

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Кваліфікаційна наукова праця
на правах рукопису

БАРАНОВ ОЛЕКСІЙ МИКОЛАЙОВИЧ

УДК 004.85

ДИСЕРТАЦІЯ

**МОДЕЛІ АНСАМБЛЕВОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ
СТЕКІНГУ НЕЙРОННИХ МЕРЕЖ**

Спеціальність: 122 – Комп'ютерні науки

Галузь знань: 12 – Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії.

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

_____ О.М. Баранов

Науковий керівник:

Галчонков Олег Миколайович, кандидат технічних наук, доцент

Одеса – 2026

АНОТАЦІЯ

Баранов О.М. Моделі ансамблевої класифікації зображень на основі стекінгу нейронних мереж. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки. – Національний університет «Одеська політехніка», МОН України, Одеса, 2026.

Дисертаційна робота присвячена розробці моделей і методів підвищення ефективності ансамблевої класифікації зображень на основі стекінгу нейронних мереж для використання у застосуваннях з невеликими розмірами зображень та невеликими наборами даних для навчання.

Метою дисертаційної роботи є підвищення якості класифікації та/або зменшення обсягу обчислень при використанні ансамблевої класифікації зображень зі стекінгом нейронних мереж шляхом розробки нейромережових моделей і методів.

Для досягнення мети в дисертаційній роботі розв'язуються такі завдання:

1. Провести аналіз підходів ансамблевої класифікації зображень, які забезпечують підвищене співвідношення якості класифікації та необхідного обсягу обчислень і здатних в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

2. Розробити модель першого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж та метод визначення її структурно-параметричних ознак.

3. Розробити модель другого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж.

4. Дослідити використання розроблених моделей і методу, при вирішенні задачі класифікації зображень в автономних робототехнічних комплексах.

Об'єкт дослідження – процес ансамблевої класифікації зображень в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

Предмет дослідження – моделі і методи ансамблевої класифікації зображень зі стекінгом нейронних мереж.

Результати дисертаційного дослідження ґрунтуються на використанні теорії нейронних мереж, теорії матриць, математичної статистики і теорії ймовірностей, обчислювальних методів лінійної алгебри та функціонального аналізу в частині розкладання у ряди функцій і апроксимації.

На основі проведених теоретичних і експериментальних досліджень розв’язано **актуальне наукове завдання**:

- розробка моделі і методу, що дозволяють збільшувати кількість нейронних мереж на першому ступені в ансамблевих класифікаторах зі стекінгом НМ;
- розвиток моделей для об’єднання результатів базових нейронних мереж на другому ступені ансамблевого класифікатора зі стекінгом НМ.

Отримані такі **нові наукові результати**:

1. Вдосконалено модель першого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних, шляхом додавання до заданої кількості базових класифікаторів їх клонів, отриманих навчанням на вибірках зображень, які створено за визначеними операторами перетворення, що дозволило запропонувати метод визначення структурно-параметричних ознак моделі.
2. Вперше запропоновано метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зображень зі стекінгом нейронних мереж, які побудовані з використанням операторів повороту та циклічного зсуву вхідних зображень, що дозволило створити ефективний ансамблевий класифікатор для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних.
3. Вперше запропоновано модель нейронної мережі MLP-KAN, яка відрізняється від моделі багатошарового перцептрона заміною першого прихованого шару MLP на шар Колмогорова-Арнольда (KAN), що дало змогу використовувати два типи апроксимації одночасно, а саме, за рахунок навчання функцій активації та навчання ваг нейронів, що дозволило зменшити обсяг вагових

коефіцієнтів, що навчаються, при збереженні якості класифікації на наборі тестових зображень.

4. Вдосконалено модель другого ступеня ансамблевого класифікатора зі стекінгом НМ, шляхом заміни класифікатора MLP на адаптовану модель MLP-KAN з урахуванням особливостей вихідного вектору ознак першого ступеня ансамблевого класифікатора, що дало змогу зменшити обсяг обчислень при збереженні якості класифікації зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних.

Дисертаційна робота виконана відповідно до плану науково-дослідних робіт Національного університету «Одеська політехніка» згідно з грантом НФДУ № 131/0161 «Використання штучного інтелекту для підвищення якості визначення вибухових пристроїв мультисенсорним дроном-міношукачем» (довідка впровадження № 1009/64-07 від 11.06.2026), в межах виконання науково-дослідної роботи №723-43 «Розробка інтелектуальних сенсорів дистанційного комплексного моніторингу деградації ґрунтів повоєнних територій» (довідка впровадження № 1009/64-07 від 11.06.2026).

Практичне значення отриманих у дисертаційній роботі результатів полягає в:

1. Розробці моделей і метода підвищення якості класифікації зображень у практичних застосуваннях з обмеженим набором даних для навчання, обмеженою обчислювальною потужністю, що використовується, та які потребують обробки в реальному масштабі часу.

2. Розробці моделей ансамблевої класифікації зі стекінгом нейронних мереж для зменшення обсягу вагових коефіцієнтів, що навчаються, при збереженні якості класифікації об'єктів на зображеннях.

3. Розробці моделей та метода побудови архітектури обробки вхідних даних у класифікаторах зображень, що відрізняються підвищеною здатністю до розпаралелювання обчислень з метою реалізації обробки в реальному масштабі часу.

Результати дисертаційної роботи були апробовані та впроваджені:

– у програмному забезпеченні дрона-міношукача, розробленого Національним університетом «Одеська політехніка» за грантом № 131/0161 Національного фонду досліджень України «Використання штучного інтелекту для підвищення якості визначення вибухових пристроїв мультисенсорним дроном-міношукачем» (довідка впровадження № 1009/64-07 від 11.06.2026);

– у теоретичних розробках перспективних моделей обробки сигналів в межах виконання науково-дослідної роботи №723-43 «Розробка інтелектуальних сенсорів дистанційного комплексного моніторингу деградації ґрунтів повоєнних територій» (довідка впровадження № 1009/64-07 від 11.06.2026),

– у навчальному процесі Національного університету «Одеська політехніка» (довідка впровадження № 1008/64-07 від 11.06.2026).

Впровадження розроблених моделей і метода забезпечує підвищення ефективності ансамблевої класифікації зі стекінгом за рахунок доповнення структури першого ступеня класифікатора клонами базових класифікаторів і вдосконалення другого ступеня ансамблевого класифікатора шляхом заміни прихованого шару моделі MLP на шар KAN, що дозволило зменшити кількість навчальних вагових коефіцієнтів моделі при збереженні якості класифікації об'єктів на зображеннях.

Запропоновані моделі та метод дають змогу здійснювати побудову ефективного ансамблевого класифікатора відповідно до особливостей вхідних даних.

Результати дисертаційних досліджень опубліковано у 8 наукових працях, які включають 6 статей (з них 2 у фаховому періодичному науковому журналі України з технічних наук категорії «Б», 4 статті у періодичних наукових журналах України з технічних наук, що індексуються у наукометричній базі Scopus), 2 тези доповідей, що були представлені в матеріалах наукових конференцій, що індексуються у наукометричній базі Scopus.

Ключові слова: нейронні мережі, мережа Колмогорова-Арнольда, нейрони, навчальні функції активації, ансамблева класифікація, радіально-базисні функції, багатошаровий перцептрон, трансформер.

Список публікацій здобувача

1. Galchonkov O., Baranov O., Babych M., Kuvaieva V., Babych Y. Improving the quality of object classification in images by ensemble classifiers with stacking. *Eastern-European Journal of Enterprise Technologies*, 2023, 3 (9 (123)), 70–77.
DOI: <https://doi.org/10.15587/1729-4061.2023.279372> (Indexed in SCOPUS)
2. Galchonkov O., Baranov O., Chervonenko P., Babilunga O. Reducing the volume of computations when building analogue of neural networks for the first stage of an ensemble classifier with stacking. *Eastern-European Journal of Enterprise Technologies*, 2024, 2 (9 (128)), 27–35.
DOI: <https://doi.org/10.15587/1729-4061.2024.299734> (Indexed in SCOPUS)
3. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babych M. Development of a neural network with a layer of trainable activation functions for the second stage of the ensemble classifier with stacking. *Eastern-European Journal of Enterprise Technologies*, 2024, 5 (9 (131)), 6–13.
DOI: <https://doi.org/10.15587/1729-4061.2024.311778> (Indexed in SCOPUS)
4. Galchonkov O., Baranov O., Maslov O., Babych M., Baskov I. MLP-KAN: implementation of the Kolmogorov-Arnold layer in a multilayer perceptron. *Eastern-European Journal of Enterprise Technologies*, 2025, 3 (4 (135)), 34–41.
DOI: <https://doi.org/10.15587/1729-4061.2025.328928> (Indexed in SCOPUS)
5. Galchonkov O., Baranov O., Baskov I. Reducing the amount of computations at the second stage of an ensemble classifier with stacking on mine-detection drones. *Applied Aspects of Information Technology*. 2024; Vol.7 No.4: 370–377.
DOI: <https://doi.org/10.15276/aait.07.2024.26> (Indexed in Copernicus)
6. Galchonkov O., Baranov O., Baskov I. Improving the quality of mine detection by a mine-detecting drone by selecting an architecture for classifying objects in visible-light camera images. *Applied Aspects of Information Technology*, 2026; Vol.9 No.1: 9–24.
DOI: <https://doi.org/10.15276/aait.09.2026.01> (Indexed in Copernicus)
7. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babilunga O., Shunin O. Using Stacked Ensemble Classifiers to Speed Up Object Recognition in Images for

Mine Detection Drones. Proceedings of 2024 IEEE 7th International Conference on ACTUAL PROBLEMS OF UNMANNED AERIAL VEHICLES DEVELOPMENTS (APUAVD). October 22-24, 2024 Kyiv, Ukraine. P.262-265.

DOI: <https://ieeexplore.ieee.org/xpl/conhome/10765843/proceeding> (Indexed in SCOPUS)

8. Galchonkov O., Baranov O., Bascov I. Building a Stacked Ensemble Object Classifier on Images Using the Base CCT Classifier and Its Clones. Proceedings of 2025 IEEE 8th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC). October 21-24, 2025 Kyiv, Ukraine. P.280-284.

DOI: 10.1109/MSNMC61130.2025.11399700 (Indexed in SCOPUS)

ABSTRACT

Baranov O.M. Models of ensemble image classification based on the stacking of neural networks. – Qualifying scientific work on the rights of a manuscript.

Dissertation for the scientific degree of Doctor of Philosophy in specialty 122 – Computer Science. – National University "Odesa Polytechnic", Ministry of Education and Science of Ukraine, Odesa, 2026.

The dissertation is devoted to the development of models and methods for improving the efficiency of ensemble image classification based on the stacking of neural networks for use in applications with small image sizes and small datasets for training.

The aim of the dissertation is to improve the quality of classification and/or reduce the volume of computations when using ensemble image classification with stacking of neural networks through the development of neural network models and methods.

To achieve this aim, the following tasks are solved in the dissertation:

1. To analyze the approaches to ensemble object classification in images that provide an increased ratio of classification quality to the required volume of computations and are capable of operating under conditions of low image resolution and a deficit of training data.
2. To develop a first-stage model of an ensemble classifier with the stacking of neural networks and a method for determining its structural-parametric features.
3. To develop a second-stage model of an ensemble classifier with the stacking of neural networks.
4. To investigate the use of the developed models and method in solving the problem of image classification in autonomous robotic complexes.

Object of research – the process of ensemble image classification under conditions of low image resolution and a deficit of training data.

Subject of research – models and methods of ensemble image classification with the stacking of neural networks.

The results of the dissertation research are based on the use of neural network theory, matrix theory, mathematical statistics and probability theory, computational

methods of linear algebra, and functional analysis in terms of series expansion of functions and approximation.

Based on the conducted theoretical and experimental research, an **actual scientific task** has been solved:

- the development of new models and a method that allow increasing the number of neural networks at the first stage in ensemble classifiers with NN stacking;
- the development of new models for combining the results of primary neural networks at the second stage of an ensemble classifier with NN stacking.

The following **new scientific results** were obtained:

1. The model of an ensemble classifier with stacking of neural networks for low-resolution images under conditions of a training data deficit has been improved by adding to a specified number of basic classifiers their clones obtained by training on image samples created by defined transformation operators, which allowed to propose a method for determining the structural-parametric features of the model.

2. For the first time, a method for determining the structural-parametric features of the first-stage model of an ensemble classifier with stacking of neural networks constructed using rotation and cyclic shift of input images operators is proposed, which allows creating an efficient ensemble classifier for low-resolution images under conditions of a training data deficit.

3. For the first time, an MLP-KAN neural network model is proposed, which differs from the multilayer perceptron (MLP) model by replacing the first hidden layer of the MLP with a Kolmogorov-Arnold layer (KAN), which allowed using two types of approximation simultaneously, namely, the training of activation functions and the training of neuron weights, which allowed reducing the volume of trainable weight coefficients while maintaining the quality of classification on a set of test images.

4. The second-stage model of an ensemble classifier with NN stacking has been improved by using the developed MLP-KAN neural network model, taking into account the features of the output feature vector of the first stage of the ensemble classifier, which made it possible to significantly increase the efficiency of ensemble classification with NN stacking.

The dissertation was carried out in accordance with the scientific research plan of the National University "Odesa Polytechnic" under the NRFU grant No. 131/0161 "Use of artificial intelligence to improve the quality of explosive device detection by a multisensor mine-detecting drone" (implementation certificate № 1009/64-07 dated 11.06.2026), within the framework of scientific research work No. 723-43 "Development of intelligent sensors for remote complex monitoring of soil degradation in post-war territories" (implementation certificate № 1009/64-07 dated 11.06.2026).

The practical significance of the results obtained in the dissertation lies in:

1. The development of models and a method for improving the quality of image classification in practical applications with a limited dataset for training, limited computational power used, and requiring real-time processing.
2. The development of models for ensemble classification with stacking of neural networks to reduce the volume of trainable weight coefficients while maintaining the quality of object classification in images.
3. The development of models and a method for constructing the input data processing architecture in image object classifiers, which are characterized by an increased ability to parallelize computations in order to implement real-time processing.

The results of the dissertation work were tested and implemented:

- in the software of a mine-detecting drone developed by the National University "Odesa Polytechnic" under grant No. 131/0161 of the National Research Foundation of Ukraine "Use of artificial intelligence to improve the quality of explosive device detection by a multisensor mine-detecting drone" (implementation certificate № 1009/64-07 dated 11.06.2026);
- in the theoretical developments of prospective signal processing models within the framework of scientific research work No. 723-43 "Development of intelligent sensors for remote complex monitoring of soil degradation in post-war territories" (implementation certificate № 1009/64-07 dated 11.06.2026);
- in the educational process of the National University "Odesa Polytechnic" (implementation certificate № 1009/64-07 dated 11.06.2026).

