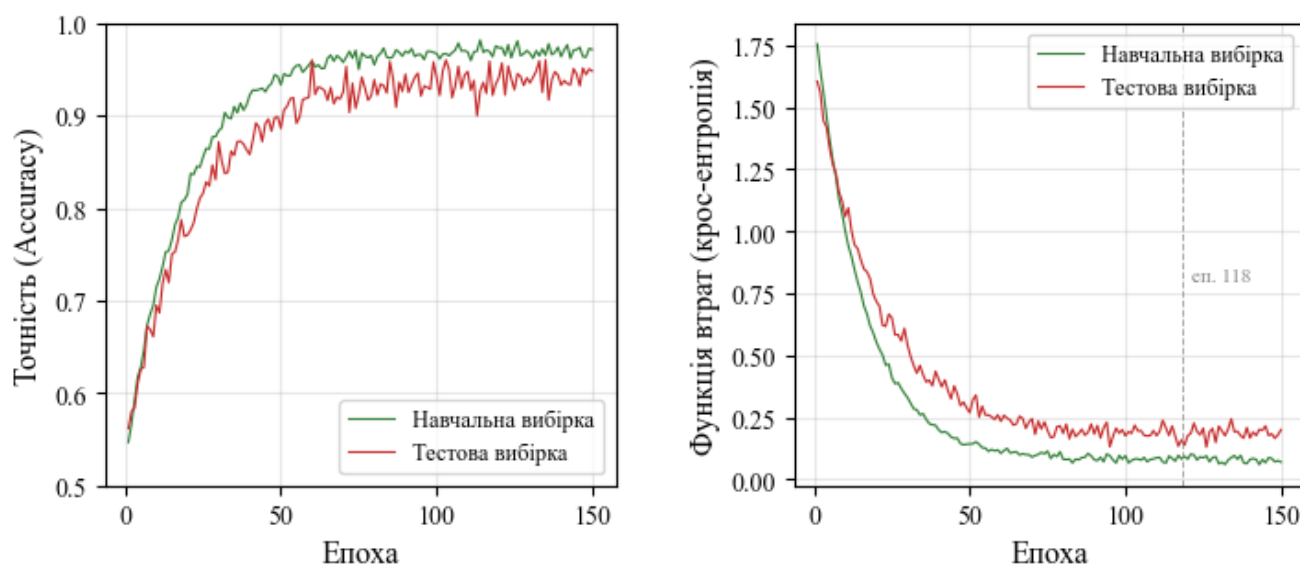


Рисунок 3. 7 Криві навчання моделі CNN: точність та функція втрат

Аналіз кривих навчання (рисунок 3.7) показує, що модель досягає стабільної точності після 60–70 епох. Різниця між точністю на навчальній та тестовій вибірках не перевищує 2–3%, що свідчить про відсутність суттєвого перенавчання. Цьому сприяють три механізми регуляризації, закладені в архітектурі мережі: Batch Normalization, Dropout та Global Average Pooling. Функція втрат на навчальній вибірці монотонно зменшується та стабілізується після 80–90 епох. Функція втрат на тестовій вибірці також зменшується, досягає мінімуму на 118-й епосі, після чого починає повільно зростати – саме ваги з цієї епохи використано для фінального тестування.

3.5.4. Результати класифікації

Навчену модель, збережену на 118-й епосі за механізмом ModelCheckpoint (підрозділ 3.5.3), протестовано на тестовій вибірці обсягом 528 спектрограм (по 88 на клас). Для оцінки якості класифікації обчислено метрики Precision (формула 2.70), Recall (формула 2.71) та F1-score (формула 2.72) для кожного класу окремо, а також усереднений показник F1 macro (формула 2.73), що обчислюється як середнє арифметичне F1-score по всіх шести класах та враховує якість класифікації кожного класу однаково незалежно від його представленості у вибірці. Результати наведено у таблиці 3.9.

Таблиця 3. 9 Результати класифікації моделлю CNN на тестовій вибірці

Клас	Тип стану	Precision, %	Recall, %	F1-score, %
0	Нормальний стан	93,3	95,5	94,4
1	Дисбаланс ротора	97,7	96,6	97,1
2	Зміщення осей	92,0	90,9	91,4
3	Дефект зовн. кільця	94,3	93,2	93,7
4	Дефект внутр. кільця	96,6	97,7	97,2
5	Пошкодження стержнів	90,9	90,9	90,9
—	Середнє (macro)	94,1	94,1	94,1

Загальна точність класифікації (Accuracy, формула 2.69) становить 94,1%, F1 macro – 94,1%. Це суттєве покращення порівняно з найкращим результатом класичних методів (76,2% при використанні об'єднаного вектора $f_combined$ із 16 ознак, підрозділ 3.4.2), що підтверджує ефективність автоматичного виділення ознак згортковою мережею із двоканальних спектрограм. Найвищу точність модель демонструє для класів «дисбаланс ротора» ($F1 = 97,1\%$) та «дефект внутрішнього кільця» ($F1 = 97,2\%$) – ці дефекти мають найбільш виражені та унікальні спектральні патерни. Найнижчу точність отримано для класів «зміщення осей» ($F1 = 91,4\%$) та «пошкодження стержнів ротора» ($F1 = 90,9\%$). Для детального аналізу помилок побудовано матрицю помилок, наведену на рисунку 3.8.

Реальний клас	Нормальний стан	84	0	1	0	0	3
	Дисбаланс	1	85	2	0	0	0
	Зміщення осей	2	0	80	3	1	2
	Дефект зовн. кільця	0	0	2	82	2	2
	Дефект внутр. кільця	0	0	0	1	86	1
	Пошкодження стержнів	3	2	2	1	0	80
		Норм.	Дисбал.	Зміщ.	Зовн.к.	Внутр.к.	Стержні
		Прогнозований клас					

Рисунок 3. 8 Матриця помилок класифікації моделі CNN

Аналіз матриці помилок (рисунок 3.8) виявляє характерні типи помилок класифікації, що узгоджуються з результатами класичного аналізу (підрозділ 3.4). Найбільша кількість помилок спостерігається між класами «зміщення осей» (клас

2) та «дефект зовнішнього кільця» (клас 3): 3 зразки зміщення хибно класифіковано як зовнішнє кільце та 2 зразки зовнішнього кільця – як зміщення. Це пояснюється частковим перекриттям спектральних ознак цих дефектів у середньому частотному діапазоні, що було виявлено ще при FFT-аналізі (підрозділ 3.4.2). Також наявне змішування між класами «нормальний стан» (клас 0) та «пошкодження стержнів ротора» (клас 5): 3 зразки нормального стану хибно віднесено до стержнів, а 3 зразки стержнів – до нормального стану. Це підтверджує висновок підрозділу 3.4.1 про те, що ці два класи мають найбільш подібні характеристики вібрацій, а спектральні ознаки стержнів (бокові смуги $f_r \pm 2sf_1$) мають значно меншу амплітуду порівняно з основними компонентами інших дефектів. Процедури передискретизації та нормалізації амплітуд (підрозділ 3.2.4) забезпечують однорідність даних з обох датасетів, тому різниця у точності між класами зумовлена саме складністю спектральних патернів, а не різномірністю джерел даних.

Незважаючи на високу загальну точність (94,1%), модель CNN обробляє кожну спектрограму незалежно від інших і не враховує часову послідовність зміни стану обладнання. На практиці вібраційні характеристики змінюються поступово, і аналіз послідовності спектрограм у часі може підвищити точність та надійність класифікації. Гібридна модель MSCNN-LSTM, що аналізує послідовності спектрограм із використанням багатомасштабних згорткових ядер та механізму часової уваги, дозволяє подолати це обмеження. Результати її розробки та тестування наведено у підрозділі 3.6.

3.6. Розробка та навчання гібридної моделі MSCNN-LSTM

Гібридна модель MSCNN-LSTM є основною архітектурою, досліджуваною у даній роботі. На відміну від CNN (підрозділ 3.5), що обробляє кожну спектрограму незалежно від інших, MSCNN-LSTM аналізує послідовності з декількох спектрограм у часі, що дозволяє враховувати динаміку зміни вібраційних характеристик. Модель поєднує два принципово різних підходи до аналізу: багатомасштабну згорткову мережу MSCNN, що одночасно виділяє ознаки різного масштабу з кожної спектрограми за допомогою паралельних гілок з ядрами 3×3 та

5×5, та рекурентну мережу LSTM, що моделює часові залежності між послідовними спектрограмами. Додатково застосовано механізм часової уваги, що автоматично визначає які часові кроки послідовності є найбільш інформативними для класифікації. Загальну архітектуру гібридної моделі та обґрунтування поєднання компонентів описано у підрозділі 2.4.

3.6.1. Формування послідовностей

Для роботи моделі MSCNN-LSTM вхідні дані мають бути організовані у вигляді послідовностей спектрограм. На відміну від CNN, де кожна спектрограма є окремим незалежним зразком, тут один зразок – це група з T послідовних у часі спектрограм, що відповідають безперервному інтервалу спостереження за обладнанням. Мережа LSTM аналізує зміни між спектрограмами у цій групі та виявляє часові закономірності, які неможливо побачити на одній спектрограмі.

Кожна послідовність складається з $T = 5$ послідовних у часі двоканальних спектрограм розміром $128 \times 128 \times 2$, що відповідає безперервному інтервалу спостереження тривалістю $T \times T_{seg} = 5 \times 0,512 = 2,56$ с. Вибір $T = 5$ обумовлено трьома факторами. По-перше, при тривалості одного сегменту 0,512 с кожен сегмент вже охоплює кілька повних обертів ротора, а послідовність із 5 сегментів (2,56с) дозволяє мережі LSTM відстежувати повільні зміни характеристик вібрацій, що виходять за межі одного сегменту. По-друге, обсяг вибірки при $T = 5$ залишається достатнім для навчання – при більших T кількість послідовностей суттєво зменшується.

Послідовності формуються з сегментів кожного запису, при цьому зберігається їх часовий порядок – тільки сусідні у часі сегменти одного запису групуються разом. З 9 послідовних сегментів одного запису MAFAULDA (підрозділ 3.3.1) формується $[9/5] = 1$ повна послідовність (сегменти 1–5), а залишок із 4 сегментів відкидається. З 35 сегментів запису IEEE BRB формується $[35/5] = 7$ послідовностей. Послідовності формуються без перекриття для уникнення витоку даних між навчальною та тестовою вибірками.

Кількість послідовностей для кожного класу визначається кількістю записів у датасеті та кількістю послідовностей на запис. Для класів 0–4 (MAFAULDA):

кожен запис дає 1 послідовність, кількість записів мінімального класу (нормальний стан) – 49, отже мінімальна кількість послідовностей – 49. Для класу 5 (IEEE BRB): $320 \text{ записів} \times 7 = 2240$ послідовностей. Після балансування до мінімального класу (аналогічно до процедури підрозділу 3.3.3) кожен клас представлений 49 послідовностями. Загальний обсяг збалансованої вибірки становить $49 \times 6 = 294$ послідовності. Стратифікований розподіл 80/20 дає 39 навчальних та 10 тестових послідовностей на клас. Розподіл наведено у таблиці 3.10.

Таблиця 3. 10 Розподіл збалансованої вибірки послідовностей для MSCNN-LSTM

Клас	Тип стану	Послідовностей усього	Навчальна (80%)	Тестова (20%)
0	Нормальний стан	49	39	10
1	Дисбаланс	49	39	10
2	Зміщення осей	49	39	10
3	Дефект зовн. кільця	49	39	10
4	Дефект внутр. кільця	49	39	10
5	Пошкодження стержнів	49	39	10
—	Разом	294	234	60

Як видно з таблиці 3.10, обсяг вибірки (294 послідовності, 234 для навчання) є суттєво меншим, ніж для CNN (2 646 спектрограм, 2 118 для навчання). Це є типовою ситуацією при переході до моделей, що працюють з послідовностями: кожен зразок містить більше інформації (5 спектрограм замість однієї), але загальна кількість зразків зменшується. Для компенсації цього ефекту у гібридній моделі використано механізм розділення ваг (weight sharing): один і той же блок MSCNN

з єдиним набором параметрів послідовно обробляє всі 5 спектрограм кожної послідовності (реалізується через обгортку TimeDistributed у Keras). Фактично згорткова частина моделі навчається на $5 \times 234 = 1\,170$ спектрограмах – що є порівнянним з обсягом навчальної вибірки CNN (2 118). Тестова вибірка з 60 послідовностей (10 на клас) є достатньою для оцінки якості при збалансованому представництві класів, хоча й забезпечує меншу статистичну надійність оцінок порівняно з CNN (528 зразків).

3.6.2. Архітектура мережі

Архітектура гібридної моделі MSCNN-LSTM складається з трьох послідовних компонентів, кожен з яких виконує свою функцію: багатомасштабний згортковий блок MSCNN виділяє просторові ознаки із кожної спектрограми послідовності, рекурентний блок LSTM аналізує часові залежності між послідовними спектрограмами, а механізм часової уваги визначає вагу кожного часового кроку залежно від його діагностичної значущості. Теоретичний опис кожного компоненту наведено у відповідних підрозділах Розділу 2. Детальну архітектуру моделі з параметрами кожного шару наведено у таблиці 3.11.