The implementation of the developed models and method provides an increase in the efficiency of ensemble classification with stacking by supplementing the structure of the first stage of the classifier with clones of basic classifiers and improving the second stage of the ensemble classifier by replacing the hidden layer of the MLP model with a KAN layer, which allowed reducing the number of trainable weight coefficients of the model while maintaining the quality of object classification in images.

The proposed models and method allow building an efficient ensemble classifier according to the features of the input data.

The results of the dissertation research are published in 8 scientific papers, which include 6 articles (2 of them in a professional periodic scientific journal of Ukraine in technical sciences of category "B", 4 articles in periodic scientific journals of Ukraine in technical sciences indexed in the Scopus scientometric database), and 2 report abstracts presented in the proceedings of scientific conferences indexed in the Scopus scientometric database.

Keywords: neural networks, Kolmogorov-Arnold networks, neurons, trainable activation functions, ensemble classification, radial basis functions, multilayer perceptron, transformer.

List of publications

1. Galchonkov O., Baranov O., Babych M., Kuvaieva V., Babych Y. Improving the quality of object classification in images by ensemble classifiers with stacking. Eastern-European Journal of Enterprise Technologies, 2023, 3 (9 (123)), 70–77.

DOI: <https://doi.org/10.15587/1729-4061.2023.279372> (Indexed in SCOPUS)

2. Galchonkov O., Baranov O., Chervonenko P., Babilunga O. Reducing the volume of computations when building analogue of neural networks for the first stage of an ensemble classifier with stacking. Eastern-European Journal of Enterprise Technologies, 2024, 2 (9 (128)), 27–35.

DOI: <https://doi.org/10.15587/1729-4061.2024.299734> (Indexed in SCOPUS)

3. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babych M. Development of a neural network with a layer of trainable activation functions for the

second stage of the ensemble classifier with stacking. *Eastern-European Journal of Enterprise Technologies*, 2024, 5 (9 (131)), 6–13.

DOI: <https://doi.org/10.15587/1729-4061.2024.311778> (Indexed in SCOPUS)

4. Galchonkov O., Baranov O., Maslov O., Babych M., Baskov I. MLP-KAN: implementation of the Kolmogorov-Arnold layer in a multilayer perceptron. *Eastern-European Journal of Enterprise Technologies*, 2025, 3 (4 (135)), 34–41. <https://doi.org/10.15587/1729-4061.2025.328928> (Indexed in SCOPUS)

5. Galchonkov O., Baranov O., Baskov I. Reducing the amount of computations at the second stage of an ensemble classifier with stacking on mine-detection drones. *Applied Aspects of Information Technology*. 2024; Vol.7 No.4: 370–377.

DOI: <https://doi.org/10.15276/aait.07.2024.26> (Indexed in Copernicus)

6. Galchonkov O., Baranov O., Baskov I. Improving the quality of mine detection by a mine-detecting drone by selecting an architecture for classifying objects in visible-light camera images. *Applied Aspects of Information Technology*, 2026; Vol.9 No.1: 9–24.

DOI: <https://doi.org/10.15276/aait.09.2026.01> (Indexed in Copernicus)

7. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babilunga O., Shunin O. Using Stacked Ensemble Classifiers to Speed Up Object Recognition in Images for Mine Detection Drones. *Proceedings of 2024 IEEE 7th International Conference on ACTUAL PROBLEMS OF UNMANNED AERIAL VEHICLES DEVELOPMENTS (APUAVD)*. October 22-24, 2024 Kyiv, Ukraine. P.262-265.

DOI: <https://ieeexplore.ieee.org/xpl/conhome/10765843/proceeding> (Indexed in SCOPUS)

8. Galchonkov O., Baranov O., Bascov I. Building a Stacked Ensemble Object Classifier on Images Using the Base CCT Classifier and Its Clones. *Proceedings of 2025 IEEE 8th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC)*. October 21-24, 2025 Kyiv, Ukraine. P.280-284.

DOI: [10.1109/MSNMC61130.2025.11399700](https://doi.org/10.1109/MSNMC61130.2025.11399700) (Indexed in SCOPUS)

ЗМІСТ

Перелік скорочень, умовних позначень, одиниць і термінів.....	16
Вступ.....	17
1 АНАЛІЗ СТАНУ АНСАМБЛЕВОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ СТЕКІНГУ НЕЙРОННИХ МЕРЕЖ.....	24
1.1 Аналіз методів побудови нейромережових ансамблевих класифікаторів зображень в умовах низької роздільної здатності та дефіциту навчальних даних.....	24
1.2 Формальне визначення задачі ансамблевої класифікації зображень зі стекінгом нейронних мереж.....	26
1.3 Аналіз відомих моделей нейронних мереж першого ступеня ансамблевого класифікатора зі стекінгом.....	29
1.4 Аналіз підходів до зменшення обсягу обчислень на другому ступені ансамблевого класифікатора.....	40
1.5 Постановка завдання дослідження.....	46
Висновки до розділу 1.....	47
2 РОЗРОБКА МОДЕЛІ ПЕРШОГО СТУПЕНЯ АНСАМБЛЕВОГО КЛАСИФІКАТОРА ЗІ СТЕКІНГОМ НА ОСНОВІ КЛОНІВ НЕЙРОННИХ МЕРЕЖ ТА МЕТОДУ ВИЗНАЧЕННЯ ЇЇ ОЗНАК.....	49
2.1 Модель ансамблевого класифікатора зі стекінгом на основі клонів нейронних мереж.....	49
2.2 Використанням оператора повороту для створення клонів нейронних мереж на першому ступені ансамблевого класифікатора зі стекінгом нейронних мереж.....	54
2.3 Використання оператора циклічного зсуву для створення клонів нейронних мереж.....	65
2.4 Порівняння якості класифікації при використанні на першому ступені моделі ансамблевого класифікатора зі стекінгом клонів НМ, отриманих за допомогою операторів зсуву і повороту зображень.....	72

2.5 Метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зі стекінгом на основі клонів нейронних мереж.....	76
Висновки до розділу 2.....	78
3 РОЗРОБКА МОДЕЛІ MLP-KAN ТА ЇЇ ВИКОРИСТАННЯ НА ДРУГОМУ СТУПЕНІ АНСАМБЛЕВОГО КЛАСИФІКАТОРА ЗІ СТЕКІНГОМ.....	81
3.1 Модель нейронної мережі MLP-KAN для задач класифікації загального типу.....	81
3.2 Вдосконалення моделі нейронної мережі для другого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж.....	95
Висновки до розділу 3.....	105
4 ДОСЛІДЖЕННЯ ЗАПРОПОНОВАНИХ МОДЕЛЕЙ, МЕТОДУ, ПРОГРАМНИХ ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ В АВТОНОМНИХ РОБОТОТЕХНІЧНИХ КОМПЛЕКСАХ.....	107
4.1 Особливості класифікації зображень з використання нейронних мереж в автономних робототехнічних комплексах.....	107
4.2 Опис використовуваного набору даних із зображеннями вибухових пристроїв.....	108
4.3 Класифікація зображень на наявність мін з використанням базових моделей нейронних мереж.....	112
4.4 Обговорення результатів дослідження ефективності використання різних нейронних мереж для виявлення мін дроном.....	134
4.5 Дослідження якості класифікації зображень на наявність мін з використанням запропонованого ансамблевого класифікатора.....	135
4.6 Дослідження швидкодії запропонованого ансамблевого класифікатора зображень при обробці відеопослідовності.....	136
Висновки до розділу 4.....	138
Висновки.....	140
Список використаних джерел.....	143
Додаток А Приклади зображень з наборів даних MNIST, CIFAR10.....	155
Додаток Б Список публікацій здобувача.....	157

Додаток В Акти впровадження.....	159
----------------------------------	-----

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ОДИНИЦЬ І ТЕРМІНІВ

НМ - нейронна мережа.

CCT - Compact Convolutional Transformer.

CNN - Convolutional Neural Network.

EATNet - External Attention Transformer Network.

FLOPS - Floating-point operations per second.

FNet - Fourier Network.

FPGA - Field-programmable gate array.

FPN - Feature Pyramid Networks.

GBM - Gradient Boosting Machine.

gMLP - Gated Multilayer Perceptron.

MLP - Multilayer Perceptron.

LN - Layer Normalization.

KAN - Kolmogorov-Arnold Networks.

RBF - Radial basis function.

SwinTr - Swin Transformer.

ViT - Visual Transformer.

ВСТУП

Актуальність теми. Нейронні мережі (НМ) знаходять усе ширше застосування в усіх сферах людського життя. Цьому сприяє постійне підвищення ефективності їхньої роботи. Однак у більшості випадків це досягається за рахунок використання більш глибоких НМ, більш потужніших суперкомп'ютерів і більш великих наборів даних.

Водночас у багатьох нових застосуваннях не вдається сформулювати дуже великі набори даних для навчання і вимагається обробка сигналів у реальному масштабі часу. Наприклад, набір зображень зразків крові під мікроскопом BloodMNIST [9] містить лише 17092 зображення. Аналогічний набір даних для дослідження крові з метою виявлення паразитів малярії [10] містить 27588 зображень. Набір зображень із комп'ютерного томографа для навчання НМ класифікації COVID-19 містить 6752 зображення легень 4154 пацієнтів [11, 12].

Для НМ розроблено велику кількість методів компресії або проріджування [9], що зменшують обсяг обчислень. Проте, при застосуванні їх до дуже великих НМ, залишковий обсяг обчислень усе одно залишається досить значним для реалізації на звичайних комп'ютерах, а тим більше у реальному масштабі часу. Тому набув розвитку окремий напрямок у галузі НМ – НМ з високим співвідношенням якісних характеристик до необхідного обсягу обчислень. Відмінними рисами таких НМ є структура, розрахована на обробку зображень в умовах низької роздільної здатності зображень та дефіциту навчальних даних, здатність навчатися з нуля на невеликих наборах даних за відносно невеликий час, можливість реалізації на бортовому обладнанні з обмеженими обчислювальними можливостями в реальному масштабі часу. Характерним прикладом використання таких НМ є автономні робототехнічні комплекси, зокрема дрони-міношукачі. Кількість таких комплексів у різних галузях стрімко зростає.

Тому актуальною є розробка нових модифікацій моделей НМ, які відповідають критерію: підвищене співвідношення якості класифікації до

необхідного обсягу обчислень, а також можливість навчання в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

Додаткові можливості в цьому напрямі відкриває об'єднання НМ в ансамблі. При цьому найбільш перспективним є використання стекінгу. Цей архітектурний прийом дає змогу нарощувати можливості НМ за рахунок збільшення її ширини. Найбільший ефект від використання стекінгу досягається, коли об'єднуються НМ з різною архітектурою та різними способами обробки вхідних сигналів.

Однак, як показав проведений аналіз, основною перешкодою під час розв'язання цього завдання є обмежена кількість відомих ефективних, з погляду співвідношення якості класифікації та обсягу обчислень, НМ, що відрізняються одна від одної структурою обробки вхідних сигналів. У зв'язку з цим, для підвищення ефективності ансамблевих класифікаторів актуальним науковим завданням є:

- розвиток нових моделей і методів, що дозволяють збільшувати кількість НМ на першому ступені в ансамблевих класифікаторах зі стекінгом;
- розвиток нових моделей і методів для об'єднання результатів первинних НМ на другому ступені ансамблевого класифікатора зі стекінгом.

Зв'язок дисертаційної роботи з науковими програмами, планами, темами. Дисертаційна робота виконана відповідно до плану робіт Національного університету «Одеська політехніка» згідно з грантом НФДУ № 131/0161 «Використання штучного інтелекту для підвищення якості визначення вибухових пристроїв мультисенсорним дроном-міношукачем» (довідка впровадження № 1009/64-07 від 11.06.2026), в межах виконання науково-дослідної роботи №723-43 «Розробка інтелектуальних сенсорів дистанційного комплексного моніторингу деградації ґрунтів повоєнних територій» (довідка впровадження № 1009/64-07 від 11.06.2026).

Мета і завдання дослідження. Метою дисертаційної роботи є підвищення якості класифікації та зменшення обсягу обчислень при використанні ансамблевої класифікації зображень зі стекінгом нейронних мереж шляхом розробки нейромережеских моделей і методів.

Для досягнення поставленої мети в дисертаційній роботі розв'язуються такі завдання:

1. Провести аналіз підходів ансамблевої класифікації зображень, які забезпечують підвищене співвідношення якості класифікації та необхідного обсягу обчислень і здатних навчатися в умовах низької роздільної здатності зображень та дефіциту навчальних даних.
2. Розробити модель першого ступеня ансамблевого класифікатора зі стекінгом НМ та метод визначення її структурно-параметричних ознак.
3. Розробити модель другого ступеня ансамблевого класифікатора зі стекінгом НМ.
4. Дослідити використання розроблених моделей і методу, при вирішенні задачі класифікації зображень в автономних робототехнічних комплексах.

Об'єкт дослідження – процес ансамблевої класифікації зображень в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

Предмет дослідження – моделі і методи ансамблевої класифікації зображень зі стекінгом нейронних мереж.

Наукова новизна отриманих результатів.

На основі проведених теоретичних і експериментальних досліджень отримано такі нові наукові результати:

1. Вдосконалено модель першого ступеня ансамблевого класифікатора зі стекінгом НМ для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних, шляхом додавання до заданої кількості базових класифікаторів їх клонів отриманих навчанням на вибірках зображень, які створено за визначеними операторами перетворення, що дозволило запропонувати метод визначення структурно-параметричних ознак моделі.
2. Вперше запропоновано метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зображень зі стекінгом НМ, які побудовані з використанням операторів повороту та циклічного зсуву вхідних зображень, що дозволило створити ефективний ансамблевий класифікатор

для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних.

3. Вперше запропоновано модель НМ MLP-KAN, яка відрізняється від моделі багат шарового перцептрона (англ. Multilayer Perceptron, MLP) заміною першого прихованого шару MLP на шар Колмогорова-Арнольда (англ. Kolmogorov-Arnold Networks, KAN), що дало змогу використовувати два типи апроксимації одночасно, а саме, за рахунок навчання функцій активації та навчання ваг нейронів, що дозволило зменшити обсяг вагових коефіцієнтів, що навчаються, при збереженні якості класифікації на наборі тестових зображень.

4. Вдосконалено модель другого ступеня ансамблевого класифікатора зі стекінгом НМ, шляхом заміни класифікатора MLP на адаптовану модель MLP-KAN з урахуванням особливостей вихідного вектору ознак першого ступеня ансамблевого класифікатора, що дало змогу зменшити обсяг обчислень при збереженні якості класифікації зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних.

Практичне значення отриманих теоретичних результатів.

Практичне значення отриманих у дисертаційній роботі результатів полягає в:

1. Розробці моделей і методів підвищення якості класифікації зображень у практичних застосуваннях з обмеженим набором даних для навчання, обмеженою обчислювальною потужністю, що використовується, та які потребують обробки в реальному масштабі часу.

2. Розробці моделей ансамблевої класифікації зі стекінгом НМ для зменшення обсягу вагових коефіцієнтів, що навчаються, при збереженні якості класифікації зображень.

3. Розробці моделей та методів побудови архітектури обробки вхідних сигналів у класифікаторах зображень, що відрізняються підвищеною здатністю до розпаралелювання обчислень з метою реалізації обробки в реальному масштабі часу.

Результати дисертаційної роботи були апробовані та впроваджені:

– у програмному забезпеченні дрона-міношукача, розробленого Національним університетом «Одеська політехніка» за грантом № 131/0161 Національного фонду досліджень України «Використання штучного інтелекту для підвищення якості визначення вибухових пристроїв мультисенсорним дроном-міношукачем» (довідка впровадження № 1009/64-07 від 11.06.2026);

– у теоретичних розробках перспективних архітектур обробки сигналів в межах виконання науково-дослідної роботи №723-43 «Розробка інтелектуальних сенсорів дистанційного комплексного моніторингу деградації ґрунтів повоєнних територій» (довідка впровадження № 1009/64-07 від 11.06.2026);

– у навчальний процес Національного університету «Одеська політехніка» (довідка впровадження № 1008/64-07 від 11.06.2026).

Впровадження розроблених моделей і методів забезпечує підвищення ефективності ансамблевої класифікації зі стекінгом за рахунок доповнення структури першого ступеня класифікатора клонами базових класифікаторів і синтезу другого ступеня ансамблевого класифікатора, що дозволяє підвищувати якість класифікації за мінімально можливого збільшення кількості навчальних вагових коефіцієнтів.

Запропоновані моделі та методи дають змогу здійснювати синтез архітектури ансамблевого класифікатора відповідно до особливостей вхідних даних.

Особистий внесок здобувача.

Автор самостійно отримав наукові результати, що виносяться на захист у межах дисертаційного дослідження.

У роботах, виконаних у співавторстві, здобувач представив такі нові наукові результати:

[1] – запропоновано метод синтезу клонів базових класифікаторів першого ступеня за рахунок афінного перетворення вхідних зображень, запропоновано метод порівняння структурних відмінностей базових класифікаторів і синтезованих клонів;

[2] – запропоновано метод синтезу клонів базових класифікаторів першого ступеня за рахунок зсуву випадково вибраних рядків або стовпців вхідних зображень;

[3] – запропоновано модель архітектури НМ MLP-KAN, що одночасно використовує апроксимальні властивості навчальних функцій активації та нейронів із фіксованою функцією активації;

[4] – запропоновано метод синтезу НМ для другого ступеня ансамблевого класифікатора шляхом модифікації НМ MLP-KAN відповідно до особливостей вихідних сигналів першого ступеня ансамблевого класифікатора;

[5] – запропоновано метод формалізації процесів вибору кроку навчання, інтервалу обнулення зв'язків і кількості одночасно видалених вагових коефіцієнтів під час оптимізації структури нейронної мережі MLP;

[6] – запропоновано метод вибору класифікатора, який обробляє зображення з камери дрона-міношукача залежно від обчислювальної потужності процесорів на борту дрона та якості класифікації наявності мін на зображеннях.

Апробація матеріалів дисертації.

Основні положення дисертаційної роботи були представлені на конференціях:

– 2024 IEEE 7th International Conference on ACTUAL PROBLEMS OF UNMANNED AERIAL VEHICLES DEVELOPMENTS (APUAVD). October 22-24, 2024 Kyiv, Ukraine [7];

– 2025 IEEE 8th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC). October 21-24, 2025 Kyiv, Ukraine [8].

Публікації.

Результати дисертаційних досліджень опубліковано у 6 наукових працях, які включають 6 статей (з них 2 - у фаховому періодичному науковому журналі України з технічних наук категорії «Б», 4 статті - у періодичних наукових журналах України з технічних наук, що індексуються в наукометричній базі Scopus) та 2 тези доповідей, що були представлені в матеріалах наукових конференцій, які індексуються в наукометричній базі Scopus.

Структура та обсяг дисертації.

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел, який містить 114 найменувань на 12 сторінках, і трьох додатків на 6 сторінках. Загальний обсяг дисертації становить 159 сторінок, зокрема 144 сторінок основного тексту, 11 сторінок анотацій, 55 ілюстрацій і 26 таблиць.

1 АНАЛІЗ СТАНУ АНСАМБЛЕВОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ СТЕКІНГУ НЕЙРОННИХ МЕРЕЖ

Представлені в цьому розділі результати базуються на працях [1-8].

1.1 Аналіз методів побудови нейромережевих ансамблевих класифікаторів зображень в умовах низької роздільної здатності та дефіциту навчальних даних

Незважаючи на те, що основним напрямом розвитку НМ на сьогодні є розробка надвеликих НМ, які навчаються на дуже великих наборах даних, більшість нових застосувань ще не мають таких великих наборів даних для навчання. Тому актуальним є розробка класифікаторів на основі нейронних мереж, які навчаються в умовах низької роздільної здатності зображень та дефіциту навчальних даних. і відзначаються високим співвідношенням якості класифікації до необхідного обсягу обчислень. Одним із прикладів таких додатків є дрони-міношукачі, оснащені камерами, зображення з яких процесори дрону обробляють у реальному масштабі часу [13].

У межах цього підходу широкого поширення набули ансамблеві класифікатори. Вони відрізняються використанням найбільших архітектурних компонентів – цілих НМ, які комбінуються тим чи іншим способом.

Для забезпечення високої результативної якості класифікації зображень, які відносяться до неструктурованих даних [14], необхідно як покращувати параметри, використовуваних в ансамблевих класифікаторах НМ, так і вдосконалювати методи комплексування вихідних даних окремих НМ.

Процес навчання ансамблевих класифікаторів зображень зі стекінгом НМ, як правило, складається з двох або більше етапів. На першому етапі навчаються базові класифікатори. На другому етапі навчається метакласифікатор, на вхід якого надходять результати класифікації базових класифікаторів.

Одним із найпростіших різновидів ансамблів такого типу є беггінг (англ. Bagging – від Bootstrap aggregation) [15]. Основним недоліком беггінгу є те, що процедура бутстрепа порушує балансування класів у похідних наборах вхідних даних. Це призводить до того, що первинні класифікатори налаштовані по-різному. Однак цього може виявитися недостатньо для суттєвого підвищення якості класифікації ансамблевим класифікатором порівняно з якістю класифікації первинних класифікаторів. Ефективність беггінгу може проявлятися у випадку некорельованості помилок, яких припускаються різні первинні класифікатори, проте на практиці це трапляється не часто, що й призводить до незначного збільшення підсумкової якості класифікації.

Широкого поширення на практиці набув механізм підвищення градієнта (англ. Gradient Boosting Machine, GBM) [16]. Алгоритм передбачає послідовне об'єднання простих слабких моделей, де кожна наступна модель виправляє помилки попередньої. Підхід до класифікації зображень на основі GBM демонструє вищі результати під час обробки структурованих і табличних даних порівняно з логістичною регресією, випадковим лісом і нейронними мережами, а також потребує менше даних для навчання. Вибір необхідної кількості використовуваних моделей зазвичай здійснюється експериментальним шляхом для заданої множини вхідних зображень. Крім того, сучасні реалізації (XGBoost [17], LightGBM [18]) мають вбудовані механізми регуляризації для запобігання перенавчанню. Недоліком градієнтного бустингу є низька ефективність роботи з неструктурованими даними (зображеннями та текстом). НМ значно краще справляються з такими типами даних.

Альтернативою градієнтному бустингу для обробки зображень є ансамблевий класифікатор зі стекінгом нейронних мереж [19].

1.2 Формальне визначення задачі ансамблевої класифікації зображень зі стекінгом нейронних мереж

Ансамблевий класифікатор зі стекінгом НМ містить два ступені (рис. 1.1). На першому ступені вхідне зображення обробляють НМ. Векторні оцінки ймовірності належності зображення кожному з можливих класів надходять на другий ступінь. НМ на другому ступені аналізує повні вектори з виходів НМ першого ступеня та формує результуючий вектор оцінок. Максимальне значення компоненти в цьому векторі вказує на найбільш ймовірний клас, до якого належить вхідне зображення.

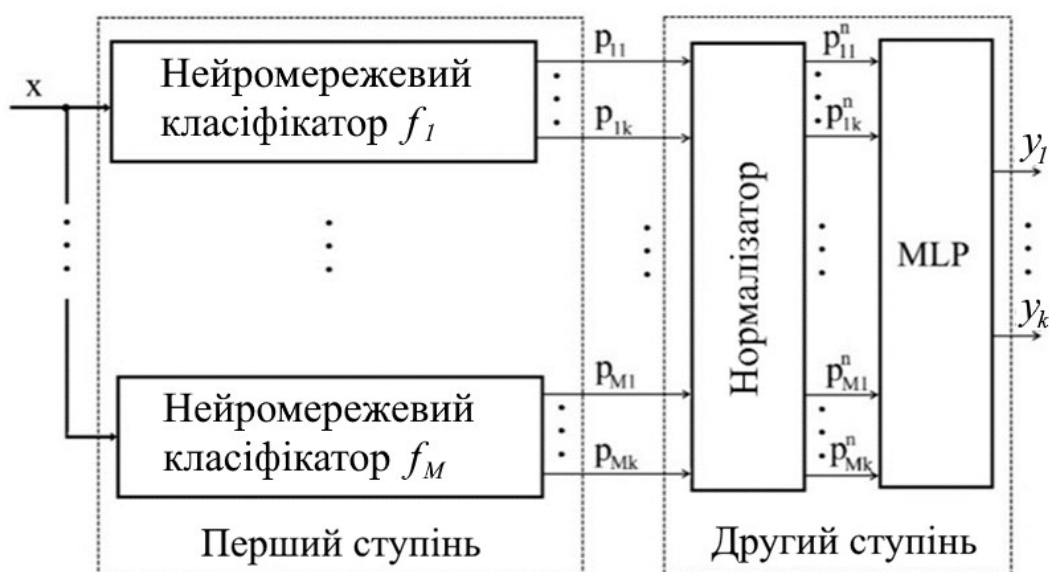


Рисунок 1.1 – Базова модель ансамблевого класифікатора зі стекінгом

Ансамблева класифікація на основі стекінгу – це ієрархічна схема навчання, в якій набір базових моделей (перший ступінь) формує новий ознаковий простір – $\mathbf{P}$, а мета-модель (другий ступінь) навчається відображати цей простір у підсумкові мітки класів – $\mathbf{Y}$ [14].

Кожна з M різних НМ першого ступеня $\{f_1, f_2, \dots, f_M\}$ для кожного вхідного зображення $x \in X$ видає вектор імовірностей розподілу за K класами, тобто вектор імовірності належності зображення x одному з K класів:

$$P_m = f_m(x; \theta_m) = [p_{m1}^{\square}, p_{m2}^{\square}, \dots, p_{mK}^{\square}], \quad (1.1)$$

де $p_{mk}^{\square} \in [0,1]$ і $\sum_{k=1}^K p_{mk}^{\square} = 1$;

θ_m – вектор вагових коефіцієнтів m -ої НМ першого ступеня.

Результати роботи всіх базових моделей першого рівня об'єднуються у вектор мета-ознак (конкатенація векторів імовірностей):

$$\mathbf{P} = \text{concat}(P_1, P_2, \dots, P_M) \in R^C, \quad (1.2)$$

де $C=M \times K$.

Таким чином, вхідне зображення x трансформується у вектор $\mathbf{P}$, який містить оцінки кожної НМ, а саме, ймовірності належності зображення до кожного з K класів.

На другому ступені мета-класифікатор $\mathbf{H}$, параметризований вагами ϕ , навчається вирішувати задачу мінімізації емпіричного ризику на мета-даних. Його вихідний вектор

$$\hat{y} = \mathbf{H}(\mathbf{P}, \phi). \quad (1.3)$$

Підсумкове вирішальне правило ансамблевого класифікатора зі стекінгом НМ має такий вигляд:

$$F(x) = \operatorname{argmax}_{k \in [1, \dots, K]} (\mathbf{H}(\text{concat}(P_1, \dots, P_M); \phi)). \quad (1.4)$$

Слід відмітити, що ефективність ансамблевої класифікації зі стекінгом НМ безпосередньо залежить від кореляції помилок базових моделей [14].

Основні переваги ансамблевих класифікаторів зі стекінгом полягають у такому:

- НМ на першому ступені ансамблевого класифікатора обробляють повні набори вхідних зображень для навчання, що дає змогу забезпечити їх збалансованість;

- другий ступінь ансамблевого класифікатора обробляє повні вихідні вектори першого ступеня, на відміну від багатьох інших підходів, де обробляються лише максимальні компоненти цих векторів; це дає змогу виловити та врахувати додаткову інформацію зі співвідношень усіх компонент вихідних векторів першого ступеня;

- природна розпаралелюваність обчислювальних процесів, що дозволяє ефективно реалізувати цю архітектуру на багатоядерних, багатопроекторних обчислювальних пристроях, а також у хмарі графічних сопроцесорах;

- суттєве зменшення впливу проблеми зникнення градієнта, оскільки нарощування можливостей моделі здійснюється «у ширину», а не «в глибину»;

- можливість навчання в умовах низької роздільної здатності зображень та дефіциту навчальних даних, оскільки кожен класифікатор першого ступеня має значно меншу кількість коефіцієнтів, що налаштовуються, порівняно із загальною сумою вагових коефіцієнтів усього ансамблевого класифікатора, та налаштовується окремо від інших класифікаторів.

Розглянуті переваги моделі ансамблевого класифікатора зі стекінгом дозволяють зробити висновок про доцільність і перспективність вибору даної моделі в якості базової для вирішення задачі класифікації зображень і запропонувати шляхи його подальшого розвитку.