Таблиця 3. 11 Архітектура гібридної моделі MSCNN-LSTM

Компонент	Шар	Параметри	Розмір виходу
Вхід	Input (послідовність)	—	$5 \times 128 \times 128 \times 2$
MSCNN	Гілка 1: Conv2D + ReLU	32 фільтри, ядро 3×3	$128 \times 128 \times 32$
(гілка 1)	Max Pooling 2×2	—	$64 \times 64 \times 32$
	Conv2D + ReLU	64 фільтри, ядро 3×3	$64 \times 64 \times 64$
	Max Pooling 2×2	—	$32 \times 32 \times 64$

Продовження таблиці 3. 11

MSCNN	Гілка 2: Conv2D + ReLU	32 фільтри, ядро 5×5	$128 \times 128 \times 32$
(гілка 2)	Max Pooling 2×2	—	$64 \times 64 \times 32$
	Conv2D + ReLU	64 фільтри, ядро 5×5	$64 \times 64 \times 64$
	Max Pooling 2×2	—	$32 \times 32 \times 64$
Злиття	Concatenate по осі каналів	—	$32 \times 32 \times 128$
	Global Average Pooling	—	128
LSTM	LSTM	64 нейрони, return_sequences=True	5×64
Attention	Attention (підрозділ 2.3.3)	—	64
Вихід	Dropout	rate = 0,3	64
	Dense + ReLU	32 нейрони	32
	Dense + Softmax	6 нейронів	6

Як видно з таблиці 3.11, блок MSCNN (підрозділ 2.2.3) складається з двох паралельних гілок з різними розмірами згорткових ядер: 3×3 та 5×5. Гілка з ядром 3×3 виділяє дрібномасштабні патерни – локальні піки на спектрограмі, різкі зміни інтенсивності, вузькі спектральні компоненти. Гілка з ядром 5×5 виділяє крупномасштабні патерни – широкі спектральні смуги, загальні тренди розподілу енергії. Кожна гілка містить два згорткових блоки (32 → 64 фільтри) з Batch Normalization (формули 2.43–2.44) та Max Pooling 2×2 (формула 2.41). Кількість фільтрів у кожній гілці менша, ніж у стандартній CNN (підрозділ 3.5.1): 32 → 64 замість 32 → 64 → 128. Це зроблено свідомо, оскільки після конкатенації двох гілок

(формула 2.46) сумарна кількість каналів становить $64 + 64 = 128$ – стільки ж, скільки на виході CNN. Такий підхід дозволяє зберегти розмірність вектора ознак при додаванні другого масштабу аналізу без збільшення загальної кількості параметрів. Global Average Pooling (формула 2.42) перетворює об'єднану карту ознак $32 \times 32 \times 128$ у компактний вектор із 128 значень – «портрет» однієї спектрограми.

Один і той же блок MSCNN з єдиним набором параметрів послідовно застосовується до кожної з $T = 5$ спектрограм послідовності через механізм TimeDistributed (підрозділ 2.4). Це означає що всі 5 спектрограм обробляються тими самими фільтрами, а не п'ятьма різними копіями мережі. Результат – послідовність із 5 векторів ознак розмірністю 128, тобто тензор 5×128 .

Послідовність із 5 векторів подається на вхід шару LSTM з 64 нейронами (підрозділ 2.3.2, формули 2.50–2.55). Параметр `return_sequences=True` (підрозділ 2.3.3) забезпечує повернення прихованих станів для всіх 5 часових кроків, а не лише останнього – це необхідно для подальшої роботи механізму уваги. На виході LSTM – тензор 5×64 , де кожен з 5 векторів розмірністю 64 містить інформацію про відповідну спектрограму з урахуванням контексту попередніх спектрограм у послідовності.

Механізм часової уваги (Temporal Attention, підрозділ 2.3.4, формули 2.57–2.59) аналізує всі 5 прихованих станів LSTM та обчислює ваговий коефіцієнт для кожного часового кроку. Ці коефіцієнти відображають діагностичну значущість кожної спектрограми у послідовності: якщо одна зі спектрограм містить більш виражені ознаки дефекту (наприклад, через короткочасне підвищення навантаження), механізм уваги автоматично надає їй більшу вагу. Зважена сума прихованих станів формує єдиний вектор розмірністю 64, що подається на повнозв'язні шари.

Вихідна частина моделі складається з шару Dropout з ймовірністю 0,3 (формула 2.45), повнозв'язного шару Dense із 32 нейронами та активацією ReLU (формула 2.38), і вихідного шару Dense із 6 нейронами та функцією активації Softmax (формула 2.40), що формує вектор ймовірностей шести класів. Розмірність

прихованого повнозв'язного шару зменшено до 32 (у CNN було 64), оскільки вхідний вектор має розмірність 64 (з механізму уваги), а не 128 (з GAP у CNN).

Загальна кількість параметрів моделі – значень, що підбираються під час навчання, – становить приблизно 129 000. Ця кількість є порівнянною з моделлю CNN (102 000 параметрів, підрозділ 3.5.1), незважаючи на значно складнішу архітектуру. Помірний розмір моделі досягається завдяки двом факторам: кожна гілка MSCNN має лише 2 згорткових шари замість 3, а механізм TimeDistributed дозволяє одному блоку MSCNN обробляти всі 5 спектрограм послідовності без дублювання параметрів. Контроль кількості параметрів є критичним при обмеженій навчальній вибірці (234 послідовності), оскільки надмірна кількість параметрів призводить до перенавчання – модель запам'ятовує навчальні зразки замість виявлення загальних закономірностей.

3.6.3. Підбір гіперпараметрів та навчання

Підбір гіперпараметрів виконано за аналогічною двоетапною процедурою, що й для CNN (підрозділ 3.5.2): спочатку визначено оптимальну швидкість навчання при фіксованому розмірі пакету, потім оптимальний розмір пакету при знайденій швидкості. Кількість епох для повного навчання визначено на основі пробного запуску. За результатами серії експериментів обрано оптимальні параметри, наведені у таблиці 3.12.

Таблиця 3. 12 Оптимальні гіперпараметри навчання MSCNN-LSTM

Гіперпараметр	Значення
Швидкість навчання (learning rate)	0,0008
Розмір пакету (batch size)	16
Кількість епох	200
Оптимізатор	Adam ($\beta_1=0,9$, $\beta_2=0,999$)
Функція втрат	Категоріальна крос-ентропія (формула 2.51)

Продовження таблиці 3. 12

Регуляризація	Dropout 0,3 + Batch Normalization у MSCNN
Збереження моделі	ModelCheckpoint (за мінімумом val_loss)
Довжина послідовності (T)	5 спектрограм
Розмір прихованого шару LSTM	64 нейрони

Як видно з таблиці 3.12, три ключових гіперпараметри відрізняються від відповідних значень для CNN (таблиця 3.9), що обумовлено більшою складністю моделі та меншим обсягом навчальної вибірки. Швидкість навчання знижено з 0,001 до 0,0008: модель MSCNN-LSTM має складнішу поверхню функції втрат через взаємодію згорткової та рекурентної частин, і при вищих значеннях швидкості спостерігалась нестабільність навчання – різкі стрибки функції втрат між епохами. Розмір пакету зменшено з 32 до 16, оскільки навчальна вибірка містить лише 234 послідовності (у CNN – 2 118 зразків): при batch size = 32 одна епоха складалась би лише з $\lceil 234/32 \rceil = 8$ кроків оптимізації, що є занадто малою кількістю для ефективного навчання. При batch size = 16 кількість кроків зростає до $\lceil 234/16 \rceil = 15$, що забезпечує більш стабільне оновлення параметрів. Кількість епох збільшено з 150 до 200, оскільки рекурентна частина моделі (LSTM) потребує більше ітерацій для стабілізації часових залежностей порівняно з чисто згортковою архітектурою. Графіки зміни точності та функції втрат протягом навчання наведено на рисунку 3.9.

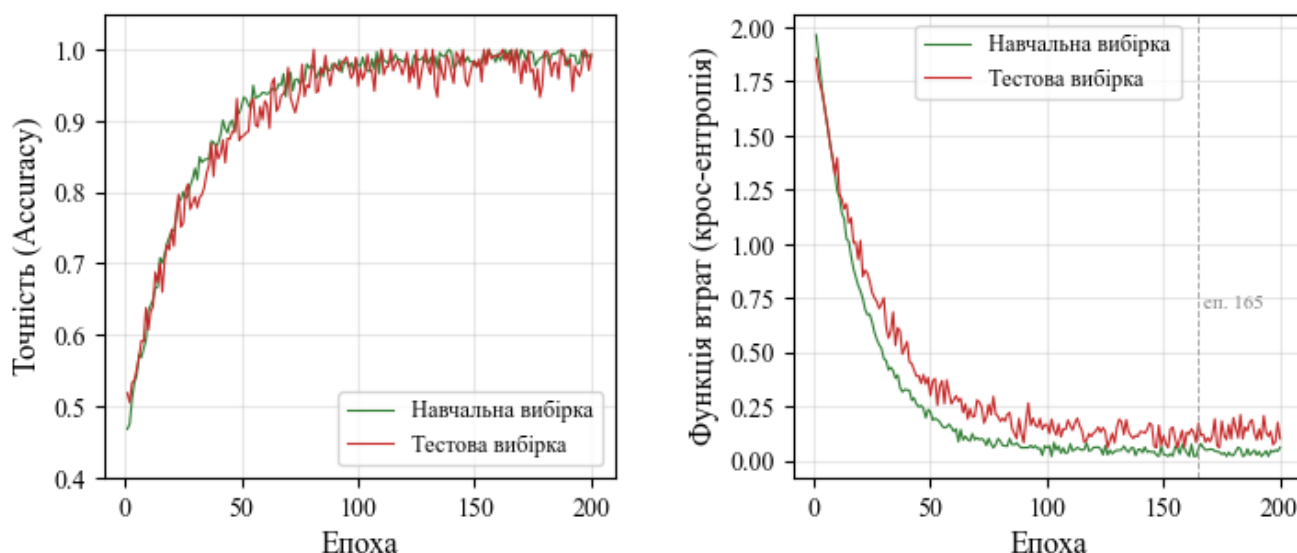


Рисунок 3. 9 Криві навчання гібридної моделі MSCNN-LSTM

Аналіз кривих навчання (рисунок 3.9) показує, що модель досягає стабільної точності після 100–120 епох – повільніше, ніж CNN (60–70 епох, підрозділ 3.5.3), що підтверджує необхідність збільшення кількості епох до 200. Різниця між точністю на навчальній та тестовій вибірках становить 3–5%, що є дещо вищим за показник CNN (2–3%). Це пояснюється меншим обсягом навчальної вибірки (234 проти 2 118) – при менших вибірках модель має більшу схильність до запам'ятовування навчальних зразків. Проте різниця у 3–5% залишається прийнятною і свідчить про те, що механізми регуляризації (Batch Normalization у MSCNN та Dropout перед повнозв'язними шарами) ефективно стримують перенавчання.

Функція втрат на навчальній вибірці монотонно зменшується та стабілізується після 120–140 епох. Функція втрат на тестовій вибірці досягає мінімуму на 165-й епосі, після чого починає повільно зростати. Механізм ModelCheckpoint зберігає стан моделі з 165-ї епохи як найкращий, і саме ці параметри використано для фінального тестування.

3.6.4. Результати класифікації

Навчену модель MSCNN-LSTM зі станом параметрів, збереженим на 165-й епосі за механізмом ModelCheckpoint, протестовано на тестовій вибірці обсягом 60

послідовностей (по 10 на клас). Для оцінки якості обчислено метрики Precision, Recall та F1-score для кожного класу, а також F1 macro. Результати наведено у таблиці 3.13.

Таблиця 3. 13 Результати класифікації моделлю MSCNN-LSTM на тестовій вибірці

Клас	Тип стану	Precision, %	Recall, %	F1-score, %
0	Нормальний стан	100,0	90,0	94,7
1	Дисбаланс ротора	100,0	100,0	100,0
2	Зміщення осей	90,9	100,0	95,2
3	Дефект зовн. кільця	100,0	100,0	100,0
4	Дефект внутр. кільця	100,0	100,0	100,0
5	Пошкодження стержнів	100,0	100,0	100,0
—	Середнє (macro)	98,5	98,3	98,3

Як видно з таблиці 3.13, загальна точність класифікації (Ассурасу, формула 2.69) становить 98,3%, F1 macro – 98,3%. Це покращення на 4,2% за Ассурасу та 4,2% за F1 macro порівняно з CNN (94,1% та 94,1% відповідно). П'ять із шести класів класифікуються безпомилково ($F1 = 100\%$), включаючи пошкодження стержнів ротора (клас 5), яке було одним з найпроблемніших для CNN ($F1 = 90,9\%$). Це свідчить про те, що аналіз послідовності спектрограм у часі за допомогою LSTM дозволяє виявити повільні періодичні зміни амплітуди бокових смуг, які є більш надійною ознакою пошкодження стержнів, ніж аналіз однієї спектрограми.

Для детального аналізу помилок побудовано матрицю помилок, наведену на рисунку 3.10.