В результаті аналізу джерел виділено також основні проблеми, що виникають під час реалізації ансамблевих класифікаторів зі стекінгом [19]:

- класифікатори першого ступеня повинні здійснювати структуровану обробку (відмінну від однорідної), різну для кожного класифікатора, що забезпечує появу різних помилок на різних зображеннях. У свою чергу, це дозволяє використовувати такі відмінності на другому ступені моделі для підвищення підсумкової якості класифікації. Складність полягає в тому, що кількість таких

відомих класифікаторів є обмеженою, тому необхідно розробляти нові моделі та методи збільшення кількості класифікаторів на першому ступені;

- підсумкова якість класифікації та необхідний обсяг обчислень усього ансамблевого класифікатора залежать від співвідношення між якістю класифікації та обсягом обчислень класифікаторів першого ступеня, що потребує розроблення нових моделей класифікаторів першого ступеня для забезпечення високої якості класифікації зображень за відносно невеликого обсягу обчислень;

- НМ, яка здійснює обробку на другому ступені ансамблевого класифікатора, також робить свій внесок у підсумкові параметри ефективності, тому необхідно також підвищувати якість її класифікації та зменшувати обсяг обчислень.

1.3 Аналіз відомих моделей нейронних мереж першого ступеня ансамблевого класифікатора зі стекінгом

Для забезпечення ансамблевим класифікатором зі стекінгом високої якості класифікації в умовах низької роздільної здатності зображень та дефіциту навчальних даних, на першому ступені мають бути використані базові класифікатори, що забезпечують:

- високу якість класифікації;
- невеликий обсяг обчислень;
- структуровану обробку вхідних зображень.

Остання вимога стосується фундаментального принципу побудови ефективних ансамблів зі стекінгом НМ – принципу диверсифікації (різноманітності) помилок [14]. З практичного досвіду відомо, що якщо всім класифікаторам дати однакові дані (однаковий набір ознак) і використати схожі алгоритми навчання, вони стануть «однорідними». Це призведе до того, що класифікатори будуть помилятися на одних і тих самих прикладах. Для ансамблевих класифікаторів зі стекінгом це не прийнятно, бо мета-класифікатор не зможе виправити помилку, яку зробили всі «одноставно». Тобто, вимога

структурованої обробки вхідного набору зображень передбачає, що вона повинна бути різною для кожного з класифікаторів і забезпечувати різноманітність помилок у класифікаторів.

1.3.1 Згорткові нейронні мережі

Згорткові нейронні мережі (англ. Convolutional Neural Net, CNN) були запропоновані в роботі [20] і являють собою послідовно з'єднані згорткові шари, нелінійні шари, шари об'єднання та на виході – повнозв'язані шари.

У згорткових шарах використовуються навчальні ядра, з якими виконується операція згортки. Вони виділяють ознаки на зображеннях, що дають змогу розпізнавати об'єкти. У першому згортковому шарі розпізнаються найдрібніші ознаки, у другому – більш великі тощо.

Повнозв'язані шари формують результуючий вектор імовірностей належності зображення кожному з класів. Приклад CNN для розпізнавання рукописних цифр із набору MNIST [21] наведено на рис. 1.2 [22].

Необхідно зазначити, що при індивідуальному використанні CNN демонструє високу якість класифікації, однак обробка в CNN є неструктурованою, тобто всі елементи зображення обробляються однаково. Тому її доцільно включати до складу першого ступеня ансамблевого класифікатора лише за умови, що всі інші класифікатори матимуть структуровану обробку. В іншому випадку можна очікувати не підвищення, а погіршення підсумкової якості класифікації.

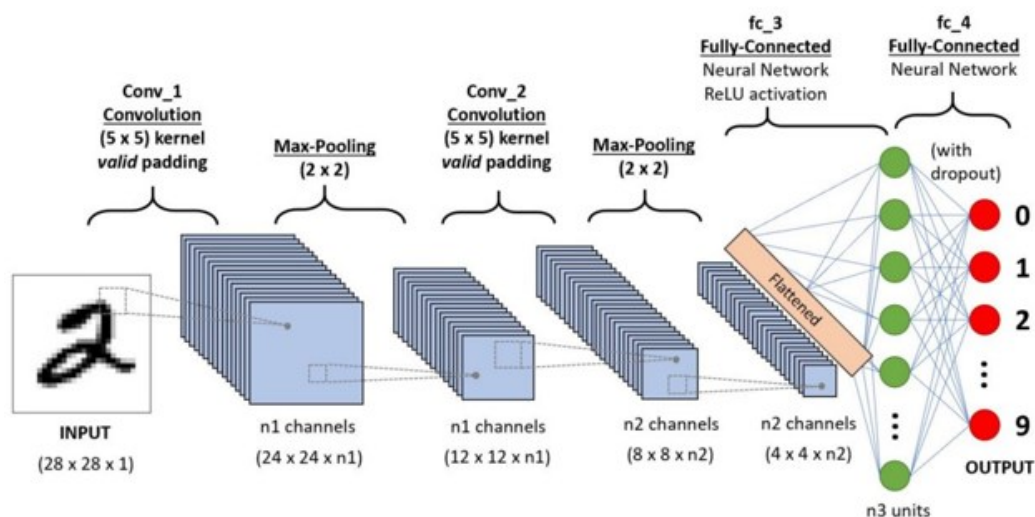


Рисунок 1.2 – Модель CNN для розпізнавання рукописних цифр (джерело: [22])

1.3.2 Класифікатори на базі трансформерів

В основі переважної більшості сучасних нейронних мереж лежить модель Transformer [23]. Було показано, що такі НМ демонструють суттєво вищу якість класифікації зображень порівняно з різними модифікаціями CNN та іншими моделями. Однак співвідношення між якістю класифікації та обсягом обчислень у них поступається ансамблевому класифікатору зі стекінгом, який використовує типовий набір класифікаторів на першому ступені [19]. Це пояснюється тим, що модель трансформера спочатку була розроблена для обробки текстових даних, а пряме застосування моделі трансформера для обробки зображень (англ. Visual Transformer, ViT) характеризується надмірним обсягом обчислень [23]. В результаті були розроблені модифікації трансформерів, адаптовані спеціально для обробки зображень.

У статті [24] запропоновано ViT-Lite – компактну версію ViT, яка перед трансформером містить блоки розбиття, згортку та шар зменшення розмірності. Далі до моделі було введено нову операцію Sequence Pooling, що об'єднує послідовності даних усередині трансформера. Це дозволило поєднати релевантну інформацію, яка міститься в різних частинах вхідного зображення. Усі ці зміни забезпечили компактність моделі, що отримала назву компактний згортковий трансформер (англ. Compact Convolution Transformer, CCT) (рис. 1.3).

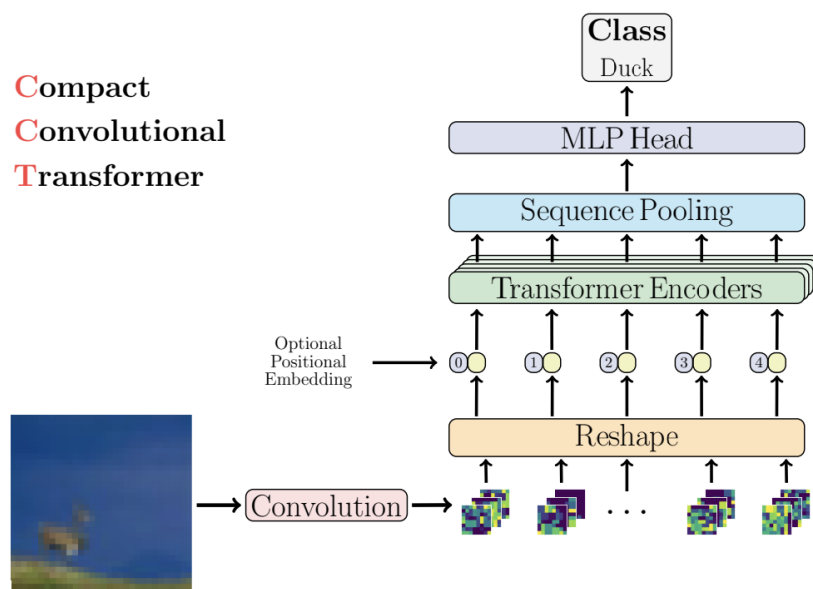


Рисунок 1.3 – Модель CCT (джерело: [24])

У [25] модифікацію трансформера здійснено за рахунок заміни механізму самоуваги на механізм зовнішньої уваги (англ. External Attention Transformer, EATNet).

Механізми уваги відіграють винятково важливу роль в архітектурі трансформерів. Він реалізує обчислення довгострокових залежностей у межах однієї вибірки на основі попарних зв'язків між усіма позиціями. Це одна з найбільш обчислювально витратних компонент моделі трансформера, яка має квадратичну складність. У [25] запропоновано механізм зовнішньої уваги, що має лінійну складність. Багаточисленні експерименти підтвердили, що під час обробки зображень механізм зовнішньої уваги забезпечує такі ж або навіть вищі результати, порівняно з використанням механізму самоуваги. На рис. 1.4 показано відмінності моделей з самоувагою, спрощеною самоувагою та зовнішньою увагою. На рис. 1.5 наведено різницю в реалізації багатоголової моделі трансформера.

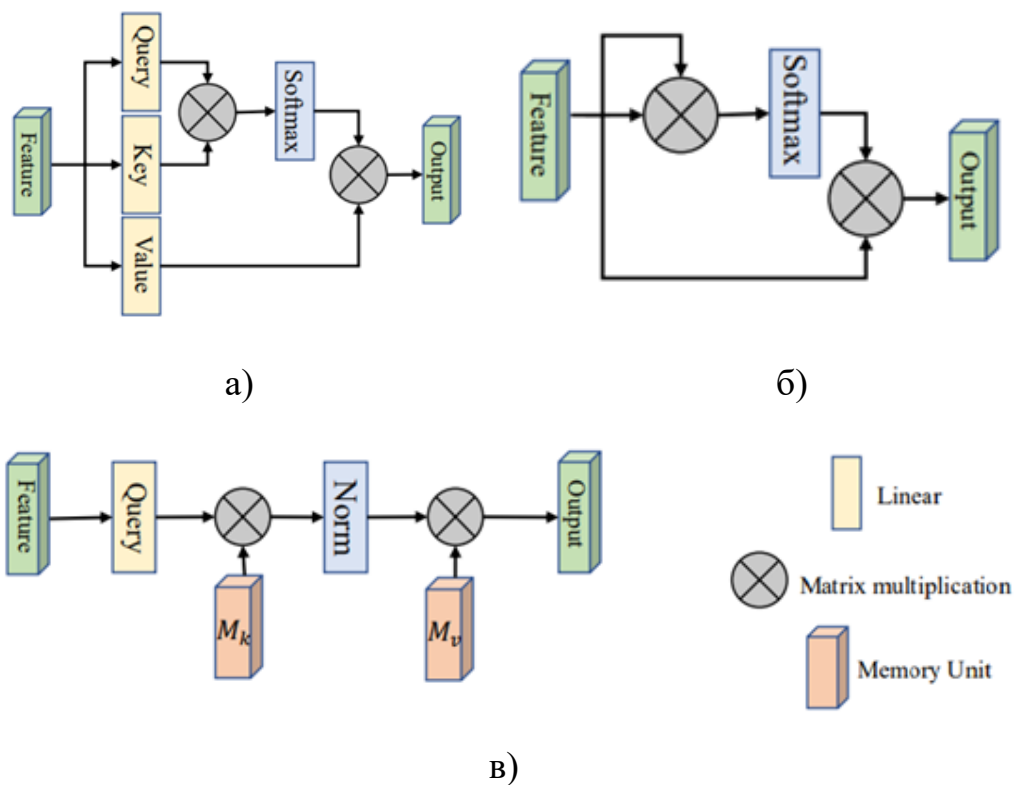
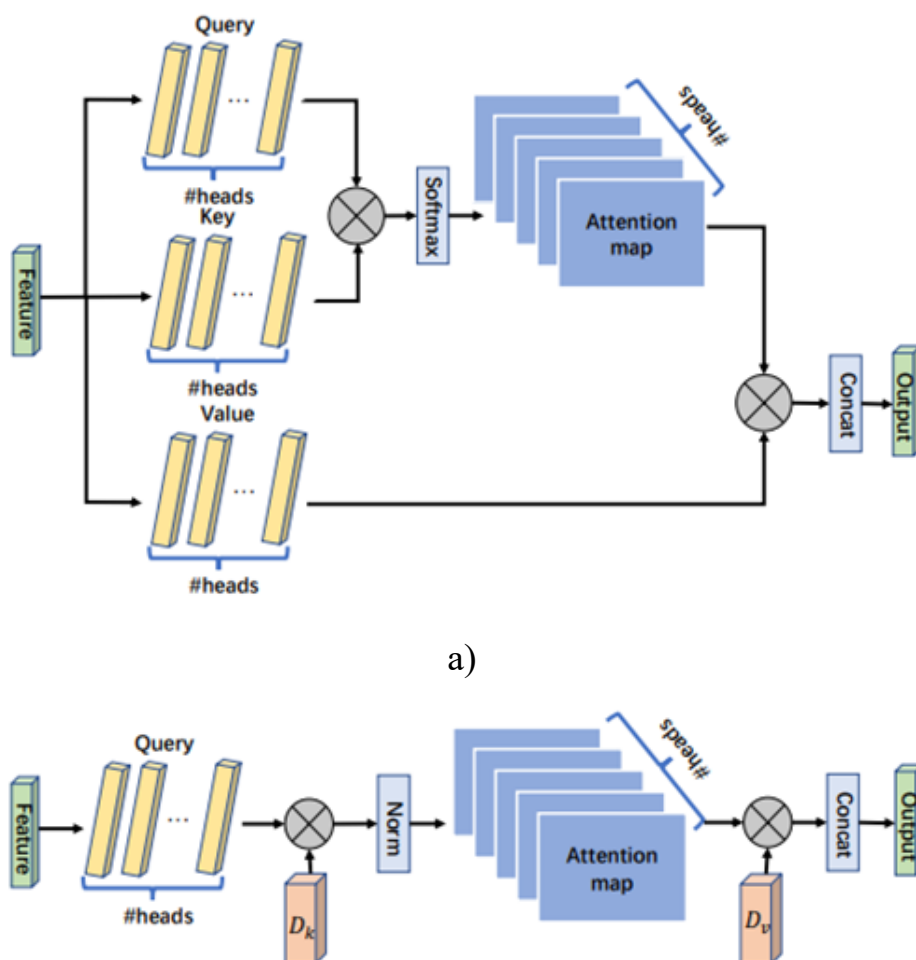


Рисунок 1.4 – Приклади моделей трансформерів з механізмами: а – самоуваги; б – спрощеної самоуваги; в – зовнішньої уваги (джерело: [25])



б)

Рисунок 1.5 – Приклади реалізації багатоголових моделей з механізмами:

а – самоуваги; б – зовнішньої уваги (джерело: [25])

В роботі [26] запропонована модифікація енкодера трансформера – заміна механізму внутрішньої уваги стандартним перетворенням Фур'є, яке здійснює змішування вхідних токенів (англ. Fourier Network, FNet).

Під час реалізації перетворення Фур'є використовується лише дійсна частина результату, щоб уникнути обробки комплексних чисел у наступних шарах. Використання перетворення Фур'є в цій моделі забезпечує наступним шаром достатній доступ до всіх токенів зображення.

Оскільки множення в частотній області еквівалентне згортці у часовій області, чергування в послідовних блоках енкодера прямого та оберненого перетворення Фур'є являє собою чергування множення і згортки в цих блоках. Модель нейронної мережі FNet показана на рисунку 1.6.

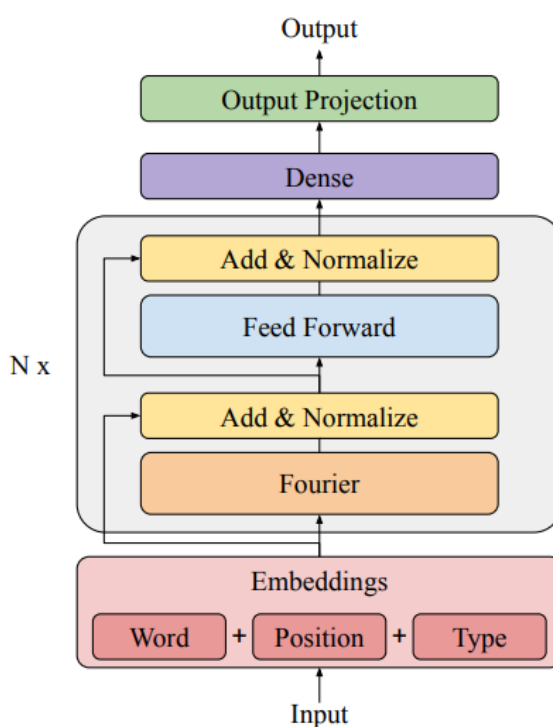
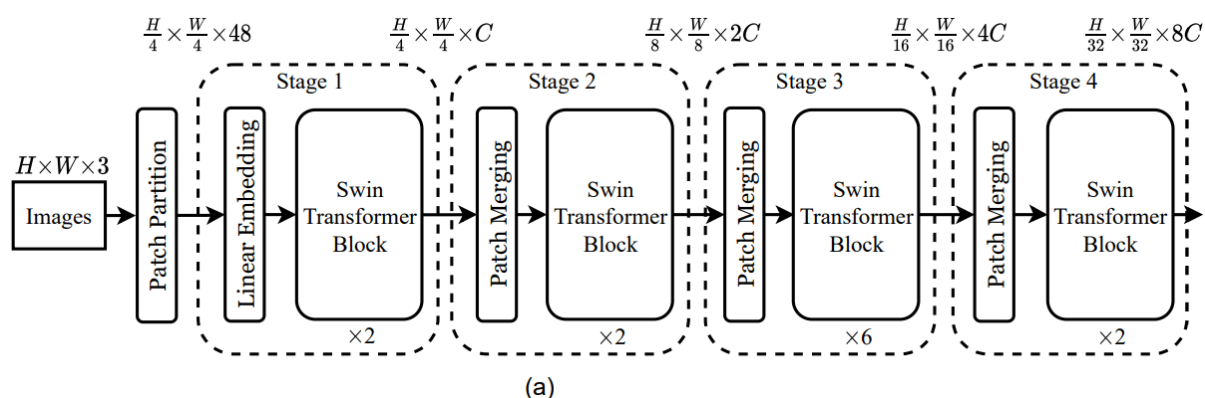


Рисунок 1.6 – Модель нейронної мережі FNet (джерело: [26])

Однією з основних проблем адаптації моделей трансформерів, спочатку розроблених для обробки текстових даних до обробки зображень, є суттєва різниця в масштабі візуальних компонентів зображень. Це є наслідком того, що класифікація зображень дуже часто є наступним етапом після виявлення об'єктів на зображеннях [27-29], отриманих на різних відстанях від об'єктів, які потрібно виявити та класифікувати. У базових моделях трансформерів [23, 30] усі токени мають фіксований масштаб і слова в текстових фрагментах мають значно меншу роздільну здатність порівняно з роздільною здатністю пікселів на зображеннях. При цьому обсяг обчислень для самоуваги пропорційний квадрату розміру зображення.

У роботі [31] запропоновано модифікацію трансформера Swin Transformer (SwinTr), яка створює ієрархічні карти ознак і має лінійну обчислювальну складність залежно від розміру зображення. Побудова карт починається з невеликих фрагментів, які потім поступово об'єднуються у масштабніші трансформовані шари. Завдяки такій побудові, модель SwinTr добре поєднується з архітектурними підходами щільної обробки пікселів, такими як пірамідальні мережі ознак (англ. Feature Pyramid Networks, FPN) [27] і U-Net [32]. Самоувага обчислюється всередині неперетинних вікон, кожне з яких має фіксовану кількість патчів (фрагментів зображення). Розмір вікон послідовно збільшується від блоку до блоку. Така структура обробки забезпечує лінійну обчислювальну складність SwinTr. Модель мережі SwinTr показана на рис. 1.7.



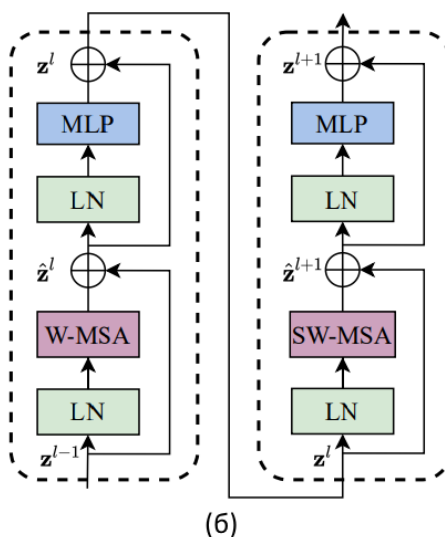


Рисунок 1.7 – Модель мережі: а – SwinTr; б – два послідовні блоки SwinTr, де W-MSA і SW-MSA – це багатоголові модулі самоконтролю зі звичайними та зміщеними конфігураціями керування вікнами відповідно (джерело: [31])

Спочатку вхідне кольорове зображення розбивається на патчі розміром 4×4 пікселі. У кожному наступному блоці патчі об'єднуються із сусідніми, що забезпечує поступове збільшення розміру вікон при збереженні сталої кількості патчів в одному вікні.

Як видно з рис. 1.7, а, замість модуля самоконтролю в SwinTr використовуються зміщені вікна. W-MSA і SW-MSA – це багатоголові модулі самоконтролю зі звичайними та зміщеними конфігураціями керування вікнами [31]. Блок SwinTrr (рис. 1.7, б) складається з модуля W-MSA, побудованого на основі звичайного вікна, за яким слідує двошаровий MLP. Перед MSA та MLP виконується нормалізація даних за допомогою шарів Layer Normalization (LN). Друга частина блоку побудована аналогічно, але використовує зміщене вікно, що забезпечує врахування всіх взаємозв'язків між елементами зображення.

1.3.3 Класифікатори на базі MLP

Особливістю НМ, побудованих на моделі трансформера з механізмом уваги, є динамічна параметризація просторових взаємодій на основі вхідних представлень [33, 34]. Це природно відповідає природі текстових повідомлень. Однак для

окремих зображень, що не утворюють єдиного відеоряду, неможливо побудувати аналогічну динамічну параметризацію просторових взаємодій. Тому у таких випадках MLP, який здійснює статичну параметризацію, може показувати високі результати [35].

У роботі [33] запропоновано модель Gated Multilayer Perceptron (gMLP), яка не містить механізму уваги та згорток. Вона складається з послідовно з'єднаних однакових блоків на базі MLP, що реалізують каналні і міжканальні проєкції, доповнені мультиплікативним стробуванням. Структура цих блоків наведена на рис. 1.8.

Кожен блок містить послідовно з'єднані елементи, а саме, проєктор каналу (англ. Channel Projection), міжканальний проєктор (англ. Spatial Gating Unit) і знову проєктор каналу.

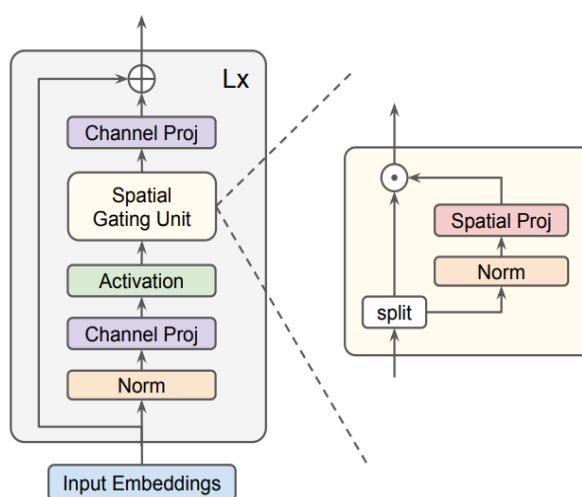


Рисунок 1.8 – Модель gMLP (джерело: [33])

Аналогічно побудована архітектура MLP-Mixer [36], яка повністю складається з класичних MLP. Однак, на відміну від звичайної обробки всього зображення, коли зображення перетворюється на довгий вектор і подається на MLP, у цьому випадку двовимірна структура зображення зберігається. У кожному блоці зображення спочатку розбивається на стовпці, кожен із яких обробляється своїм MLP. Розмір стовпців не змінюється. Результати обробки знову об'єднуються в цілісне зображення, яке розбивається на рядки. Кожен рядок, аналогічно

стовпцям, обробляється власним MLP без зміни розмірності. Після цього рядки знову об'єднуються в єдине зображення, яке подається на наступний аналогічний блок. Модель MLP-Mixer показана на рис. 1.9.

Послідовна обробка за стовпцями та рядками забезпечує врахування зв'язків між усіма елементами зображення.

Для розв'язання проблеми зникнення градієнта під час збільшення кількості блоків в архітектурі MLP-Mixer використовуються стандартні архітектурні компоненти: skip-з'єднання [37] та нормалізація шарів [38]. На виході всього класифікатора застосовується двошаровий об'єднувальний MLP.

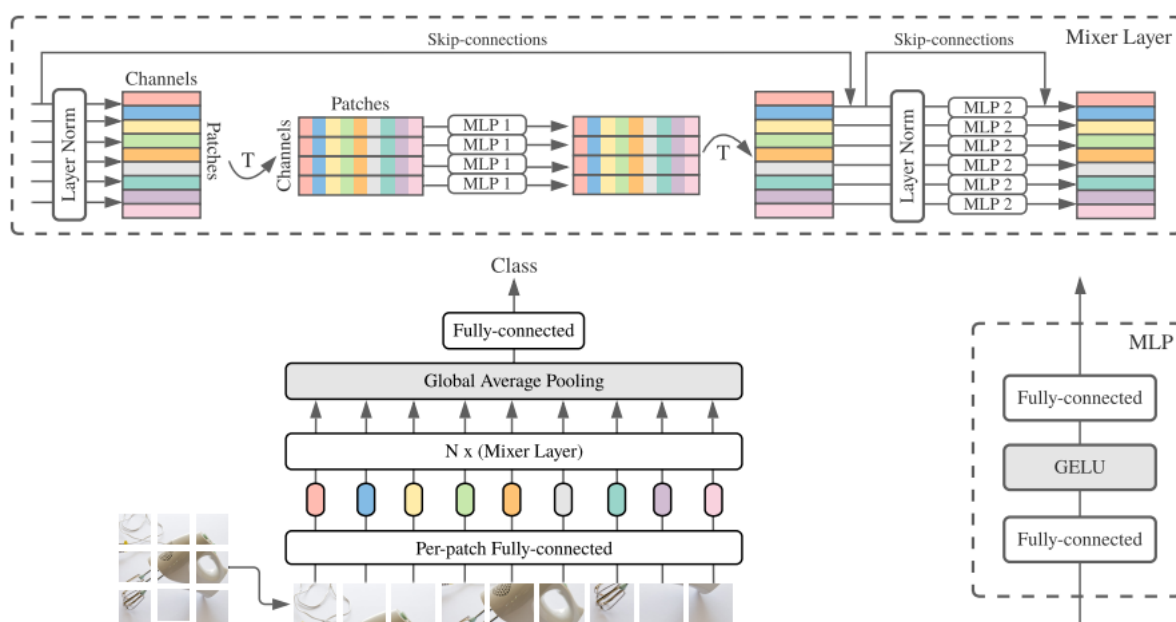


Рисунок 1.9 – Модель MLP-Mixer (джерело: [36])

1.3.4 Порівняльний аналіз характеристик розглянутих нейронних мереж

Усі НМ, що були розглянуті в підрозділі 1.3, відповідають вимогам до класифікаторів першого ступеня ансамблевого класифікатора зі стекінгом:

- мають різні алгоритми обробки вхідних зображень;
- характеризуються високим співвідношенням між якістю класифікації та необхідним обсягом обчислень;

– можуть бути навчені в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

Для проведення комп'ютерних експериментів з оцінки точності класифікації зображень в роботі використані набори даних MNIST [21] і CIFAR-10 [39].

У табл. 1.1 [19] наведено точність класифікації після 50 епох навчання та кількість навчальних вагових коефіцієнтів цих НМ для наборів даних MNIST.

Набір даних MNIST містить чорно-білі зображення рукописних цифр розміром 28×28 пікселів: 60 тис. зображень – для навчання і 10 тис. зображень – для тестування. Зображення належать до 10 класів.

Набір даних CIFAR-10 містить 50 тис. кольорових зображень для навчання та 10 тис. зображень для тестування. Кожне зображення має розмір 32×32 пікселі і належить до одного з 10 класів об'єктів (літак, пароплав, автомобіль, тощо).

Таблиця 1.1 – Порівняння характеристик нейромережевих класифікаторів

Назва нейронної мережі	MNIST			CIFAR-10		
	<i>Accuracy</i>	Кількість вагових коефіцієнтів	<i>Accuracy</i> / Кількість вагових коефіцієнтів	<i>Accuracy</i>	Кількість вагових коефіцієнтів	<i>Accuracy</i> / Кількість вагових коефіцієнтів
CNN	0.9929	34826	$2.85 \cdot 10^{-5}$	0.7687	343306	$2.24 \cdot 10^{-6}$
CCT	0.9879	406987	$2.43 \cdot 10^{-6}$	0.8021	408139	$1.97 \cdot 10^{-6}$
EATNet	0.8517	358090	$2.38 \cdot 10^{-6}$	0.6788	355530	$1.91 \cdot 10^{-6}$
FNet	0.9881	992010	$1.00 \cdot 10^{-6}$	0.7572	582410	$1.30 \cdot 10^{-6}$
gMLP	0.9876	1271818	$0.78 \cdot 10^{-6}$	0.7405	862218	$0.86 \cdot 10^{-6}$
MLP-Mixer	0.9574	629258	$1.52 \cdot 10^{-6}$	0.7674	219658	$3.49 \cdot 10^{-6}$
SwinTr	0.9711	147034	$6.60 \cdot 10^{-6}$	0.7128	151386	$4.71 \cdot 10^{-6}$

Типові приклади зображень із наборів даних MNIST і CIFAR-10 наведені у додатку А.