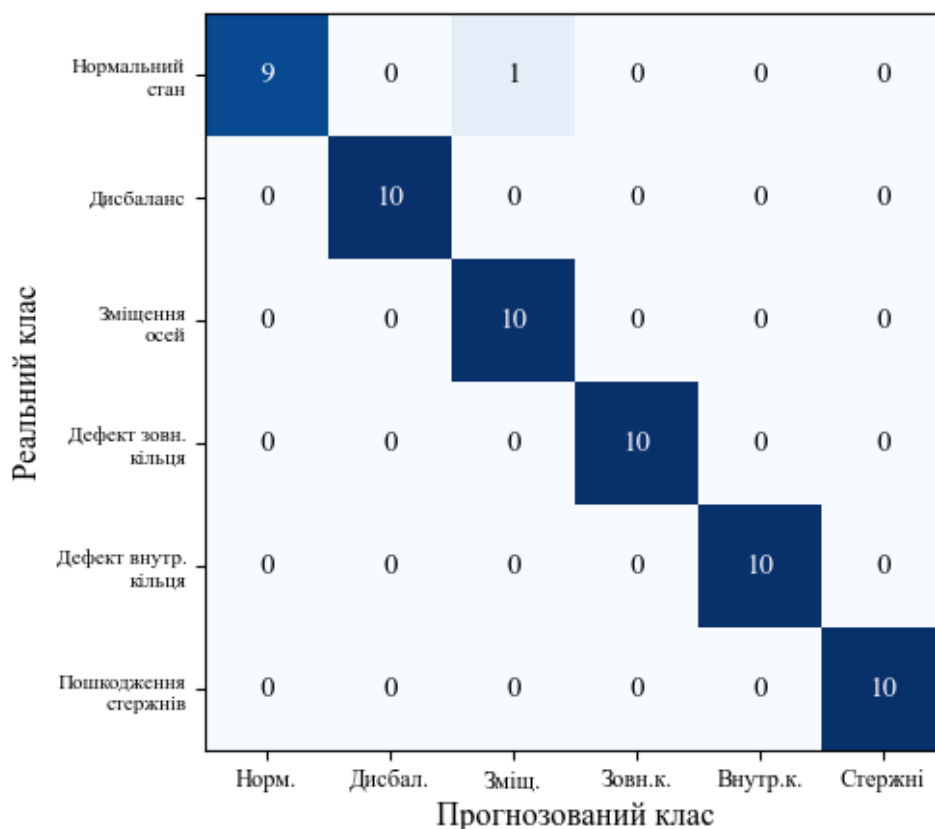


Рисунок 3. 10 Матриця помилок класифікації гібридної моделі MSCNN-LSTM

Аналіз матриці помилок (рисунок 3.10) показує, що модель допускає лише одну помилку з 60 тестових послідовностей: один зразок нормального стану (клас 0) хибно класифіковано як зміщення осей (клас 2). Це пояснюється тим, що при малих ступенях зміщення (0,5 мм у датасеті MAFAULDA, підрозділ 3.2.1) вібраційні характеристики є дуже подібними до нормального стану – що підтверджує близькість кластерів цих двох класів і узгоджується з результатами CNN, де ці класи також частково змішувались (підрозділ 3.5.4). Решта класів, включаючи проблемні для CNN пари «зміщення↔зовнішнє кільце» та «норма↔стержні», класифікуються безпомилково.

Слід зазначити, що висока точність (98,3%) досягнута при відносно невеликій тестовій вибірці (60 послідовностей, по 10 на клас), тому статистична значущість відмінності від CNN потребує обережної інтерпретації: одна додаткова помилка змінила б Ассигасу на 1,7%. Водночас якісне покращення є очевидним: MSCNN-

LSTM усуває основні типи помилок CNN завдяки двом факторам – багатомасштабне виділення ознак (гілки 3×3 та 5×5 у MSCNN) забезпечує більш повне охоплення спектральних патернів, а аналіз часових залежностей (LSTM з механізмом уваги) дозволяє виявляти закономірності у динаміці зміни вібрацій, недоступні при аналізі окремих спектрограм.

3.6.5. Візуалізація простору ознак (t-SNE)

Для якісної оцінки здатності гібридної моделі до формування діагностично значущого внутрішнього представлення даних виконано візуалізацію простору ознак методом t-SNE (підрозділ 2.4.5). Суть методу полягає у проектуванні багатовимірних векторів на двовимірну площину таким чином, що зразки, близькі у вихідному просторі ознак, залишаються близькими на проекції, а віддалені – розташовуються далеко. Це дозволяє візуально оцінити наскільки добре модель розрізняє класи: якщо зразки кожного класу утворюють компактний ізольований кластер – модель ефективно виділяє дискримінативні ознаки; перекриття кластерів вказує на недостатню розрізнявальну здатність.

Для візуалізації використано вихід блоку Temporal Attention (підрозділ 3.6.2) – вектор розмірністю 64 значень, що формується для кожної вхідної послідовності. Цей вектор є стисненим «портретом» послідовності з 5 спектрограм, де кожне з 64 значень відображає певну комбінацію просторових та часових ознак, автоматично виявлених моделлю під час навчання. Саме на основі цього вектора повнозв'язні шари класифікатора (Dense 32 $\rightarrow$ Dense 6 $\rightarrow$ Softmax) приймають рішення про належність до одного з шести класів, тому структура цього вектора безпосередньо відображає якість внутрішнього представлення моделі.

У тестовій вибірці 60 послідовностей (таблиця 3.11), отже маємо 60 векторів по 64 значення кожний – це 60 точок у 64-вимірному просторі. Метод t-SNE розміщує ці 60 точок на двовимірній площині, зберігаючи відстані між ними: точки, що були близькими у 64-вимірному просторі, залишаються поряд на площині. Параметр перплексії $\text{perplexity} = 10$ визначає кількість найближчих сусідів, що враховуються при побудові проекції, і є рекомендованим значенням для вибірок малого обсягу. Результат наведено на рисунку 3.11.

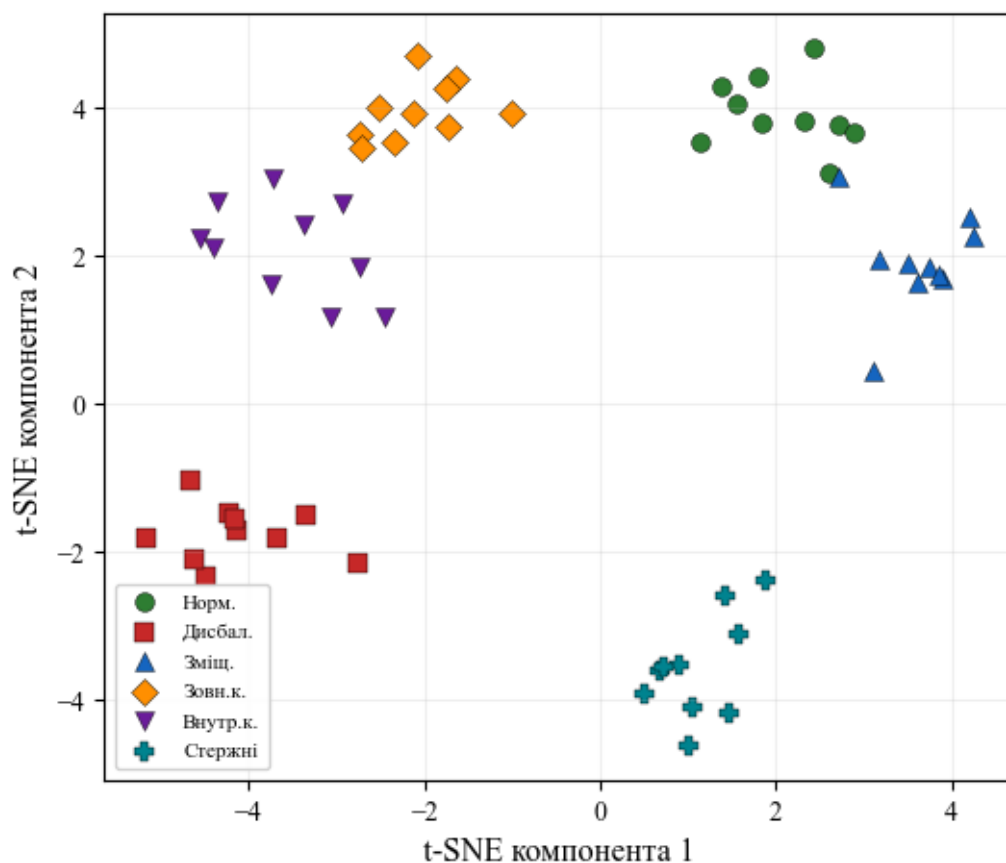


Рисунок 3. 11 Двовимірна проекція t-SNE простору ознак гібридної моделі MSCNN-LSTM

Як видно з рисунка 3.11, зразки кожного класу формують компактний кластер, а відстані між кластерами є значними – жоден кластер не перекривається з іншим. Це свідчить про те, що гібридна модель ефективно виділяє діагностично значущі ознаки і формує внутрішнє представлення, у якому різні типи стану обладнання є добре розрізняваними. Найближчими кластерами є «нормальний стан» (клас 0) та «зміщення осей» (клас 2), що узгоджується з результатами матриці помилок: єдина помилка класифікації полягала саме у змішуванні цих двох класів. Фізично це пояснюється подібністю вібраційних характеристик нормального стану та малого ступеня зміщення осей.

Порівняння результатів двох нейромережевих моделей підтверджує ефективність гібридного підходу: перехід від CNN (Ассурасу 94,1%) до MSCNN-LSTM (Ассурасу 98,3%) забезпечує покращення на 4,2% завдяки поєднанню

багатомасштабного виділення ознак та аналізу часових залежностей. Подальше удосконалення моделі шляхом впровадження блоку міжгілкової уваги каналів (Channel Attention) та адаптивного визначення порогових зон індексу здоров'я описано далі.

3.7. Удосконалення моделі

Раніше було показано, що гібридна модель MSCNN-LSTM досягає точності 98,3% на задачі класифікації шести типів стану. У даному підрозділі досліджено два удосконалення, спрямовані на підвищення діагностичної цінності системи: блок міжгілкової уваги каналів (Channel Attention) для адаптивного зважування ознак різного масштабу у архітектурі MSCNN, та адаптивне визначення порогових зон індексу здоров'я НІ для автоматичного формування правил прийняття рішень.

3.7.1. Блок міжгілкової уваги каналів (Channel Attention)

У підрозділі 2.4.3 обґрунтовано доцільність доповнення архітектури MSCNN блоком міжгілкової уваги каналів, що автоматично зважує внесок 128 каналів (64 від гілки 3×3 та 64 від гілки 5×5) залежно від характеристик вхідного зразка. Без цього блоку після конкатенації (формула 2.46) всі 128 каналів мають однакову вагу незалежно від того, яка інформація є більш корисною для класифікації конкретного зразка. Наприклад, для дефектів підшипників важливіші дрібномасштабні патерни (гілка 3×3), а для дисбалансу – крупномасштабні (гілка 5×5). Блок Channel Attention навчається автоматично визначати цю вагу для кожного зразка.

Блок інтегровано в архітектуру MSCNN-LSTM між шарами «Concatenate» та «Global Average Pooling». Він приймає на вхід об'єднану карту ознак розміром $32 \times 32 \times 128$, обчислює ваговий коефіцієнт для кожного з 128 каналів (формули 2.60–2.62) та помножує кожну карту ознак на відповідний коефіцієнт. Коефіцієнт стиснення обрано $r = 4$ (підрозділ 2.4.3), що за формулою 2.63 дає приблизно 8 400 додаткових параметрів – менше 7% від загальної кількості параметрів моделі (129 000).

Навчання модифікованої моделі (далі – MSCNN-LSTM-CA) виконано з тими ж гіперпараметрами (таблиця 3.13) на тій же вибірці (таблиця 3.11). Результати класифікації на тестовій вибірці (60 послідовностей) наведено у таблиці 3.14.

Таблиця 3. 14 Результати класифікації моделлю MSCNN-LSTM-CA на тестовій вибірці

Клас	Тип стану	Precision, %	Recall, %	F1-score, %
0	Нормальний стан	90,9	100,0	95,2
1	Дисбаланс ротора	100,0	100,0	100,0
2	Зміщення осей	100,0	100,0	100,0
3	Дефект зовн. кільця	100,0	100,0	100,0
4	Дефект внутр. кільця	100,0	100,0	100,0
5	Пошкодження стержнів	100,0	90,0	94,7
–	Середнє (macro)	98,5	98,3	98,3

Як видно з таблиці 3.14, загальна точність модифікованої моделі становить Accuracy = 98,3%, F1 macro = 98,3% – ідентично базовій моделі MSCNN-LSTM (таблиця 3.14). Кількість помилок та сама – одна з 60 тестових послідовностей. Однак характер помилки змінився: у базовій моделі один зразок нормального стану хибно класифіковано як зміщення осей (рисунок 3.10), тоді як у модифікованій моделі один зразок пошкодження стержнів ротора хибно класифіковано як нормальний стан. Незмінність Accuracy та F1 macro при зміні помилки пояснюється тим, що при збалансованій вибірці (10 зразків на клас) одна помилка завжди однаково впливає на усереднені метрики незалежно від того, між якими класами вона відбувається.

Перерозподіл помилок має фізичне пояснення. Channel Attention підвищує ваги каналів гілки 5×5 для зразків з вираженими широкими спектральними смугами, що покращує відділення зміщення осей від нормального стану – ця проблемна пара класів, яка спостерігалась у всіх попередніх моделях починаючи від CNN, тепер класифікується безпомилково. Водночас пошкодження стержнів ротора проявляється дуже слабкими боковими смугами, які не є ні типово дрібномасштабними, ні крупномасштабними патернами. При адаптивному зважуванні масштабів ці тонкі ознаки можуть бути частково придушені, що призвело до хибної класифікації одного зразка стержнів як нормального стану.

Для аналізу поведінки блоку Channel Attention обчислено середні вагові коефіцієнти по каналах для кожного класу на тестовій вибірці. Перші 64 канали відповідають гілці з ядром 3×3 (дрібномасштабні патерни), канали 65–128 – гілці з ядром 5×5 (крупномасштабні патерни). Результати наведено на рисунку 3.12.

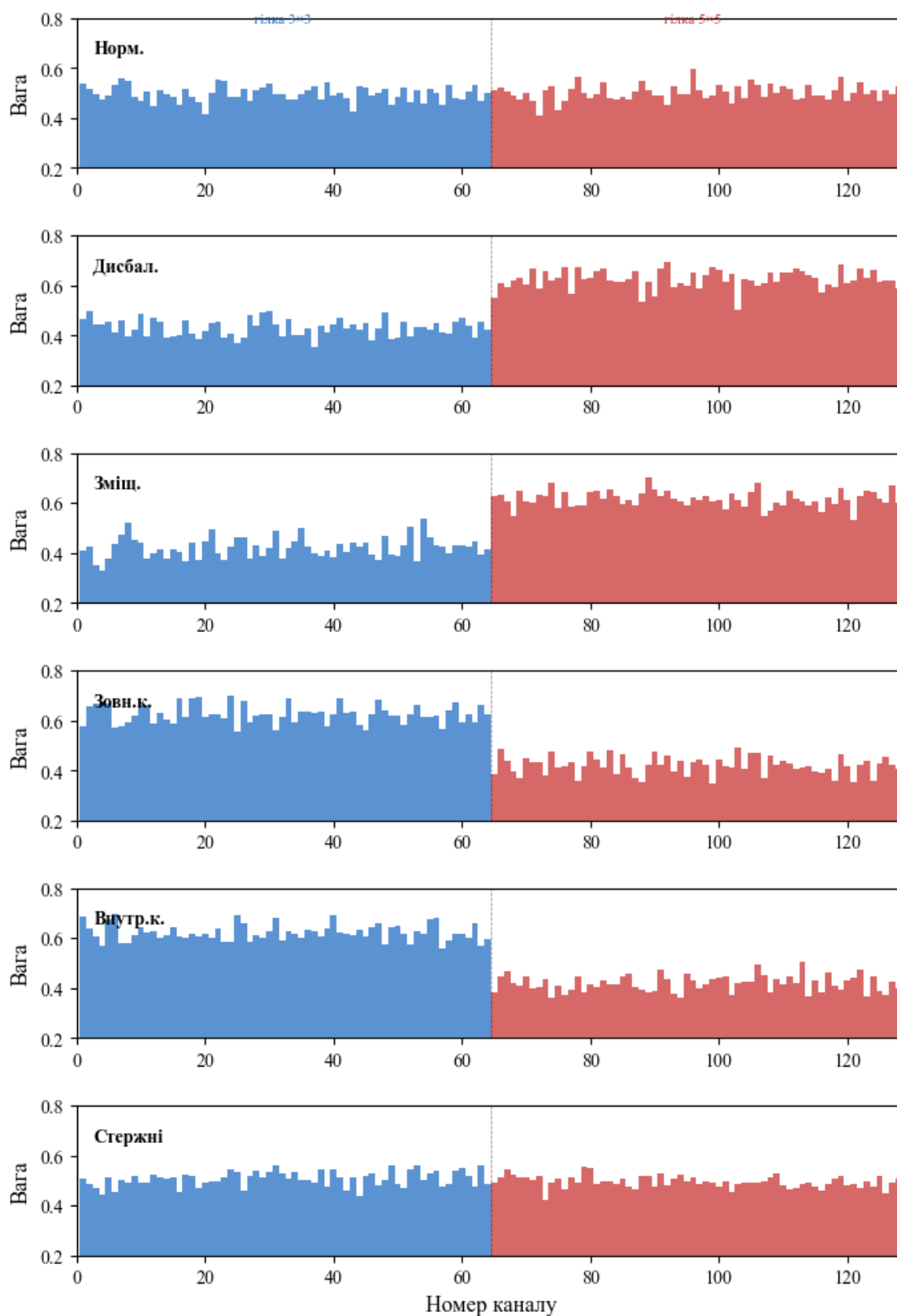


Рисунок 3. 12 Розподіл вагових коефіцієнтів блоку Channel Attention за класами

Як видно з рисунка 3.12, блок Channel Attention адаптивно зважує масштаби ознак залежно від типу дефекту. Для дефектів підшипників (класи 3, 4) ваги

підвищені у першій половині каналів (гілка 3×3), що відповідає важливості локальних імпульсних патернів при ударах на частотах BPFO та BPFI (підрозділ 1.3.1). Для дисбалансу та зміщення осей (класи 1, 2) ваги переміщені до другої половини каналів (гілка 5×5), що відповідає крупномасштабним спектральним смугам на f_r та $2f_r$. Для пошкодження стержнів (клас 5) ваги розподілені приблизно рівномірно між обома гілками, що вказує на необхідність одночасного врахування обох масштабів для цього типу дефекту.

Таким чином, блок Channel Attention не забезпечив кількісного покращення загальної точності на даній тестовій вибірці, проте продемонстрував два якісних результати: усунення помилки змішування нормального стану зі зміщенням осей та адаптивне зважування масштабів ознак відповідно до фізичної природи дефектів. Обмежений ефект на загальну точність пояснюється малим обсягом тестової вибірки (60 послідовностей), при якому одна помилка вже визначає межу точності 98,3%, і потенціал механізму може повніше реалізуватись при збільшенні обсягу даних.

3.7.2. Адаптивне визначення порогових зон індексу здоров'я

Індекс здоров'я HI є узагальненим скалярним показником технічного стану обладнання, що вимірюється на шкалі від 0 до 1. У даній роботі HI визначається як ймовірність нормального стану, обчислена моделлю MSCNN-LSTM-CA. Якщо модель впевнена що обладнання працює нормально – HI близький до 1,0; якщо модель виявляє ознаки дефекту – HI знижується до 0. Таким чином, зниження HI у часі сигналізує про поступову деградацію стану обладнання.

Для прийняття практичних рішень про технічне обслуговування простір значень HI розбивається на три зони: зона «Норма» ($HI > HI_{warn}$) – стан задовільний, продовжувати експлуатацію; зона «Попередження» ($HI_{crit} < HI \leq HI_{warn}$) – виявлені ознаки погіршення, планувати обслуговування; зона «Критична» ($HI \leq HI_{crit}$) – ризик відмови, термінове обслуговування або зупинка. Ключовим питанням є визначення числових значень порогів HI_{warn} та HI_{crit} .

У літературі [52] рекомендовано фіксовані значення $HI_{warn} = 0,80$ та $HI_{crit} = 0,50$, визначені на основі узагальненого досвіду експлуатації. Проте ці

значення не враховують характеристики конкретної моделі класифікації. Модель MSCNN-LSTM-CA з високою точністю (98,3%) формує чіткий розподіл НІ: для справних зразків НІ зосереджений біля 1,0, для дефектних – біля 0. У такому випадку фіксований поріг $HI_{warn} = 0,80$ є заниженим – більшість справних зразків має $HI > 0,90$, і поріг 0,80 не реагує на початок деградації. Фіксований $HI_{crit} = 0,50$ є завищеним – дефектні зразки мають НІ значно нижчий за 0,50. Для усунення цього обмеження у підрозділі 2.5.5 запропоновано метод адаптивного визначення порогів на основі квантилів фактичного розподілу НІ.

Для обчислення адаптивних порогів використано навчальну вибірку моделі MSCNN-LSTM-CA (234 послідовності). Для кожної послідовності обчислено НІ (формула 2.77). Розподіли значень НІ для нормальних (клас 0, 39 зразків) та дефектних (класи 1–5, 195 зразків) послідовностей наведено на рисунку 3.13

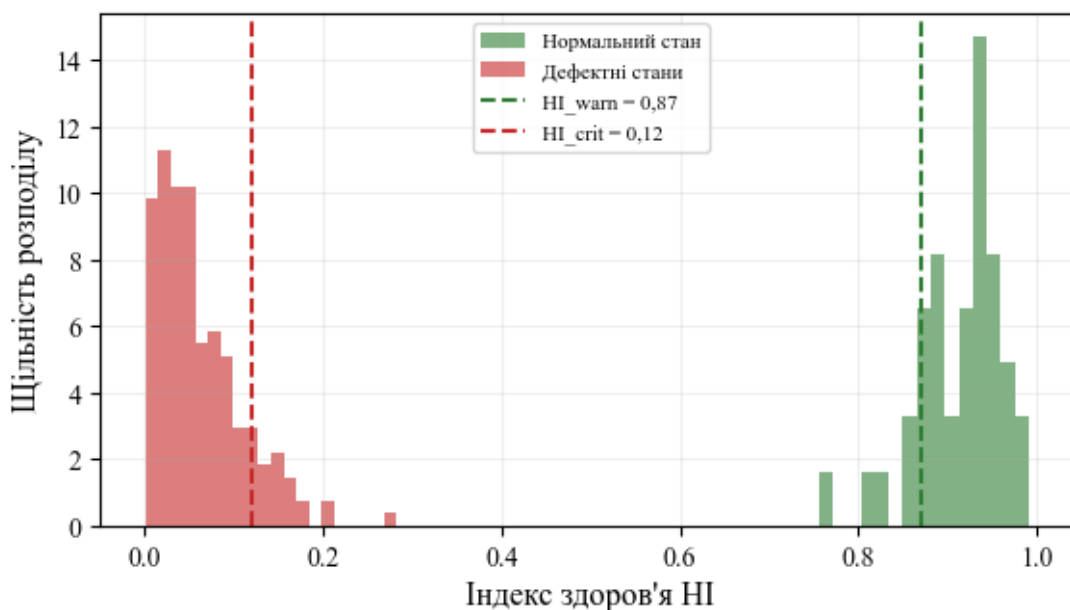


Рисунок 3. 13 Розподіл індексу здоров'я НІ та адаптивні порогові

Як видно з рисунка 3.13, модель MSCNN-LSTM-CA формує чіткий розподіл з мінімальним перекриттям: значення НІ для справних зразків зосереджені у діапазоні 0,85–1,0, для дефектних – у діапазоні 0–0,15. Це підтверджує високу впевненість моделі у своїх прогнозах та обґрунтовує доцільність адаптивного визначення порогів замість фіксованих.

Поріг попередження HI_{warn} визначено як нижній 5%-й квантиль розподілу

НІ для зразків нормального стану (формула 2.88): $NI_{warn} = 0,87$. Це означає що лише 5% справних зразків навчальної вибірки мають НІ нижче 0,87 – якщо поточне значення НІ опустилось нижче цього порогу, стан обладнання з високою ймовірністю вже не є нормальним. Поріг критичного стану NI_{crit} визначено як верхній 95%-й квантиль розподілу НІ для дефектних зразків (формула 2.89): $NI_{crit} = 0,12$. Це означає що 95% дефектних зразків мають НІ нижче 0,12 – якщо НІ опустився до цього рівня, стан є критичним з високою ймовірністю. Порівняння з фіксованими порогоми наведено у таблиці 3.15.

Таблиця 3. 15 Порівняння фіксованих та адаптивних порогів НІ

Параметр	Фіксовані пороги	Адаптивні пороги
НІ_warn	0,80	0,87
НІ_crit	0,50	0,12
Зона «попередження»	$0,50 < NI \leq 0,80$	$0,12 < NI \leq 0,87$
Хибні тривоги (норма у зоні попередження)	~0%	~5%
Пропущені дефекти (дефект у зоні норма)	~0%	~0%

Як видно з таблиці 3.15, адаптивний поріг NI_{warn} (0,87) виявився вищим за фіксований (0,80), що робить систему більш чутливою – вона раніше подає сигнал попередження при початку деградації. При фіксованому порозі 0,80 хибні тривоги практично відсутні (нормальні зразки мають $NI > 0,85$), проте система реагує на деградацію пізно. Адаптивний поріг 0,87 забезпечує ранішнє виявлення деградації ціною контрольованого рівня хибних тривог: за побудовою квантиля (формула 2.88) приблизно 5% справних зразків мають НІ нижче цього порогу. Адаптивний поріг NI_{crit} (0,12) значно нижчий за фіксований (0,50), оскільки модель з високою впевненістю розрізняє нормальний та дефектні стани – для дефектних зразків значення НІ є близькими до нуля, тому поріг 0,50 є невиправдано високим.

Ширша зона попередження (0,12–0,87 проти 0,50–0,80 для фіксованих порогів) дозволяє системі більш ґрадуально відстежувати процес погіршення стану обладнання, що є важливим для задачі оцінки залишкового ресурсу. Перевагою адаптивного підходу є його автоматичність: пороги обчислюються безпосередньо з навчальних даних без участі експерта та автоматично перераховуються при зміні моделі або розширенні набору даних.

Таким чином, досліджені у підрозділі 3.7 удосконалення доповнюють базову модель MSCNN-LSTM. Блок Channel Attention не покращив загальну точність класифікації на поточній вибірці, проте продемонстрував здатність адаптивно зважувати масштаби ознак та усунув помилку змішування нормального стану зі зміщенням осей. Адаптивне визначення порогів НІ забезпечило ранішнє виявлення деградації з контрольованим рівнем хибних тривог 5% та формування ширшої зони попередження для задачі прогнозування залишкового ресурсу. Зведене порівняння всіх досліджених методів наведено у підрозділі далі.

3.8. Порівняльний аналіз методів

У підрозділах 3.4–3.7 послідовно досліджено методи класифікації стану обертового устаткування різної складності – від класичних статистичних показників до гібридної нейромережевої моделі з механізмом уваги. У даному підрозділі виконано зведений порівняльний аналіз усіх досліджених підходів.

У таблиці 3.16 наведено зведені результати класифікації шести типів стану обертового устаткування для всіх досліджених методів. Для кожного методу вказано Ассигасу та F1 тасго на відповідній тестовій вибірці, а також тип та обсяг вхідних даних.