В даній роботі ефективність використання НМ для класифікації зображень оцінюється двома характеристиками, а саме, якістю класифікації й обсягом

обчислень. В свою чергу для оцінки якості класифікації зображень використано метрику точності *Accuracy*, яка обчислюється як частка правильно класифікованих зображень від загальної кількості зображень.

Обсяг обчислень визначається кількістю навчасних вагових коефіцієнтів НМ, який визначається в ході програмної реалізації нейронної мережі стандартними бібліотечними методами .

З табл. 1.1 видно, що найкращими характеристиками для набору MNIST володіє CNN, але для більш складного набору даних CIFAR-10 вона поступається за якістю класифікації SST.

1.4 Аналіз підходів до зменшення обсягу обчислень на другому ступені ансамблевого класифікатора

Типовим варіантом НМ на другому ступені ансамблевого класифікатора є MLP [19]. Найефективнішими підходами для зменшення обсягу обчислень є проріджування та модифікація архітектури на основі мережі Колмогорова-Арнольда [40].

1.4.1 Проріджування зв'язків у MLP

В задачах класифікації зображень MLP має вищі показники якості класифікації порівняно зі спеціальними НМ, наприклад CNN, що пояснюється його повнозв'язністю і наявністю надлишковості в моделі MLP.

Це, в свою чергу, призводить до появи шумових складових під час навчання і формування вихідних даних, що заважає досягненню кращих результатів. Тому було розроблено багато різних підходів до проріджування зв'язків у НМ, які дають змогу зменшити обсяг обчислень зі збереженням підсумкових характеристик.

Наразі можна виділити такі основні напрями:

а) передавання знань (англ. Knowledge Distillation) [41-43] – попередньо навчається модель із великою кількістю вагових коефіцієнтів, після чого на основі її структури та значень ваг навчається модель із меншою кількістю ваг. Однак

практична реалізація навчання спрощеної моделі при цьому підході є складним завданням і часто супроводжується зниженням підсумкових показників [44];

б) зниження рангу матриці вагових коефіцієнтів (англ. Low-rank Decomposition) [45, 46] – пропонує видалення ваг, що відповідають найменшим власним числам. Проте сучасні НМ містять велику кількість вагових коефіцієнтів, а знаходження власних векторів і чисел для великих матриць є обчислювально складним завданням і може супроводжуватись значними похибками, що призводить до помітного погіршення підсумкових характеристик;

в) квантування параметрів (англ. Parameter Quantization) [47, 48] – передбачає заміну обчислень із рухомою комою на обчислення з фіксованою комою та зменшеною розрядністю. Динамічна зміна розрядності дозволила отримати високі результати для наборів MNIST і CIFAR-10. Проте будь-які неточності у масштабуванні обчислень у різних вузлах неминуче призводять до суттєвого погіршення підсумкових характеристик. Такий підхід ефективно реалізований лише на програмованих логічних матрицях (англ. Field-programmable gate array, FPGA). Реалізація ж на універсальних програмованих обчислювачах призводить до надлишкового обсягу операцій і значного уповільнення;

г) проріджування мережі (англ. Network Pruning) – із НМ видаляються окремі зв'язки або цілі структури на основі функції важливості вагових коефіцієнтів.

Як функцію важливості може бути використано матрицю других похідних функції втрат [49], але її обчислення потребує великого обсягу ресурсів, тому запропоновано низку спрощених функцій важливості [50].

Найпростішим способом проріджування є ітераційне видалення найменших за модулем вагових коефіцієнтів [51]. Такий підхід потребує розроблення спеціальних методів досягнення балансу між швидкістю та тривалістю навчання, а також вибору кількості зв'язків, що видаляються за один раз.

Надто часте проріджування може призвести до того, що мережа не встигне навчитися й будуть видалені важливі ваги. Цьому ж може сприяти надмірно велика швидкість навчання через шумові процеси оптимізації.

Надто велика кількість одночасно видалених ваг може взагалі призвести до деградації мережі, а надто рідке проріджування – істотно уповільнити процес навчання.

Ще одним критичним фактором є ризик перенавчання, який також може спричинити видалення невідповідних коефіцієнтів. Звичайна регуляризація [52, 53] призводить до зсуву значень вагових коефіцієнтів i , відповідно, до зниження максимально досяжних показників. Використання Dropout [54] не викликає зсуву ваг, проте істотно сповільнює навчання.

Найперспективнішим способом регуляризації вважається використання передспотворень вхідних зображень [55-57]. У [58] показано, що одночасне застосування ітераційного проріджування та передспотворення вхідних сигналів забезпечує регуляризацію процесу навчання НМ й дозволяє навіть підвищити підсумкові характеристики.

У [59] досліджено взаємозв'язок параметрів процесів навчання, проріджування та регуляризації. Показано, що на основі аналізу взаємозв'язку цих параметрів для конкретної моделі НМ та типу представлення вхідних зображень можливо досягати вищих підсумкових результатів ефективності класифікації.

1.4.2 Використання моделі KAN

Теоретичним обґрунтуванням використання MLP в задачах класифікації є те, що він є універсальним апроксиматором [35]. Альтернативним апроксиматором [60], який також знайшов широке застосування в НМ, є класифікатор на основі мережі KAN [40].

MLP використовує як основні елементи нейрони – багатовходові суматори з навчальними ваговими коефіцієнтами на входах та незмінною нелінійністю на виході (функцією активації), яка є однаковою для всіх нейронів одного шару.

У KAN відсутні функції активації на виході суматора, і замість навчальних вагових коефіцієнтів використовуються навчальні нелінійні функції, які також називаються функціями активації.

Мінімальна кількість апроксимуючих компонент досягається тоді, коли апроксиматор відповідає апроксимованій функції. Однак у НМ зазвичай невідомий

тип апроксимованої функції, тому перспективним напрямком покращення якості класифікації є одночасне використання обох типів апроксиматорів для досягнення найкращої апроксимації з мінімальною кількістю апроксимуючих компонент.

У [40] для ряду спеціальних задач показано, що KAN потребує меншого обсягу обчислень, порівняно з MLP. У [61] розглядалися задачі навчання з підкріпленням та було отримано значно менший обсяг обчислень при зіставній якості навчання порівняно з MLP. У [62] було досліджено широкий спектр задач машинного навчання, комп'ютерного зору, обробки тексту та звуку.

Для всіх досліджених задач використання MLP забезпечувало вищі результати при меншому обсязі обчислень. Виняток становили лише задачі із символічним поданням математичних формул. Результати цих робіт підтверджують, що для різних типів задач найкраще підходить або MLP, або KAN. Проте для задач складного типу, до яких належать задачі класифікації об'єктів на зображеннях, одночасне використання MLP і KAN в архітектурі трансформера [63] дозволило отримати вищі результати, ніж при використанні лише MLP або лише KAN.

У роботі [64] розглядалося завдання прогнозування часових рядів. Запропонована двошарова архітектура з модифікацією KAN використовувала сплайни для аналізу часових закономірностей між послідовними відліками даних. Обсяг обчислень виявився суттєво меншим, ніж у двошарового MLP, однак навчання цієї модифікації KAN тривало значно довше. У зв'язку з цим у [64] рекомендовано для застосувань із швидко змінними часовими залежностями використовувати чистий MLP. У [65] у рекурентній нейронній мережі (англ. Recurrent Neural Network) вихідні шари MLP були замінені на KAN. Це дозволило в задачі прогнозування сонячної радіації та температури повітря підвищити якість прогнозування й зменшити обсяг обчислень. У шарах KAN нелінійні функції апроксимувалися сплайнами.

У роботі [66] запропоновано гібридну модель, що використовує тривимірні, двовимірні та одновимірні шари KAN для задачі класифікації гіперспектральних зображень. Це дозволило отримати співставні характеристики за меншого обсягу

обчислень у порівнянні з результатами в [67], де застосовувалися згорткові нейронні мережі і трансформери. У [68] для цього ж застосування використано дифузійну нейронну мережу, яка використовувала KAN. Експериментальні результати показали покращення якості зображень під час визначення місць посадки та розвідки ресурсів у дослідженнях далекого космосу. Однак слід зазначити, що моделі, запропоновані в роботах [66] і [68], є надлишковими для задачі класифікації об'єктів на зображеннях.

У роботі [69] KAN-шар було використано на виході системи YOLOv.10. Це дозволило підвищити точність позиціювання рухомих об'єктів і зменшити обсяг обчислень під час обробки інфрачервоних знімків, у порівнянні з використанням MLP у тому самому застосуванні [69]. У роботі [70] застосовано аналогічний підхід – у CNN вихідний MLP було повністю замінено на тришаровий KAN. Модифікована CNN показала вищі результати при меншому обсязі обчислень, однак експерименти проводилися зі специфічними сигналами, які добре апроксимуються сплайнами. Експерименти для вхідних сигналів загального типу у [70] не виконувалися. У [71] CNN була модифікована шляхом вбудовування шарів KAN безпосередньо в згорткову частину НМ, тоді як на виході MLP залишився незмінним. У результаті при обробці набору даних Fashion MNIST [72] обсяг обчислень зменшився, але якість класифікації була нижчою на 1,3%, ніж у звичайної НМ. У [73] на основі KAN були розроблені нові шари – Bottleneck Convolutional Kolmogorov-Arnold та Self KAGNtention, які дозволили одночасно покращити якість класифікації та зменшити обсяг обчислень. Однак це рішення було створене спеціально для згорткових мереж і в незмінному вигляді не може бути використане в інших архітектурах НМ. У [74] запропоновано модифікацію мережі U-Net, у якій використовувалися комбінації KAN-шарів і згорткових шарів на виходах кодувальної та декодувальної частин. Експериментальні дослідження проводилися в задачі супутникового моніторингу повеней шляхом обробки радіолокаційних даних із синтезованою апертурою. Було отримано покращення підсумкових характеристик при меншому обсязі обчислень.

Таким чином, у більшості робіт пропонується або замінити MLP на KAN, або інтегрувати KAN у складну структуру моделі. Суттєве покращення характеристик спостерігається лише в тих застосуваннях, де оброблювані дані добре апроксимуються за допомогою KAN. Це, природно, супроводжується зменшенням обсягу обчислень. Тому необхідними є подальші розробки та дослідження моделей НМ, що поєднують апроксимальні властивості MLP і KAN у задачі класифікації об'єктів на зображеннях.

Під час реалізації KAN істотне значення має спосіб реалізації навчальних нелінійних функцій (функцій активації). Він впливає як на підсумкову ефективність, так і на обсяг обчислень. У базовій роботі з KAN [40] використовуються сплайни. Для цього була розроблена спеціальна бібліотека, яка, однак, погано інтегрується з такими популярними фреймворками, як PyTorch та Keras. Під час практичної реалізації навчальних функцій активації зручніше користуватися деяким рядом функцій і навчати вагові коефіцієнти, з якими вони підсумовуються для отримання потрібного вигляду функцій активації.

У роботі [75] на прикладі задачі класифікації об'єктів на зображеннях із набору MNIST було досліджено 18 рядів функцій для апроксимації навчальних функцій активації. Порівнювалася якість класифікації MLP і KAN. Усі варіанти KAN показали свою працездатність, однак якість класифікації була нижчою, ніж при використанні MLP. Це пояснюється тим, що набір даних MNIST має загальний характер, тому жоден із рядів функцій, найбільш придатних для спеціалізованих даних, не забезпечив покращення якості класифікації. У [76] було показано, що одним із найкращих варіантів заміни сплайнів є базисні радіальні функції (англ. Radial basis function, RBF), які мають універсальні апроксимаційні властивості. У [77] отримано високі результати в задачі апроксимації функцій за допомогою НМ, що використовують RBF. Ширина і центри використовуваних RBF визначалися за допомогою додаткового алгоритму кластеризації. У [78] покращення характеристик НМ з RBF було досягнуто завдяки додатковій обробці вхідних сигналів експоненціальними функціями. У [79] у задачі розпізнавання образів із нерівномірним (негаусівським) розподілом шумів використовувалися НМ з RBF та

сурогатна функція правдоподібності. Експериментальні дослідження показали покращення якості класифікації об'єктів на зображенні. У роботі [80] було запропоновано аналог моделі KAN, однак повноцінної апроксимації функцій активації RBF не проводилося. Замість цього апроксимація здійснювалася шляхом додавання або видалення вузлів у прихованих шарах залежно від їхнього внеску в підсумкові характеристики.

Таким чином, аналіз літератури показав, що проста заміна MLP на KAN забезпечує покращення характеристик лише у спеціальних випадках, тоді як спільне використання моделей НМ, що поєднують нейрони та шари KAN, дозволяє підвищити якість класифікації об'єктів на зображеннях із наборів даних загального типу. Однак відомі моделі містять складні додаткові вузли, що ускладнює їхню практичну реалізацію. З погляду зручності реалізації, перспективним є використання RBF для апроксимації навчальних функцій активації. Це має забезпечити застосування регулярних компонентів моделі, які можуть бути легко реалізовані в межах уже відомих підходів. Крім цього, аналіз методу проріджування MLP, запропонованого в [59] показав, що комбінування MLP і KAN є більш перспективним з точки зору зменшення обсягу обчислень.

1.5 Постановка завдання дослідження

Проведений аналіз методів класифікації об'єктів на зображеннях дозволяє зробити висновок про обмеженість результатів існуючих підходів і наявність перспективних напрямів, які можуть забезпечити покращення якості класифікації за відносно невеликого збільшення обсягу обчислень та з можливістю навчання в умовах низької роздільної здатності зображень та дефіциту навчальних даних.

Як базова модель класифікатора, що забезпечує найвищі підсумкові характеристики, а саме точність класифікації та кількість вагових коефіцієнтів, визначено ансамблевий класифікатор зі стекінгом НМ. Проте для нього існує низка обмежень, подолання яких вимагає розроблення нових моделей та методів.

Найсуттєвішим обмеженням є вимога, щоб однакові базові класифікатори першого ступеня забезпечували структуровану обробку, відмінну для різних класифікаторів. Кількість відомих НМ, які задовольняють цю вимогу та мають високе співвідношення якості класифікації й обсягу обчислень, є дуже обмеженою. Одним із критичних параметрів таких систем є доступна обчислювальна потужність, тому для них розроблення моделей із зменшеним обсягом обчислень за збереження якості класифікації об'єктів є одним із пріоритетних завдань.