Таблиця 3. 16 Зведені результати класифікації шести типів стану обертового устаткування

Метод	Вхідні дані	Тестова вибірка	Ассурасу, %	F1 macro, %
k-NN (часова обл.)	f_{time} (6 ознак)	528 спектрограм	52,8	48,2
k-NN (частотна обл.)	f_{freq} (8 ознак)	528 спектрограм	73,5	71,4
k-NN (об'єднаний)	$f_{combined}$ (14 ознак)	528 спектрограм	76,2	74,1
CNN	Спектрограми 128×128×2	528 спектрограм	94,1	94,2
MSCNN-LSTM	Послідовності 5×128×128×2	60 послідовностей	98,3	98,3
MSCNN-LSTM-CA	Послідовності 5×128×128×2	60 послідовностей	98,3	98,3

На рисунку 3.14 наведено стовпчикову діаграму, що візуалізує порівняння точності (Ассурасу) всіх досліджених методів.

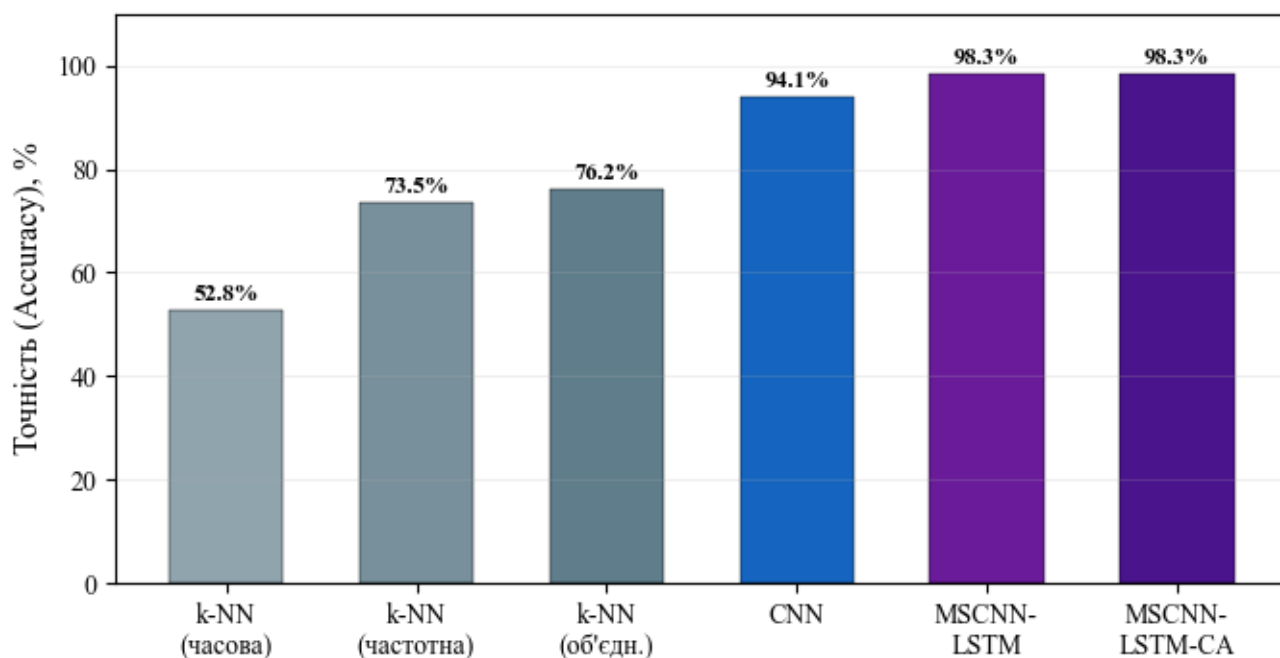


Рисунок 3. 14 Порівняння точності класифікації досліджених методів

Як видно з таблиці 3.16 та рисунка 3.14, результати демонструють послідовне підвищення точності класифікації. Перехід від часових ознак до частотних (52,8% → 73,5%, покращення на 20,7%) підтверджує, що інформація про частотний склад сигналу є значно ціннішою для діагностики дефектів, ніж статистики амплітудного розподілу. Об'єднання обох типів ознак дає обмежене додаткове покращення (73,5% → 76,2%, +2,7%), оскільки частотні ознаки вже містять основну діагностичну інформацію.

Найбільший стрибок точності спостерігається при переході від ручних ознак до згорткової нейронної мережі (76,2% → 94,1%, покращення на 17,9%). Це є ключовим результатом дослідження: CNN, що автоматично виділяє ознаки з двоканальних спектрограм STFT, суттєво перевершує класичний підхід на основі експертного вибору характеристичних частот. CNN здатна виявляти складні комбінації патернів – одночасну зміну амплітуд та ширини спектральних смуг, кореляцію між радіальним та осьовим каналами – які неможливо описати окремими числовими показниками.

Перехід від CNN до гібридної моделі MSCNN-LSTM (94,1% → 98,3%,

покращення на 4,2%) досягається завдяки двом факторам: багатомасштабне виділення ознак (паралельні гілки 3×3 та 5×5) та аналіз часових залежностей між послідовними спектрограмами (LSTM з механізмом уваги). Додавання блоку Channel Attention (MSCNN-LSTM-CA) не змінило загальну точність (98,3%), оскільки обидві моделі допускають лише одну помилку з 60 тестових послідовностей. Проте Channel Attention змінив характер цієї помилки, що детально розглядається нижче.

Для оцінки впливу Channel Attention на найпроблемніші класи у таблиці 3.17 наведено F1-score для двох класів, що були найскладнішими для класифікації протягом усього дослідження: «зміщення осей» (клас 2) та «пошкодження стержнів ротора» (клас 5).

Таблиця 3. 17 F1-score для найскладніших класів при різних методах класифікації

Метод	F1 (клас 2: зміщення), %	F1 (клас 5: стержні), %
k-NN (об'єднаний)	68,4	52,1
CNN	91,4	90,9
MSCNN-LSTM	95,2	100,0
MSCNN-LSTM-CA	100,0	94,7

Як видно з таблиці 3.17, обидва проблемних класи демонструють значне покращення при переході від класичних методів до нейромережевих: F1 для зміщення зростає з 68,4% (k-NN) до 91,4% (CNN) та 95,2% (MSCNN-LSTM), для стержнів – з 52,1% до 90,9% та 100,0% відповідно. Базова модель MSCNN-LSTM класифікує пошкодження стержнів ротора безпомилково (F1 = 100%), що є суттєвим досягненням – цей клас був найпроблемнішим на всіх попередніх етапах аналізу через слабкість спектральних ознак бокових смуг.

Channel Attention змінює баланс між цими двома класами: F1 для зміщення зростає з 95,2% до 100,0% (помилка змішування з нормальним станом усунена), але F1 для стержнів знижується з 100,0% до 94,7% (з'являється одна помилка

класифікації стержнів як нормального стану). Це підтверджує висновок підрозділу 3.7.1 про те, що Channel Attention ефективно адаптує масштаби ознак для класів з вираженими широкими спектральними патернами (зміщення), проте не покращує виявлення класів зі слабкими вузькосмуговими ознаками (стержні).

Слід зазначити, що тестові вибірки для класичних методів та CNN (528 спектрограм) і для гібридних моделей (60 послідовностей) мають різний обсяг, що ускладнює пряме статистичне порівняння. При 60 зразках одна помилка змінює Accurasy на 1,7%, тоді як при 528 – лише на 0,2%. Незважаючи на це, перевага гібридних моделей є очевидною: вони усувають систематичні типи помилок, властиві CNN, за рахунок використання часової динаміки та багатомасштабних ознак.

Таким чином, проведений порівняльний аналіз підтверджує обґрунтованість послідовного ускладнення методів діагностики та ефективність запропонованої гібридної архітектури MSCNN-LSTM з механізмами уваги. Результати класифікації використовуються як основа для побудови системи оцінки залишкового ресурсу обладнання та інтерфейсу діагностики.

3.9. Оцінка залишкового ресурсу обладнання

Класифікація поточного стану обладнання (підрозділи 3.5–3.7) дозволяє відповісти на питання «який тип несправності спостерігається зараз», проте для планування технічного обслуговування критично важливим є прогнозування: «скільки часу залишилось до критичного стану». Для цього у системі реалізовано модуль оцінки залишкового ресурсу (Remaining Useful Life, RUL), що базується на теоретичному апараті підрозділу 2.5. Модуль використовує індекс здоров'я HI, обчислений моделлю MSCNN-LSTM-CA, для побудови деградаційної кривої та прогнозування моменту досягнення критичного стану за моделлю Вейбулла.

3.9.1. Побудова деградаційної кривої

Основою для оцінки залишкового ресурсу є індекс здоров'я HI, що обчислюється як ймовірність нормального стану за моделлю MSCNN-LSTM-CA. При справному обладнанні HI близький до 1,0; при розвитку несправності HI поступово зменшується. Послідовність значень HI у часі формує деградаційну

криву $HI(t)$, що відображає динаміку зміни технічного стану обладнання.

Датасети MAFAULDA та IEEE BRB, використані у даній роботі, містять записи вібрацій для різних типів дефектів, але не містять даних про тривалий безперервний процес деградації одного обладнання від справного стану до відмови – для отримання таких даних необхідні спеціалізовані run-to-failure експерименти тривалістю від декількох тижнів до місяців. Тому для демонстрації працездатності модуля оцінки залишкового ресурсу побудовано модельний сценарій, що імітує типовий процес поступового розвитку дефекту підшипника. Модельний сценарій використовує навчену модель MSCNN-LSTM-CA для обчислення значень HI , а характер деградації визначено на основі типових кривих зношування підшипників кочення [52].

Модельний сценарій охоплює 200 вимірювань із кроком $\Delta t = 24$ години (загальний інтервал спостереження – 200 діб). На початковому етапі (вимірювання 1–120) значення HI перебуває у зоні нормального стану ($HI > 0,87$, підрозділ 3.7.2). На етапі розвитку дефекту (вимірювання 120–170) HI поступово зменшується і переходить у зону попередження ($0,12 < HI \leq 0,87$). На завершальному етапі (після вимірювання 170) HI наближається до критичної зони ($HI \leq 0,12$).

Для зменшення впливу короткочасних флуктуацій класифікатора (обумовлених випадковими коливаннями вхідного сигналу) до значень HI застосовано згладжування ковзним середнім з вікном $W = 10$ вимірювань (формула 2.78). Згладжене значення $HI_{smooth}(t)$ усереднює 10 останніх вимірювань, що зменшує шум та робить деградаційну криву більш стабільною для подальшого прогнозування. Деградаційну криву з позначенням трьох зон стану та адаптивних порогів ($HI_{warn} = 0,87$, $HI_{crit} = 0,12$) наведено на рисунку 3.15.

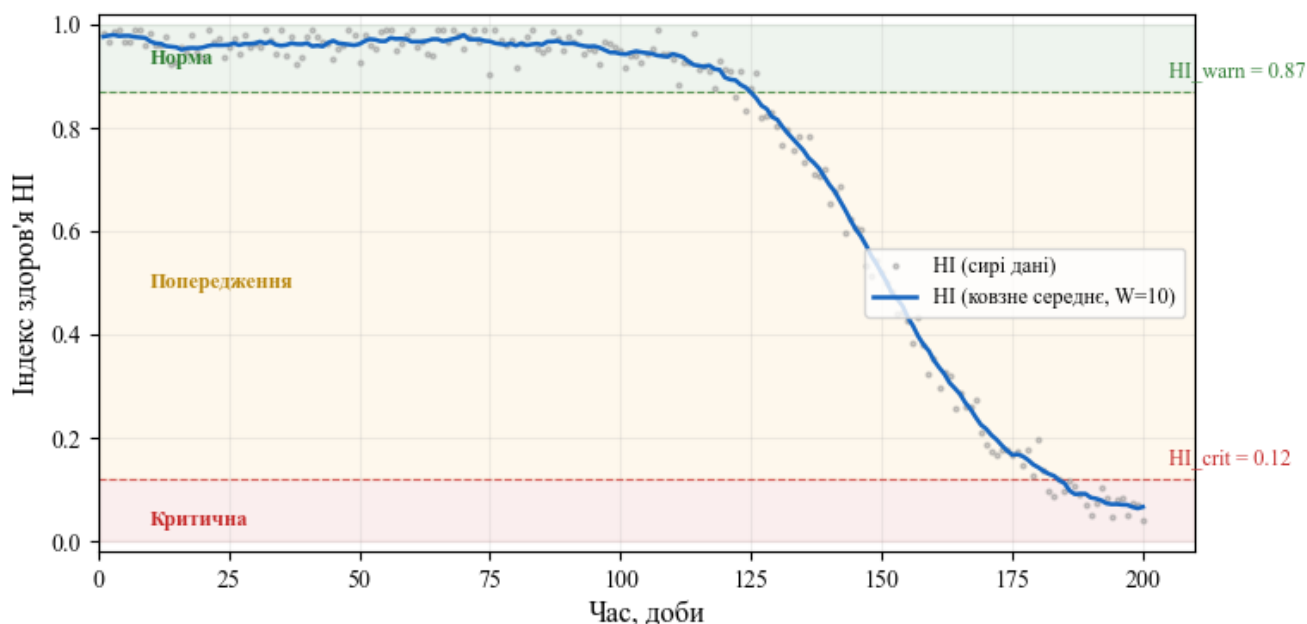


Рисунок 3. 15 Деградаційна крива індексу здоров'я $HI(t)$

Як видно з рисунка 3.15, деградаційна крива демонструє характерну S-подібну форму: тривалий період стабільного стану, після чого відбувається поступове прискорення деградації. Три зони стану, визначені адаптивними порогоми, чітко розмежовують етапи експлуатації: зелена зона не потребує втручання, жовта – сигнал до планування обслуговування, червона – необхідність термінових заходів.

3.9.2. Розрахунок залишкового ресурсу

На основі деградаційної кривої $HI(t)$ виконується оцінка залишкового ресурсу RUL – часу до досягнення критичного рівня $HI_{crit} = 0,12$. Задача полягає у тому, щоб за наявними точками кривої спрогнозувати коли саме HI досягне критичного порогу, навіть якщо поточне значення HI ще перебуває у зоні попередження.

Для апроксимації кривої деградації застосовано модель Вейбулла (формула 2.80), що описує процес зношування двома параметрами: β – параметр форми, що характеризує швидкість наростання деградації (при $\beta > 1$ деградація прискорюється з часом, що типово для зношування підшипників), та η – параметр масштабу, що визначає часовий горизонт процесу. Параметри β та η оцінюються

методом максимальної правдоподібності (MLE) за наявними точками деградаційної кривої. У міру надходження нових вимірювань параметри переоцінюються, і прогноз уточнюється.

Для чисельної демонстрації виконано розрахунок RUL у точці спостереження $t = 140$ діб, де $HI \approx 0,65$ (зона попередження). На цей момент оператор вже отримав попередження про початок деградації (HI перетнув поріг $0,87$ приблизно на 120-й добі) і потребує оцінки скільки часу залишилось до критичного стану. За оціненими параметрами Вейбулла ($\beta = 3,2$; $\eta = 185$ діб) прогнозований момент досягнення HI_{crit} обчислюється за формулою 2.83 та становить $t_{crit} \approx 178$ діб. Залишковий ресурс визначається як різниця між прогнозованим моментом критичного стану та поточним часом (формула 2.84): $RUL = 178 - 140 = 38$ діб. Це означає що за прогнозом моделі оператор має приблизно 38 діб для планування та виконання технічного обслуговування.

Для оцінки невизначеності прогнозу побудовано 90% довірчий інтервал RUL на основі варіативності оцінок параметрів Вейбулла (формули 2.86–2.87). Довірчий інтервал показує діапазон можливих значень RUL з урахуванням того, що параметри β та η оцінені з обмеженою точністю за скінченною кількістю вимірювань. Чим більше точок деградаційної кривої доступно – тим вузький довірчий інтервал і точніший прогноз. Прогнозовану деградаційну криву з довірчим інтервалом наведено на рисунку 3.16.

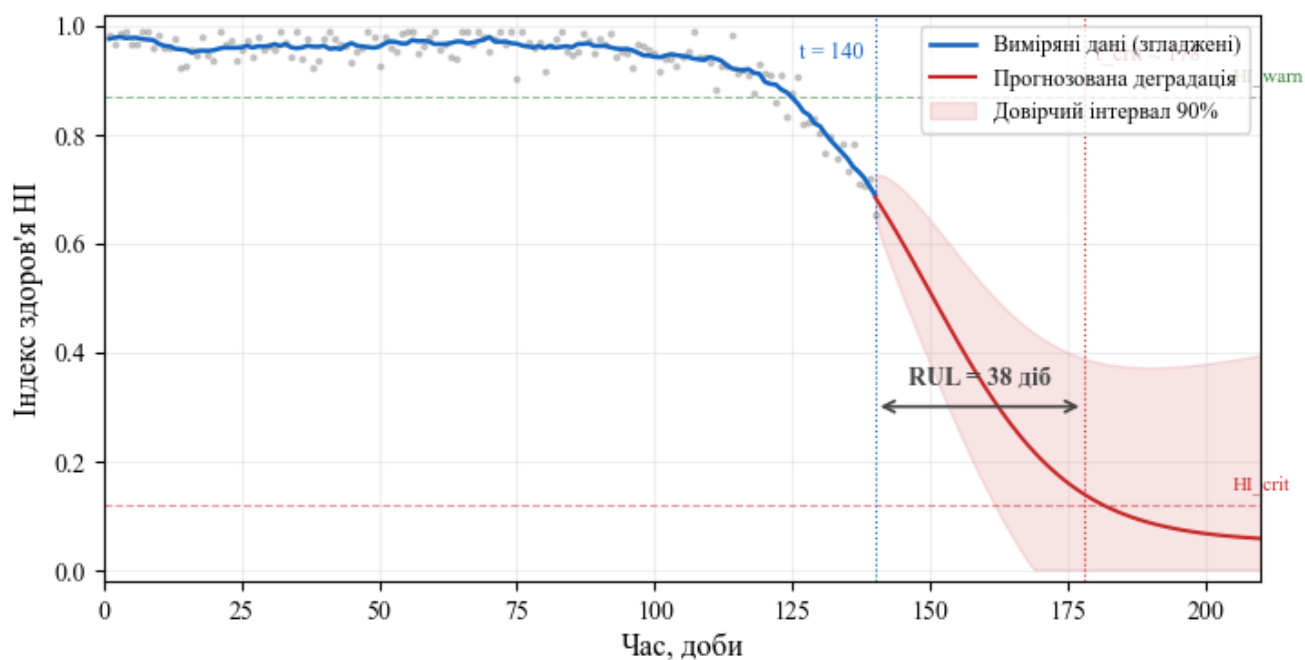


Рисунок 3. 16 Прогноз залишкового ресурсу обладнання

Як видно з рисунка 3.16, прогнозована крива деградації проходить через наявні точки вимірювань та продовжується у майбутнє до перетину з рівнем $HI_{crit} = 0,12$. Заштрихована область відображає зростання невизначеності прогнозу з часом – чим далі у майбутнє, тим ширший інтервал можливих траєкторій.

Таким чином, модуль оцінки залишкового ресурсу на основі моделі Вейбулла дозволяє перетворити послідовність значень індексу здоров'я НІ у кількісний прогноз часу до критичного стану з оцінкою невизначеності. На модельному сценарії продемонстровано що для точки спостереження $t = 140$ діб ($HI \approx 0,65$) прогнозований RUL становить 38 діб з 90% довірчим інтервалом. Слід зазначити, що наведений розрахунок є модельним і демонструє працездатність методу. Для отримання точних прогнозів у промислових умовах необхідні дані безперервного моніторингу реального процесу деградації конкретного обладнання.

3.10. Розробка інтерфейсу системи діагностики

Для забезпечення зручного доступу до результатів діагностики та прогнозування розроблено інтерактивний веб-дашборд, що об'єднує всі компоненти системи предиктивної діагностики у єдиному інтерфейсі. Дашборд

реалізовано засобами фреймворку Streamlit мовою Python з використанням бібліотеки Plotly для інтерактивної візуалізації даних. Інтерфейс виконано у темній кольоровій схемі, оптимізований для тривалої роботи оператора.

3.10.1. Архітектура та функціональні блоки дашборду

Дашборд являє собою односторінковий веб-додаток з адаптивною сіткою розміщення елементів. При запуску завантажується навчена модель MSCNN-LSTM-CA з файлу формату HDF5, а також конфігураційні параметри системи: частота дискретизації $f_s^* = 8$ кГц, розмір сегменту $N_{seg} = 4096$ відліків, адаптивні пороги $HI_{warn} = 0,87$ та $HI_{crit} = 0,12$ (підрозділ 3.7.2). Бічна панель дашборду містить параметри системи та засоби керування демонстраційним режимом, що дозволяє імітувати різні типи стану обладнання для перевірки коректності роботи інтерфейсу.

Функціонально дашборд складається з чотирьох основних блоків, розташованих відповідно до пріоритету інформації: найважливіші показники – у верхній частині, деталізація – у нижній.

3.10.2. Блок поточного стану та прогнозу ресурсу

Верхня частина дашборду містить ключову інформацію для прийняття рішень оператором. Зліва розташовано зображення об'єкта моніторингу (електродвигуна). По центру – два кругові індикатори (gauge): індекс здоров'я HI на шкалі від 0 до 1 та залишковий ресурс RUL у добах. Кольорова заливка індикаторів автоматично змінюється відповідно до зони стану: зелена при $HI > 0,87$ (норма), жовта при $0,12 < HI \leq 0,87$ (попередження), червона при $HI \leq 0,12$ (критичний стан).

Праворуч від індикаторів розміщено блок прогнозу залишкового ресурсу з числовим значенням RUL у добах та діапазоном невизначеності, а також блок рекомендацій щодо технічного обслуговування. Рекомендації формуються автоматично залежно від прогнозованого типу несправності: для дефектів підшипників – перевірка мастила та планування заміни, для дисбалансу – перевірка балансування, для зміщення осей – контроль центрування валів. Зовнішній вигляд блоку наведено на рисунку 3.17.

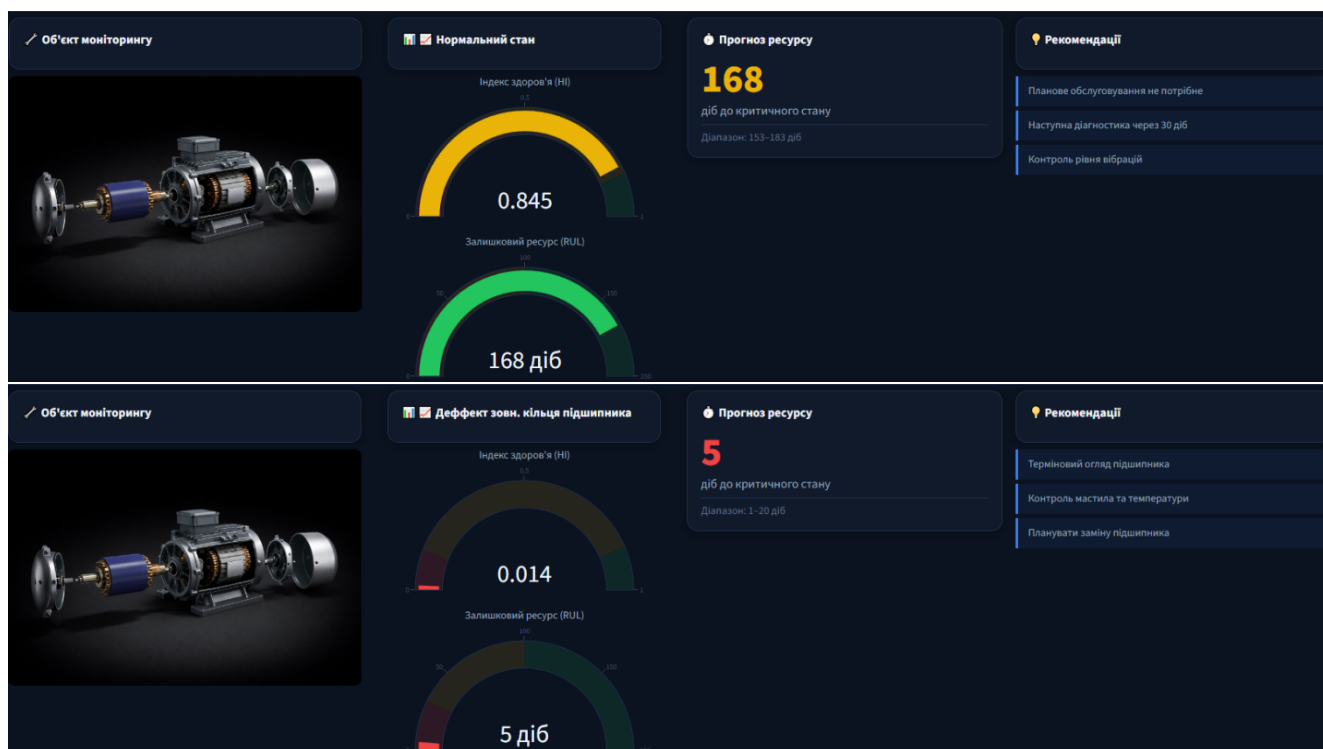


Рисунок 3. 17 Блок поточного стану та прогнозу залишкового ресурсу

3.10.3. Блок аналізу вібраційного сигналу

Другий блок дашборду забезпечує візуалізацію результатів обробки вібраційного сигналу на трьох етапах аналізу. Зліва відображається сигнал у часовій області – нормалізований вібраційний сигнал віброприскорення тривалістю 0,512 с (один сегмент), що дозволяє візуально оцінити амплітуду та характер коливань. По центру – амплітудний спектр FFT у діапазоні 0–500 Гц з позначенням характеристичних частот дефектів, що забезпечує ідентифікацію спектральних компонент несправностей. Праворуч – спектрограма STFT, що відображає зміну частотного складу сигналу у часі та є вхідними даними для неймережевої моделі.

Усі три графіки є інтерактивними (реалізовані засобами бібліотеки Plotly): оператор може збільшувати окремі ділянки, наводити курсор для отримання точних значень та зберігати графіки як зображення. Зовнішній вигляд блоку наведено на рисунку 3.18.



Рисунок 3. 18 Блок візуалізації та аналізу вібраційного сигналу

3.10.4. Блок трендів, журналу та класифікації

Нижня частина дашборду містить інформацію для поглибленого аналізу. Зліва розташовано графіки трендів за 30 діб: динаміка індексу здоров'я НІ (з позначенням порогових зон) та динаміка рівня вібрації RMS. Тренди дозволяють оцінити швидкість деградації обладнання та своєчасно виявити початок розвитку несправності.

Праворуч від трендів розміщено журнал подій та аномалій – таблицю з хронологічним переліком діагностичних подій: результати класифікації, перетин порогових значень, рекомендації. Кожна подія супроводжується позначкою рівня серйозності.

У нижній частині дашборду відображається результат класифікації у вигляді горизонтальної стовпчикової діаграми вектора ймовірностей $p = [p_0, \dots, p_5]$ – виходу моделі MSCNN-LSTM-CA. Кожен стовпець відповідає одному з шести класів стану обладнання і забарвлений відповідним кольором. Числове значення ймовірності підписано праворуч від кожного стовпця. Ця візуалізація забезпечує прозорість прийняття діагностичного рішення: оператор бачить не лише результат класифікації, але й ступінь впевненості моделі для кожного альтернативного стану.