Актуальність проблеми дослідження полягає у необхідності створення ансамблевих класифікаторів зображень із підвищеним співвідношенням якості класифікації й обсягу обчислень. Прикладом практичного застосування таких класифікаторів може бути їх використання в автономних робототехнічних комплексах, що працюють у реальному масштабі часу. Одним із прикладів таких систем є дрон, який автономно обстежує територію з метою виявлення мін.

У зв'язку з цим, для підвищення ефективності ансамблевих класифікаторів актуальним науковим завданням є:

- розвиток нових моделей і методів, що дозволяють збільшувати кількість НМ на першому ступені в ансамблевих класифікаторах зі стекінгом для забезпечення високої якості класифікації зображень;
- розвиток нових методів і моделей для об'єднання результатів первинних НМ на другому ступені ансамблевого класифікатора зі стекінгом для забезпечення зменшення обсягу обчислень.

Висновки до розділу 1

1. Основною проблемою ансамблевих класифікаторів зі стекінгом є забезпечення на першому ступені достатньої кількості класифікаторів, які б відрізнялися між собою структурованістю обробки вхідних зображень і мали високе співвідношення якості класифікації та обсягу обчислень.

2. Було проведено аналіз відомих НМ, які можуть бути використані на першому ступені ансамблевого класифікатора зі стекінгом НМ. Виконано

порівняння їхніх моделей, якості класифікації та обсягу обчислень. Виявлено, що для наборів даних, подібних до MNIST, найвищу якість класифікації та найменший обсяг обчислень забезпечує CNN, а для наборів даних, подібних до CIFAR-10, найкращим вибором за цими параметрами є ССТ.

3. Проведений аналіз підходів, які є перспективними для зменшення обсягу обчислень на другому ступені ансамблевого класифікатора зі стекінгом НМ, дозволив виділити два основні напрями – використання НМ типу KAN та ітераційне проріджування. Проста заміна MLP на KAN у загальному випадку призводить до погіршення підсумкових характеристик, тоді як комбінації елементів MLP і KAN дозволяють досягти покращення результатів. Аналіз літератури показав, що комбінування MLP і KAN є більш перспективним з точки зору зменшення обсягу обчислень ніж проріджування.

4. Аналіз способів практичної реалізації шарів KAN показав, що найперспективнішим підходом є апроксимація навчальних функцій активації рядами функцій. Одним із найвдаліших варіантів вибору таких функцій є RBF. Однак відомі моделі НМ, що використовують RBF, або розроблені для спеціалізованих застосувань, або містять додаткові вузли, які ускладнюють практичну реалізацію. Отже, необхідно розробити та дослідити спеціальні моделі НМ, які поєднують MLP і KAN та ефективно інтегрують апроксимуючі властивості обох підходів для підвищення якості класифікації зображень.

2 РОЗРОБКА МОДЕЛІ ПЕРШОГО СТУПЕНЯ АНСАМБЛЕВОГО КЛАСИФІКАТОРА ЗІ СТЕКІНГОМ НА ОСНОВІ КЛОНІВ НЕЙРОННИХ МЕРЕЖ ТА МЕТОДУ ВИЗНАЧЕННЯ ЇЇ ОЗНАК

Перший розділ був присвячений вибору перспективної базової нейромережевої моделі для класифікації зображень та визначенню напрямів дослідження для підвищення співвідношення якості класифікації й обсягу обчислень.

Другий розділ спрямований на розробку нових рішень щодо збільшення кількості класифікаторів на першому ступені ансамблевого класифікатора зі стекінгом. Необхідність цього зумовлена тим, що результуюча якість класифікації ансамблевого класифікатора істотно залежить від кількості та ефективності класифікаторів першого ступеня. Однак кількість відомих НМ, які можуть бути ефективно використані на цьому ступені, є обмеженою.

Пропонується розв'язання проблеми за рахунок створення клонів НМ. В даному розділі вирішуються такі завдання:

- розробка та дослідження нової моделі першого ступеня ансамблевого класифікатора зі стекінгом на основі клонів НМ;
- розробка та дослідження методу визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зі стекінгом на основі клонів НМ, отриманих з використанням операторів повороту та циклічного зсуву зображень.

Представлені в цьому розділі результати базуються на працях [1], [2].

2.1 Модель ансамблевого класифікатора зі стекінгом на основі клонів нейронних мереж

Аналіз літератури показав, що однією з основних проблем ансамблевих класифікаторів є невелика кількість відомих НМ, які володіють високим співвідношенням якості класифікації та необхідного обсягу обчислень, що

підходять для використання на першому ступені ансамблевого класифікатора зі стекінгом [25].

Додатковою вимогою до класифікаторів першого ступеня є *структурованість* обробки вхідних зображень, яка на відміну від однорідної забезпечує те, що різні класифікатори помиляються на різних зображеннях. У той же час, підвищення якості класифікації ансамблевого класифікатора зі стекінгом в першу чергу пов'язане з підвищенням кількості класифікаторів на першому ступені.

Для вирішення цієї проблеми пропонується нова модель ансамблевого класифікатора зі стекінгом НМ, яка відрізняється від базової моделі (див. рис. 1.1) додаванням клонів первинних нейромережових класифікаторів на перший ступінь ансамблю (рис. 2.1).

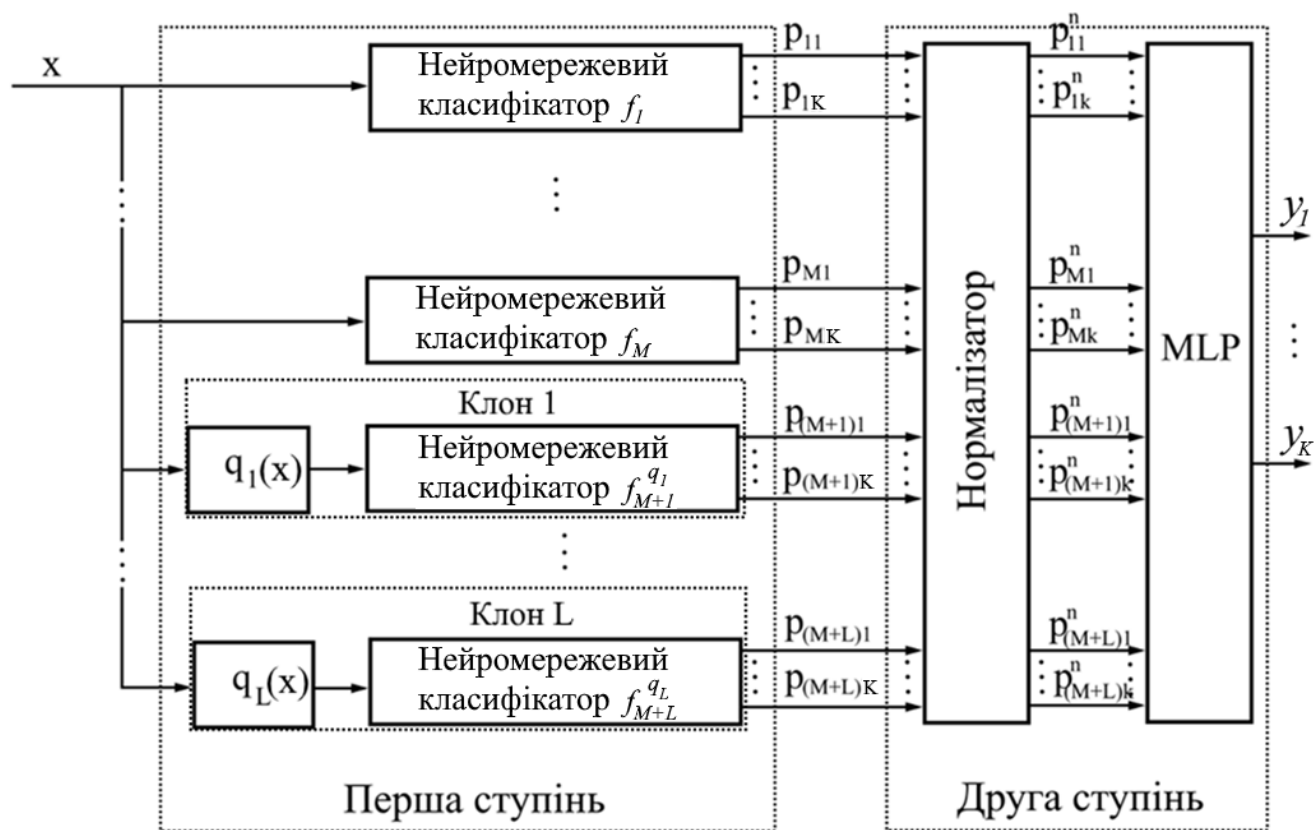


Рисунок 2.1 – Запропонована модель ансамблевого класифікатора зі стекінгом на основі клонів НМ

Структурно-параметричними ознаками першого ступеня запропонованої моделі ансамблевого класифікатора зі стекінгом НМ є:

- множина типів вибраних базових класифікаторів з їхніми власними параметрами – $\mathbf{F}(\Theta)$;
- загальна кількість класифікаторів (базових і клонів) на першому ступені – $N = M + L$;
- множина типів базових класифікаторів, що входять до складу клонів – $\mathbf{F}_{\square}^q(\Theta)$;
- множина типів і параметри операторів перетворення, що входять до складу клонів – $\mathbf{Q} = (q_1, q_2, \dots, q_L)$.

Набір структурно-параметричних ознак першого ступеня моделі ансамблевого класифікатора зі стекінгом НМ можна представити у вигляді кортежу:

$$V = \langle \mathbf{F}(\Theta), N, \mathbf{F}_{\square}^q(\Theta), \mathbf{Q} \rangle. \quad (2.1)$$

Найкращий набір ознак першого ступеня підбирається під особливості оброблюваних зображень і об'єктів на них, в рамках наявних обмежень обчислювальної потужності обладнання ансамблевого класифікатора. Для визначення ознак моделі першого ступеня ансамблевого класифікатора потрібно розробити відповідний метод.

На рис. 2.1 представлено ансамблевий класифікатор зі стекінгом НМ, у якому на першому ступені використовуються M базових класифікаторів і L клонів. На другому ступені застосовується нормалізатор вихідних векторів і багатошаровий перцептрон, який формує підсумкове рішення ансамблю.

На відміну від базової моделі ансамблевого класифікатора зі стекінгом (див. підрозділ 1.2) в запропонованій моделі ансамблевого класифікатора перший ступінь представлений множиною з M нейронних мереж і L клонів нейронних мереж $\{f_1, f_2, \dots, f_M, f_{M+1}^{q_1}, f_{M+2}^{q_2}, \dots, f_{M+L}^{q_L}\}$.

Клон відрізняється від базового класифікатора тим, що до його складу входить, окрім самого базового класифікатора, і оператор перетворення вхідного зображення $q_j(x)$, де $j = 1 \dots L$. Наявність у складі першого ступеня ансамблевого класифікатора клона моделі базового класифікатора має забезпечувати структурованість обробки та високе співвідношення якості класифікації і необхідного обсягу обчислень.

Завданням операторів перетворення вхідного зображення $q_j(x)$ є зміна множини зображень, на яких робить помилки клон, таким чином, щоб клони і базовий класифікатор значною мірою робили помилки класифікації на різних зображеннях. Додатковою вимогою до оператора перетворення є невеликий обсяг обчислень, у порівнянні з обсягом обчислень для базового класифікатора.

Кожна з M різних НМ першого ступеня для кожного вхідного зображення $x \in X$ видає вектор імовірностей розподілу за K класами, тобто вектор імовірності належності зображення x одному з K класів (див. (1.1)).

Кожен з L клонів НМ першого ступеня для кожного вхідного зображення $x \in X$ також видає вектор імовірностей розподілу за K класами, тобто вектор імовірності належності зображення x одному з K класів

$$P_l^{q_l} = f_l(x; \theta_l) = [p_{l1}^{\square}, p_{l2}^{\square}, \dots, p_{lK}^{\square}], \quad (2.2)$$

де $p_{lk}^{\square} \in [0,1]$ і $\sum_{k=1}^K p_{lk}^{\square} = 1$, $l = 1, \dots, L$;

θ_l – вектор вагових коефіцієнтів l -ого клона НМ першого ступеня.

Результати роботи всіх базових моделей і клонів НМ першого рівня об'єднуються у вектор мета-ознак (конкатенація векторів імовірностей):

$$\mathbf{P} = \text{concat}(P_1, P_2, \dots, P_M, P_{M+1}^{q_1}, P_{M+2}^{q_2}, \dots, P_{M+L}^{q_L}) \in R^C, \quad (2.3)$$

де $P_1, \dots, P_M$ – вихідні вектори M базових класифікаторів;

$P_{M+1}^{q_1}, \dots, P_{M+L}^{q_L}$ – вихідні вектори L клонів НМ;

$$C=N \times K;$$

$N = M + L$ – загальна кількість класифікаторів на першому ступені ансамблевого класифікатору зі стекінгом.

Таким чином, вхідне зображення x трансформується у вектор $\mathbf{P}$, який містить оцінки кожної НМ і кожного клону НМ про ймовірність належності зображення до кожного з K класів.

На другому ступені ці дані нормалізуються і подаються на вхід до MLP. Нормалізація виконується для кожного класифікатора окремо:

$$a_i(x) = \max_j \{ |p_{ij}(x)| \}, \quad (2.4)$$

$$p_{ij}^n(x) = p_{ij}(x)/a_i(x), \quad (2.5)$$

де i – номер класифікатора, $i = 1, 2, \dots, N, N = M + L$.

MLP складається з 3 шарів [21]:

- вхідний шар, розмірності $(M+L) \cdot K$;
- прихований шар, розмірності $(M+L-1) \cdot K$, функція активації *Relu*;
- вихідний шар, розмірності K , функція активації *softmax*.

Клас поточного зображення x визначається за індексом виходу MLP з максимальним значенням:

$$class(x) = \arg \left\{ \max_j \left(y_j(x) \right) \right\}, \quad (2.6)$$

де $y_j(x)$ – мітки класів на виході MLP для зображення x на вході.

У даній роботі запропоновано використовувати на першому ступені ансамблевого класифікатору в якості операторів перетворення – оператор повороту зображень (T) та оператор циклічного зсуву (S) випадкових наборів рядків або стовпців вхідного зображення. При цьому кут повороту або набір рядків/стовпців для зсуву для кожного з клонів є постійним і не змінюється при навчанні та при обробці вхідних зображень. Визначення ознак моделі ансамблевого класифікатору

зі стекінгом з використанням обраних операторів перетворення потребує додаткового дослідження з метою формалізації цього процесу.