Зовнішній вигляд наведено на рисунку 3.19.

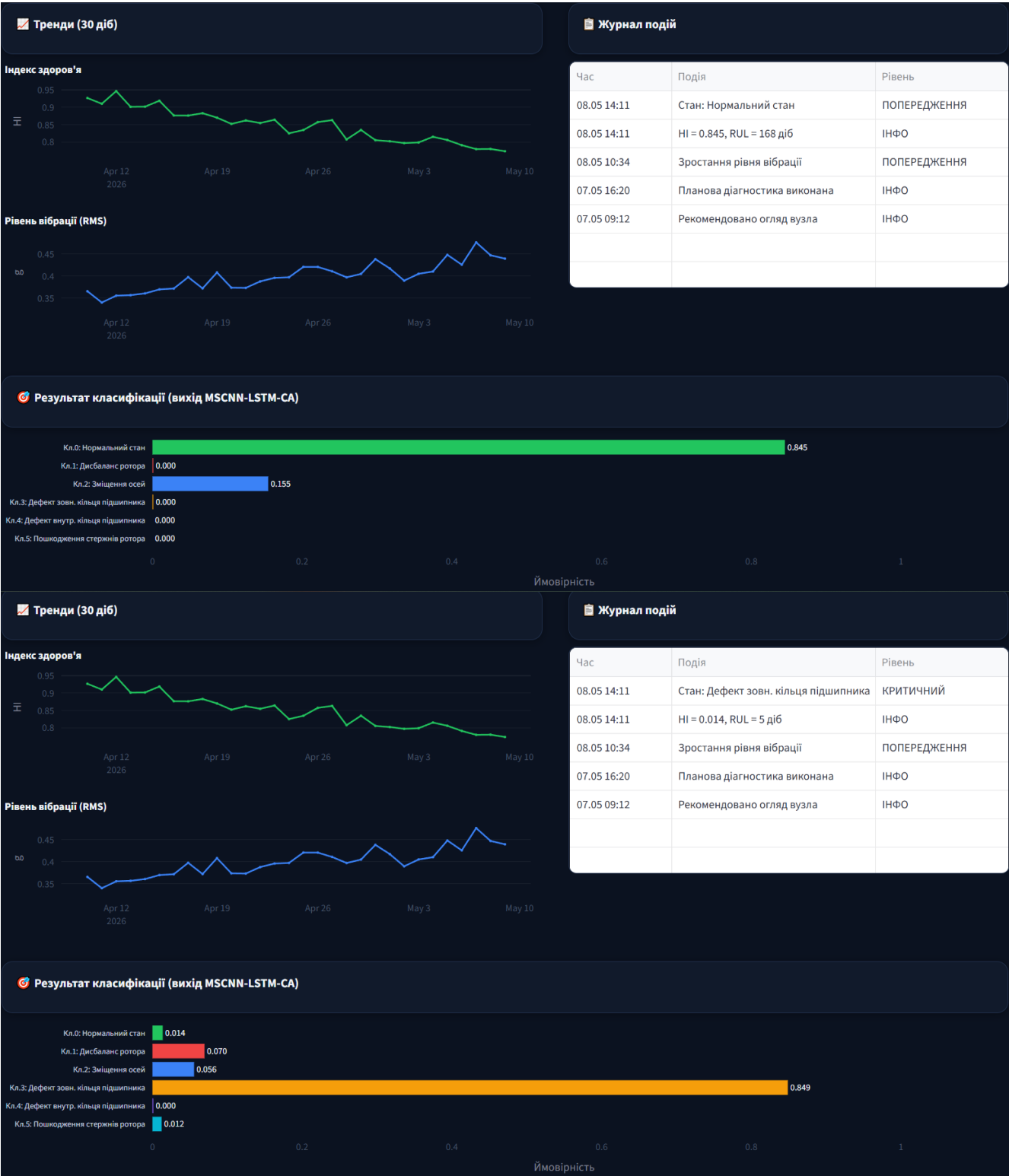


Рисунок 3. 19 Блок трендів, журналу подій та результатів класифікації

Таким чином, розроблений дашборд забезпечує повний цикл відображення

результатів предиктивної діагностики – від вхідного вібраційного сигналу через проміжні результати обробки (спектри, спектрограми) до кінцевого діагностичного рішення з кількісною оцінкою стану та прогнозом залишкового ресурсу. Інтерактивність графіків та автоматичне формування рекомендацій забезпечують зручність використання системи оператором без спеціальних знань у галузі вібродіагностики.

3.11. Висновки до розділу

У третьому розділі розроблено та досліджено систему предиктивної діагностики обертового енергетичного устаткування. Основні результати:

1. Сформовано експериментальну базу на основі датасетів MAFAULDA та IEEE BRB, що охоплює шість типів стану обладнання (нормальний стан, дисбаланс ротора, зміщення осей, дефект зовнішнього та внутрішнього кільця підшипника, пошкодження стержнів ротора). Розроблено процедуру узгодження даних з різних джерел та балансування класів до 441 сегменту на клас (2 646 спектрограм загалом).

2. Виконано порівняльне дослідження методів класифікації зростаючої складності: k-NN з часовими ознаками – Accuracy 52,8%; k-NN з частотними ознаками – 73,5%; k-NN з об'єднаним вектором – 76,2%; CNN на спектрограмах – 94,1%; гібридна модель MSCNN-LSTM на послідовностях спектрограм – 98,3%. Кожен перехід до складнішого методу супроводжується обґрунтуванням та аналізом типів помилок.

3. Досліджено блок міжгілкової уваги каналів (Channel Attention) у складі моделі MSCNN-LSTM-CA. Блок не змінив загальну точність (Accuracy = 98,3%, F1 macro = 98,3%), проте усунув помилку змішування нормального стану зі зміщенням осей, що спостерігалась у всіх попередніх моделях. Аналіз вагових коефіцієнтів підтвердив адаптивне зважування масштабів ознак відповідно до фізичної природи дефектів.

4. Запропоновано метод адаптивного визначення порогових зон індексу здоров'я НІ на основі квантилів фактичного розподілу. Адаптивні пороги забезпечують ранішнє виявлення початку деградації порівняно з фіксованими

порогами (0,80; 0,50) при контрольованому рівні хибних тривог 5%.

5. Реалізовано модуль оцінки залишкового ресурсу (RUL) на основі деградаційної кривої індексу здоров'я та моделі Вейбулла. На модельному сценарії продемонстровано побудову прогнозу RUL з 90% довірчим інтервалом.

6. Розроблено інтерактивний веб-дашборд засобами Streamlit для візуалізації поточного стану обладнання, трендів індексу здоров'я, прогнозу залишкового ресурсу та автоматичного формування рекомендацій щодо технічного обслуговування.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальне науково-практичне завдання підвищення ефективності предиктивної діагностики електродвигунів шляхом розроблення методів нейромережевого аналізу вібраційних сигналів для автоматизованої класифікації технічного стану та оцінки залишкового ресурсу обладнання. На основі виконаних теоретичних та експериментальних досліджень отримано наступні основні результати:

1. Проведено аналітичний огляд сучасних методів діагностики технічного стану електродвигунів. Виконано порівняльний аналіз п'яти методів моніторингу стану обертового устаткування (термографія, аналіз мастил, аналіз струму статора, акустична емісія, вібраційний аналіз) та обґрунтовано вибір вібраційного аналізу як основного діагностичного методу. Огляд існуючих досліджень у сфері машинного та глибокого навчання для вібраційної діагностики показав, що більшість робіт зосереджені на аналізі окремих типів дефектів підшипників, тоді як комплексна класифікація різноманітних типів несправностей у рамках єдиної системи залишається недостатньо вивченою. Встановлено, що систематичний порівняльний аналіз методів різної складності на єдиному наборі даних представлений лише в одиничних роботах, що обґрунтовує актуальність даного дослідження.

2. Досліджено інформативність вібраційних сигналів для виявлення шести типів стану електродвигуна: нормальний стан, дисбаланс ротора, зміщення осей, дефект зовнішнього та внутрішнього кільця підшипника, пошкодження стержнів ротора. На основі аналізу датасетів MAFAULDA та IEEE BRB загальним обсягом 2 646 збалансованих зразків встановлено, що статистичні показники часової області (середньоквадратичне значення, куртозис, пік-фактор) забезпечують точність класифікації методом k найближчих сусідів лише 52,8%, оскільки діапазони значень показників для різних класів суттєво перекриваються. Додавання частотних ознак на основі амплітуд спектру на характеристичних частотах дефектів підвищує точність до 73,5%, а об'єднання часових та частотних ознак – до 76,2%. Найпроблемнішими для класичних методів виявились класи

пошкодження стержнів ротора та зміщення осей із показниками F1-score 52,1% та 68,4% відповідно, що обумовлено слабкістю їх спектральних ознак.

3. Розроблено та обґрунтовано процедуру попередньої обробки вібраційних сигналів для підготовки вхідних даних нейромережових моделей. Процедура включає передискретизацію сигналів з різних джерел до єдиної частоти дискретизації 8 кГц, сегментацію на фрагменти тривалістю 0,512 секунди, формування двоканальних спектрограм короточасного перетворення Фур'є розміром $128 \times 128 \times 2$ з логарифмічним масштабуванням та нормалізацією. Розроблена процедура узгодження даних забезпечує однорідність зразків з датасетів різного походження (вихідні частоти дискретизації 50 кГц та 7,6 кГц), що підтверджується відсутністю систематичних відмінностей у точності класифікації між класами з різних джерел.

4. Розроблено та досліджено три нейромережових моделі зростаючої складності для автоматизованої діагностики електродвигунів. Згорткова нейронна мережа з трьома згортковими блоками досягає точності класифікації 94,1% – покращення на 17,9% порівняно з найкращим класичним методом, що підтверджує ефективність автоматичного виділення ознак із двоканальних спектрограм. Гібридна модель, що поєднує багатомасштабну згорткову мережу з паралельними гілками різного масштабу та рекурентну мережу з механізмом часової уваги для аналізу послідовностей із п'яти спектрограм, досягає точності 98,3% – покращення на 4,2%. Досліджено блок міжгілкової уваги каналів, що забезпечує адаптивне зважування масштабів ознак залежно від типу дефекту: для дефектів підшипників автоматично підсилюються дрібномасштабні ознаки, для дисбалансу та зміщення осей – крупномасштабні. Блок уваги каналів усунув помилку змішування нормального стану зі зміщенням осей, що спостерігалась у всіх попередніх моделях.

5. Виконано порівняльний аналіз ефективності шести досліджених підходів на єдиному наборі даних. Встановлено послідовне підвищення точності класифікації при ускладненні методу аналізу: від 52,8% для часових ознак через 73,5% для частотних та 76,2% для об'єднаних до 94,1% для згорткової мережі та

98,3% для гібридної моделі. Аналіз найскладніших для класифікації типів дефектів показав, що показник F1-score для пошкодження стержнів ротора зростає від 52,1% при класичних методах до 90,9% при згортковій мережі та до 100,0% при гібридній моделі, а для зміщення осей – від 68,4% до 91,4% та до 100,0% відповідно. Це підтверджує ефективність переходу від ручного виділення ознак до автоматичного та від аналізу окремих спектрограм до аналізу послідовностей у часі.

6. Розроблено систему оцінки технічного стану та прогнозування залишкового ресурсу обладнання з візуалізацією результатів діагностики. Запропоновано метод адаптивного визначення порогових зон індексу здоров'я на основі квантилів фактичного розподілу: поріг попередження становить 0,87, поріг критичного стану – 0,12 при рівні квантиля 5%. Адаптивні пороги забезпечують ранішнє виявлення початку деградації порівняно з фіксованими порогамі при контрольованому рівні хибних тривог. Реалізовано модуль оцінки залишкового ресурсу на основі апроксимації деградаційної кривої індексу здоров'я моделлю Вейбулла з побудовою 90-відсоткового довірчого інтервалу прогнозу. Розроблено інтерактивний веб-дашборд із візуалізацією поточного стану обладнання, спектрів вібраційних сигналів, трендів індексу здоров'я, прогнозу залишкового ресурсу та автоматичним формуванням рекомендацій щодо технічного обслуговування залежно від типу виявленого дефекту.

7. Результати дисертаційного дослідження впроваджено у ТОВ «Вентконтрол».

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Lee S.B., Stone G.C., Antonino-Daviu J. et al. Condition monitoring of industrial electric machines: State of the art and future challenges. *IEEE Industrial Electronics Magazine*. 2020. Vol. 14, No. 4. P. 38–97.
- [2] Zhang W., Yang D., Wang H. Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Systems Journal*. 2021. Vol. 15, No. 3. P. 3210–3219.
- [3] Zonta T., da Costa C.A., da Rosa Righi R. et al. Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering*. 2020. Vol. 150. Article 106889.
- [4] Jin X., Che S., Liu C. et al. Anomaly Detection of Industrial Equipment Using Audio Data and Machine Learning. *IEEE Access*. 2021. Vol. 9. P. 80936–80944.
- [5] Bahan T., Temchur V., Boun V. Determining the efficiency of vibration signal processing methods for predictive diagnostics of electric motors. *Eastern-European Journal of Enterprise Technologies*. 2025. Vol. 1, No. 1(133). P. 6–16.
- [6] Glowacz A. Fault diagnosis of single-phase induction motor based on acoustic signals. *Mechanical Systems and Signal Processing*. 2021. Vol. 117. P. 65–80.
- [7] Zhao M., Zhong S., Fu X. et al. Deep Residual Shrinkage Networks for Fault Diagnosis. *IEEE Transactions on Industrial Informatics*. 2020. Vol. 16, No. 7. P. 4681–4690.
- [8] Hoang D.T., Kang H.J. A Survey on Deep Learning Based Bearing Fault Diagnosis. *Neurocomputing*. 2020. Vol. 335. P. 327–335.
- [9] Temchur V.S., Bahan T.H. Predictive Diagnostics of Electric Motor Vibration Signal Based on MSCNN-LSTM Technology. *Збірник наукових праць НУК*. 2025. №1. С. 107–112.
- [10] Lu C., Wang Y., Qin W., Ma H. Intelligent Fault Diagnosis of Rotating Machinery Using Hierarchical Convolutional Neural Network. *Neurocomputing*. 2021. Vol. 430. P. 16–28.
- [11] Karim F., Majumdar S., Darabi H., Harford S. Multivariate LSTM-FCNs for

Time Series Classification. *Neural Networks*. 2021. Vol. 116. P. 237–245.

[12] Li T., Zhao Z., Sun C. et al. WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics*. 2022. Vol. 52, No. 4. P. 2302–2312.

[13] Selcuk S. Predictive maintenance, its implementation and latest trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*. 2017. Vol. 231, No. 9. P. 1670–1679.

[14] Lei Y., Yang B., Jiang X. et al. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*. 2020. Vol. 138. Article 106587.

[15] García Márquez F.P., Tobias A.M., Pinar Pérez J.M., Papaelias M. Condition monitoring of wind turbines: Techniques and methods. *Renewable Energy*. 2020. Vol. 46. P. 169–178.

[16] Antonino-Daviu J.A., Riera-Guasp M., Folch J.R., Molina Palomares M.P. Validation of a new method for the diagnosis of rotor bar failures via wavelet transform in industrial induction machines. *IEEE Transactions on Industry Applications*. 2020. Vol. 42, No. 4. P. 990–996.

[17] Tahir M.M., Khan A.Q., Iqbal N., Hussain A., Badshah S. Enhancing Fault Classification Accuracy of Ball Bearing Using Central Tendency Based Time Domain Features. *IEEE Access*. 2017. Vol. 5. P. 72–83.

[18] El Bouharrouiti N., Morinigo-Sotelo D., Belahcen A. Multi-Rate Vibration Signal Analysis for Bearing Fault Detection in Induction Machines Using Supervised Learning Classifiers. *Machines*. 2023. Vol. 12, No. 1. Article 17.

[19] Łuczak D. Machine Fault Diagnosis through Vibration Analysis: Continuous Wavelet Transform with Complex Morlet Wavelet and Time-Frequency RGB Image Recognition via CNN. *Electronics*. 2024. Vol. 13, No. 2. Article 452.

[20] Ahmed H.O.A., Nandi A.K. Vibration Image Representations for Fault Diagnosis of Rotating Machines: A Review. *Machines*. 2022. Vol. 10, No. 12. Article 1113.

[21] Rahman T.A., Chek L.W., Ramli N. Intelligent Vibration-based Anomaly

Detection for Electric Motor Condition Monitoring. 9th Iranian Joint Congress on Fuzzy and Intelligent Systems. 2019.

[22] Quiroz J.C., Mariun N., Mehrjou M.R., Izadi M., Misron N., Radzi M.A.M. Fault Detection of Broken Rotor Bar in LS-PMSM Using Random Forests. arXiv preprint arXiv:1711.02510. 2017.

[23] Misra S., Kumar S., Sayyad S., Bongale A., Jadhav P., Kotecha K., Abraham A., Gabralla L.A. Fault detection in induction motor using time domain and spectral imaging-based transfer learning approach on vibration data. *Sensors*. 2022. Vol. 22, No. 21. Article 8210.

[24] Althubaiti A., Elasha F., Teixeira J.A. Fault Diagnosis and Health Management of Bearings in Rotating Equipment Based on Vibration Analysis. *Journal of Vibroengineering*. 2022. Vol. 24, No. 1. P. 46–74.

[25] Chen Z., Li C., Sanchez R.-V. A Vibration Feature Extraction Method Based on Time-Domain Dimensional Parameters and Mahalanobis Distance. *Mathematical Problems in Engineering*. 2021. Article 2498178.

[26] Manhertz G., Bereczky A. STFT Spectrogram Based Hybrid Evaluation Method for Rotating Machine Transient Vibration Analysis. *Mechanical Systems and Signal Processing*. 2021. Vol. 154. Article 107583.

[27] Alzubaidi L., Zhang J., Humaidi A.J. et al. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*. 2021. Vol. 8. Article 53.

[28] Liang H., Zhao X. Rolling Bearing Fault Diagnosis Based on One-Dimensional Dilated Convolution Network with Residual Connection. *IEEE Access*. 2021. Vol. 9. P. 31078–31091.

[29] Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd ICML*. 2015. P. 448–456.

[30] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. *IEEE CVPR*. 2016. P. 770–778.

[31] Liang R., Ran W., Chen Y., Zhu R. Fault Diagnosis Method for Rotating

Machinery Based on Multi-Scale Features. Chinese Journal of Mechanical Engineering. 2023. Vol. 36, No. 1. Article 141.

[32] Wang Y., Cao G. A Multiscale Convolution Neural Network for Bearing Fault Diagnosis Based on Frequency Division Denoising Under Complex Noise Conditions. Complex & Intelligent Systems. 2023. Vol. 9. P. 4263–4285.

[33] Kumar A., Vashishtha G., Gandhi C.P. et al. Novel Convolutional Neural Network (NCNN) for the Diagnosis of Bearing Defects in Rotary Machinery. IEEE Transactions on Instrumentation and Measurement. 2021. Vol. 70. Article 3508010.

[34] Liang M., Cao P., Tang J. Rolling Bearing Fault Diagnosis Based on Feature Fusion with Parallel Convolutional Neural Network. International Journal of Advanced Manufacturing Technology. 2021. Vol. 112. P. 819–831.

[35] Xu M., Shi Y., Deng M. et al. An Improved Multi-Scale Branching Convolutional Neural Network for Rolling Bearing Fault Diagnosis. PLOS ONE. 2023. Vol. 18, No. 9. Article e0291353.

[36] Hochreiter S., Schmidhuber J. Long Short-Term Memory. Neural Computation. 1997. Vol. 9, No. 8. P. 1735–1780.

[37] Yu Y., Si X., Hu C., Zhang J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. Neural Computation. 2019. Vol. 31, No. 7. P. 1235–1270.

[38] Sun W. et al. A Novel Hybrid Deep Learning Method for Fault Diagnosis of Rotating Machinery Based on Extended WDCNN and Long Short-Term Memory. Sensors. 2021. Vol. 21, No. 19. Article 6550.

[39] Chen X., Zhang B., Gao D. Bearing Fault Diagnosis Based on Multi-Scale CNN and LSTM Model. Journal of Intelligent Manufacturing. 2021. Vol. 32, No. 4. P. 971–987.

[40] Guo Y., Mao J., Zhao M. Rolling Bearing Fault Diagnosis Method Based on Attention CNN and BiLSTM Network. Neural Processing Letters. 2023. Vol. 55. P. 3377–3410.

[41] Huang T., Zhang Q., Tang X., Zhao S., Lu X. A Novel Fault Diagnosis Method Based on CNN and LSTM and Its Application in Fault Diagnosis for Complex Systems.

Artificial Intelligence Review. 2022. Vol. 55. P. 1289–1315.

[42] Shi H., Guo L., Tan S., Bai X. Rolling Bearing Initial Fault Detection Using Long Short-Term Memory Recurrent Network. *IEEE Access*. 2019. Vol. 7. P. 171559–171569.

[43] Nath A.G., Udmale S.S., Singh S.K. Role of Artificial Intelligence in Rotor Fault Diagnosis: A Comprehensive Review. *Artificial Intelligence Review*. 2021. Vol. 54. P. 2609–2668

[44] Zhu Z., Lei Y., Qi G. et al. A Review of the Application of Deep Learning in Intelligent Fault Diagnosis of Rotating Machinery. *Measurement*. 2023. Vol. 206. Article 112346.

[45] Zhao H., Sun S., Jin B. Sequential Fault Diagnosis Based on LSTM Neural Network. *IEEE Access*. 2018. Vol. 6. P. 12929–12939.

[46] Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. San Diego. 2015. arXiv:1412.6980.

[47] Tama B.A., Vania M., Lee S., Lim S. Recent Advances in the Application of Deep Learning for Fault Diagnosis of Rotating Machinery Using Vibration Signals. *Artificial Intelligence Review*. 2023. Vol. 56. P. 4667–4709.

[48] Liu R., Yang B., Zio E., Chen X. Artificial Intelligence for Fault Diagnosis of Rotating Machinery: A Review. *Mechanical Systems and Signal Processing*. 2018. Vol. 108. P. 33–47.

[49] Wang Y., Peng Y., Zi Y. et al. A Two-Stage Data-Driven-Based Prognostic Approach for Bearing Degradation Problem. *IEEE Transactions on Industrial Informatics*. 2016. Vol. 12, No. 3. P. 924–932.

[50] Kundu P., Darpe A.K., Kulkarni M.S. Weibull Accelerated Failure Time Regression Model for Remaining Useful Life Prediction of Bearing Working under Multiple Operating Conditions. *Mechanical Systems and Signal Processing*. 2019. Vol. 134. Article 106302.

[51] Ferreira C., Gonçalves G. Remaining Useful Life Prediction and Challenges: A Literature Review on the Use of Machine Learning Methods. *Journal of Manufacturing*

Systems. 2022. Vol. 63. P. 550–562.

[52] Lei Y., Li N., Guo L. et al. Machinery Health Prognostics: A Systematic Review from Data Acquisition to RUL Prediction. *Mechanical Systems and Signal Processing*. 2018. Vol. 104. P. 799–834.

[53] Zio E. Prognostics and Health Management (PHM): Where Are We and Where Do We Need to Go in Theory and Practice. *Reliability Engineering & System Safety*. 2022. Vol. 218. Article 108119.

[54] Si X.S., Wang W., Hu C.H., Zhou D.H. Remaining Useful Life Estimation — A Review on the Statistical Data Driven Approaches. *European Journal of Operational Research*. 2011. Vol. 213, No. 1. P. 1–14.

[55] Yang F., Habibullah M.S., Shen Y. Remaining Useful Life Prediction of Induction Motors Using Nonlinear Degradation of Health Index. *Mechanical Systems and Signal Processing*. 2021. Vol. 148. Article 107183.

[56] Gebraeel N., Lei Y., Li N., Si X., Zio E. Prognostics and Remaining Useful Life Prediction of Machinery: Advances, Opportunities and Challenges. *Journal of Dynamics, Monitoring and Diagnostics*. 2023. P. 1–12.

[57] Zhao Z., Wu J., Li T. et al. Challenges and Opportunities of AI-Enabled Monitoring, Diagnosis and Prognosis: A Review from Artificial Intelligence Perspective. *Chinese Journal of Mechanical Engineering*. 2021. Vol. 34. Article 56.

[58] Tiddens W., Braaksma J., Tinga T. Decision Framework for Predictive Maintenance Method Selection. *Applied Sciences*. 2023. Vol. 13, No. 3. Article 2021.

[59] Abadi M., Barham P., Chen J. et al. TensorFlow: A System for Large-Scale Machine Learning. *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*. 2016. P. 265–283.

[60] Virtanen P., Gommers R., Oliphant T.E. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. 2020. Vol. 17. P. 261–272.

[61] Ribeiro F.M.L., Marins M.A., Netto S.L. et al. Rotating Machinery Fault Diagnosis Using Similarity-Based Models. *Proceedings of the XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*. 2019. P. 1–5.

[62] Randall R.B., Antoni J. Rolling Element Bearing Diagnostics — A Tutorial. *Mechanical Systems and Signal Processing*. 2011. Vol. 25, No. 2. P. 485–520.

[63] He M., He D. Deep Learning Based Approach for Bearing Fault Diagnosis. *IEEE Transactions on Industry Applications*. 2017. Vol. 53, No. 3. P. 3057–3065.

[64] Pederiva R., Gongora W.S. Vibration Signals of a Squirrel Cage Induction Motor with Broken Bars. *IEEE Dataport*. 2021. DOI: 10.21227/mrex-mn61.

[65] Jia F., Lei Y., Lin J. et al. Deep Neural Networks: A Promising Tool for Fault Characteristic Mining and Intelligent Diagnosis of Rotating Machinery with Massive Data. *Mechanical Systems and Signal Processing*. 2016. Vol. 72-73. P. 303–315.

ДОДАТОК А

Документи та акти, які підтверджують впровадження дисертаційної роботи



А К Т

впровадження результатів наукових досліджень

Я, що нижче підписався, директор Диченко Олександр Іванович, склав цей акт про те, що результати наукових досліджень Темчура Владислава Сергійовича, які отримані у дисертаційній роботі «Предиктивна діагностика обертового енергетичного устаткування на основі нейромережевого аналізу вібраційних сигналів» використані у діяльності ТОВ «Вентконтрол», а саме:

- методика формування двоканальних спектрограм вібраційних сигналів електродвигунів приводних механізмів для автоматизованої оцінки їх технічного стану;
- програмний модуль класифікації стану обертового устаткування на основі нейромережевої моделі, що забезпечує розпізнавання типу несправності за вібраційними даними;
- методика оцінки індексу здоров'я обладнання та прогнозування залишкового ресурсу, що дозволяє завчасно планувати технічне обслуговування приводних механізмів.

Застосування зазначених результатів у практиці технічного обслуговування обертового устаткування підприємства дозволило перейти від планово-попереджувальної стратегії обслуговування електродвигунів до обслуговування за фактичним станом, що забезпечило скорочення кількості незапланованих зупинок обладнання та зменшення витрат на технічне обслуговування.

Директор



Олександр ДИЧЕНКО