2.2 Використанням оператора повороту для створення клонів нейронних мереж на першому ступені ансамблевого класифікатора зі стекінгом нейронних мереж

Комп'ютерні експерименти, що описані в цьому розділі, проводилися на наборі даних CIFAR-10 [39]. Поворот зображень здійснювався відносно центру зображення. Перерахунок значень пікселів у повернутому зображенні проводився лише в межах кола радіусом 15,5 пікселів, центр якого збігається з центром зображення (якщо прийняти відстань між пікселями за одиницю). Пікселі, розташовані поза межами цього кола (наприклад, у кутах зображення), залишалися без змін.

На рис. 2.3 наведено приклади повернутих зображень з набору даних CIFAR-10 [39], що використовуються при побудові та навчанні клонів НМ першого ступеня ансамблевого класифікатора.

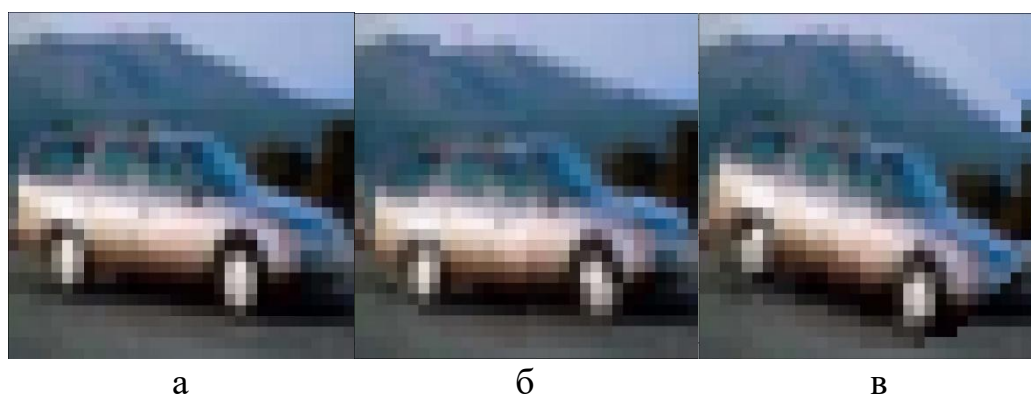


Рисунок 2.3 – Зображення автомобіля з набору CIFAR-10:

а – без повороту; б – поворот на $+3^\circ$; в – поворот на -10°

Програми базових класифікаторів першого ступеня реалізовані мовою програмування Python із використанням бібліотек NumPy, TensorFlow, Keras та

Addons і запозичені з [82]. Навчання класифікаторів першого ступеня проводилося протягом 50 епох.

Навчання MLP здійснювалося протягом 4 епох. Під час навчання MLP класифікатори першого ступеня не перенавчалися – вони використовували фіксовані вагові коефіцієнти, отримані на етапі попереднього навчання. Для кожної конфігурації ансамблевого класифікатора навчання MLP повторювалося 100 разів, при цьому початкові умови ініціалізувалися випадковим чином за методом ініціалізації Хав'єра [83]. Як кінцевий результат для кожної конфігурації використовувалося максимальне значення точності класифікації.

Навчання всіх класифікаторів виконувалося на навчальних даних набору CIFAR-10. Підсумкова точність класифікації зображень оцінювалась на тестовій виборці за метрикою Ассурасу. Усі обчислення проводилися на хмарній платформі Google Colab [84].

2.2.1 Дослідження якості класифікації об'єктів на зображеннях з застосуванням оператора повороту вхідних зображень на прикладі MLP-Mixer

Виходячи з того, що однією із основних вимог до класифікаторів першого ступеня є здійснення помилок на інших зображеннях, на відміну від решти класифікаторів цього ж ступеня, ефективність пропонованих способів створення клонів оцінювалась за ступенем відмінності підмножини зображень, на яких робить помилки клон, і таких самих підмножин для інших клонів і базових класифікаторів.

Для оцінювання ефективності способу створення клонів застосовано таку методику:

- порівняння кількості помилок, що збігаються між усіма клонами і базовими класифікаторами, від яких вони створені;
- визначення розміру групи, що складається з одного базового класифікатора та його клонів, при якій немає зображень, на яких робили б помилку одночасно всі елементи групи;

– порівняння результуючої точності (*Accuracy*) ансамблевого класифікатора, що має на першому ступені лише базові класифікатори з результуючою точністю ансамблевого класифікатора, де на першому ступені додатково до базових класифікаторів додано клони.

У табл. 2.2 та 2.3 наведено кількість збігів помилок на виході MLP-Mixer при використанні зображень із навчальної вибірки без повороту та з різними кутами повороту (кут повороту вказано в градусах).

Таблиця 2.2 – Кількість збігів помилок, при використанні оператора повороту зображення (+кут) для створення клону

MLP-Mixer	Кут повороту зображення							
	0°	T(+3°)	T(+5°)	T(+10°)	T(+15°)	T(+20°)	T(+25°)	T(+30°)
0°	8314	5103	5419	5734	4374	6261	6217	6395
T(+3°)	5103	10427	6722	6778	4915	7659	7792	8133
T(+5°)	5419	6722	11805	8043	5254	8984	9045	9297
T(+10°)	5734	6778	8043	14090	5386	11148	11171	11286
T(+15°)	4374	4915	5254	5386	7649	5844	5764	6012
T(+20°)	6261	7659	8984	11148	5844	22825	18685	18852
T(+25°)	6217	7792	9045	11171	5764	18685	26481	22425
T(+30°)	6395	8133	9297	11286	6012	18852	22425	29209

Мережа MLP-Mixer у проведених експериментах містила чотири базові шари. Порівнянню підлягали:

- кількість збігів помилок між клонами та базовим класифікатором;
- кількість збігів помилок, що здійснюються групою клонів.

В табл. 2.4 наведено кількість збігів помилок, які одночасно присутні в усіх варіантах класифікаторів, залежно від набору варіантів клонів із різними кутами повороту зображень на вході.

Таблиця 2.3 – Кількість збігів помилок, при використанні оператора повороту зображення (– кут) для створення клону

MLP-Mixer	Кут повороту зображення							
	0°	T(–3°)	T(–5°)	T(–10°)	T(–15°)	T(–20°)	T(–25°)	T(–30°)
0°	8314	4814	4517	5733	5091	5139	5122	5040
T(–3°)	4814	8387	4813	5964	5295	5380	5405	5395
T(–5°)	4517	4813	7943	5546	4969	5026	5117	5099
T(–10°)	5733	5964	5546	13919	6817	7108	7243	7226
T(–15°)	5091	5295	4969	6817	9685	6375	6289	6255
T(–20°)	5139	5380	5026	7108	6375	10439	6819	6776
T(–25°)	5122	5405	5117	7243	6289	6819	11010	7152
T(–30°)	5080	5395	5099	7226	6255	6776	7152	11578

Таблиця 2.4 – Кількість збігів помилок, при використанні оператора повороту зображення (комбінація кутів) для створення клону

MLP-Mixer + клони MLP-Mixer (CL), що навчені на зображеннях з різним кутом повороту	Кількість збігів помилок
MLP-Mixer	8314
MLP-Mixer, CL(–3°)	4814
MLP-Mixer, CL(–3°), CL(+3°)	981
MLP-Mixer, CL(–3°), CL(+3°), CL(–5°)	146
MLP-Mixer, CL(–3°), CL(+3°), CL(–5°), CL(+5°)	29
MLP-Mixer, CL(–3°), CL(+3°), CL(–5°), CL(+5°), CL(–10°)	2
MLP-Mixer, CL(–3°), CL(+3°), CL(5°), CL(+5°), CL(–10°), (+10°)	1
MLP-Mixer, CL(–3°), CL(+3°), CL(–5°), CL(+5°), CL(–10°), CL(+10°), CL(–15 °)	0

В результаті аналізу даних, наведених у табл. 2.2 – 2.4 встановлено, що поворот зображень, які подаються на вхід класифікатора з виражено структурованою обробкою, призводить до суттєвих відмінностей у помилках.

Наприклад, між MLP-Mixer та його клонами від 47 % до 23 % помилок не збігаються.

Також із табл. 2.4 видно, що вже в групі з базового класифікатора та 7 клонів немає жодної помилки, яку б одночасно здійснювали всі НМ цієї групи. Слід також відзначити значний розкид кількості збігів помилок, що свідчить про те, що для конкретного набору даних можна підібрати комбінацію клонів, які роблять меншу кількість збігів помилок, тобто забезпечують вищу різноманітність результатів і, відповідно, краще потенційне підвищення якості класифікації на другому ступені ансамблевого класифікатора.

2.2.2 Порівняння якості класифікації зображень при нарощуванні моделі MLP-Mixer «в глибину» та «в ширину»

Точність класифікації ансамблевого класифікатора зі стекінгом залежить як від кількості класифікаторів на першому ступені, так і від точності класифікації кожного з них. Тому було порівняно ефект як від підвищення точності класифікації базового класифікатора шляхом збільшення кількості послідовних блоків у його моделі (збільшення «в глибину»), так і від використання клонів того самого класифікатора (збільшення «в ширину»). Обидва підходи відповідають однаковому збільшенню обсягу обчислень. (Див. п.1.3.3)

В якості одиниці обсягу обчислень будемо використовувати обсяг обчислень для MLP-Mixer, що складається з одного блоку, який, у свою чергу, включає чотири базові субблоки (див. рис. 1.9). Обсяг обчислень для одного такого блоку становить близько 440 тис. операцій з рухомою комою (англ. Floating-point operations per second, FLOPs).

Ансамблевий класифікатор містив на першому ступені один MLP-Mixer (один блок), який обробляє вхідні зображення (без повороту), а також класифікатори-клони, що мають ідентичну архітектуру (один блок), але працюють із зображеннями, повернутими на різні кути. Результати порівняння наведено у табл. 2.5.

Великий розкид значень для MLP-Mixer пояснюється тим, що через значний обсяг обчислень кожна конфігурація навчалася лише один раз, починаючи зі випадкових початкових значень вагових коефіцієнтів, з зупинкою після досягнення найкращого результату. Для ансамблевого класифікатора другий ступінь навчався 100 разів, після чого обиралося найкраще значення якості класифікації.

Таблиця 2.5 – Порівняння якості класифікації MLP-Mixer при нарощуванні в глибину і ансамблевого класифікатора на основі MLP-Mixer і його клонів при нарощуванні в ширину

Кількість блоків	MLP-Mixer з нарощуванням в глибину	Ансамблевий класифікатор, на першому ступені MLP-Mixer + клони (нарощування в ширину)	
	<i>Accuracy</i>	<i>Accuracy</i> (і покращення)	Кути повороту вхідних зображень для різних клонів в градусах
1	0.7674	0.7674	MLP-Mixer
2	0.7814	0.7881 (+0.9%)	MLP-Mixer +CL(−3)
3	0.7868	0.7920 (+0.7%)	MLP-Mixer + CL(−3)+ CL(+3)
4	0.7633	0.7934 (+3.9%)	MLP-Mixer + CL(−3)+ CL(+3)+ CL(−5)
5	0.7740	0.7951 (+2.7%)	MLP-Mixer + CL(−3)+ CL(+3)+ CL(−5)+ CL(+5)
6	0.7831	0.7957 (+1.6%)	MLP-Mixer +CL(−3)+CL(+3)+CL(−5)+CL(+5)+CL(−10)
7	0.7719	0.7955 (+3.1%)	MLP-Mixer +CL(−3)+CL(+3)+CL(−5)+CL(+5)+CL(−10)+CL(+10)
8	0.7824	0.7952 (+1.6%)	MLP-Mixer+CL(−3)+CL(+3)+CL(5)+CL(+5)+CL(−10)+CL(+10)+CL(−15)

З результатів, наведених у табл. 2.5 видно, що нарощування в ширину для MLP-Mixer є більш переважним варіантом, ніж нарощування в глибину. Якість класифікації ансамблевого класифікатора в усіх варіантах була вищою і становила від 0.9% до 3.9% залежно від варіанта.

Однак необхідно зауважити, що додавання до ансамблевого класифікатора клонів 6 і 7 з великими кутами повороту вхідних зображень призвело не до поліпшення, а до погіршення результатів. Це пояснюється тим, що зі збільшенням кута повороту вхідних зображень погіршується якість класифікації клона і ефект від невідповідності помилок не перекриває це погіршення.

2.2.3 Дослідження якості класифікації при додаванні клонів класифікаторів до першого ступеня ансамблевого класифікатора

Основна мета створення та використання клонів – підвищення якості класифікації ансамблевого класифікатора за рахунок збільшення кількості класифікаторів на першому ступені. Тому становить інтерес дослідити величину цього підвищення на конкретному прикладі ансамблевого класифікатора з роботи [19]. Комп'ютерні експерименти проведено так само, як і в усьому розділі, на наборі даних CIFAR-10. Результати наведено в табл. 2.6.

У експерименті на першому ступені було вид 2 до 6 базових класифікаторів (див.табл.1.1 [19]) і до них додавалось 1 та 2 клони MLP-Mixer (стовпчики 3 та 4), або 1 та 2 клони ССТ (стовпчики 5 та 6), або 1 та 2 клони CNN (стовпчики 7 та 8).

З результатів табл. 2.6 видно, що додавання клонів класифікатора MLP-Mixer дозволяє зменшувати кількість помилок на 1%-11%, залежно від вихідного набору класифікаторів першого ступеня. Зменшення кількості помилок для клонів класифікатора ССТ знаходиться в діапазоні 2,1%-10%. Використання клонів класифікатора CNN призводило до збільшення кількості помилок від 9% до 38%. Таке погіршення результатів при використанні клонів CNN пояснюється тим, що обробка в CNN не є структурованою. У результаті клони CNN мають суттєво гіршу якість класифікації й роблять помилки на однакових зображеннях.

Таблиця 2.6 – Результати порівняння якості класифікації зображень з використанням на першому ступені ансамблевого класифікатора зі стекінгом базових моделей НМ та цих же базових моделей НМ з додаванням клонів (оператор повороту)

Базові класифікатори першого ступеня		Без використання клонів	Класифікатори та клони, що додані					
			MLP-Mixer T(-3°)	MLP-Mixer T(-3°), T(+3°)	CCT T(-3°)	CCT T(-3°), T(+3°)	CNN T(-3°)	CNN T(-3°), T(+3°)
1		2	3	4	5	6	7	8
CCT+ EATNet	<i>Accuracy</i>	0.8184	0.8365	0.8390	0.8354	0.8374	0.7815	0.8024
	Кількість помилок	1816	1635	1610	1646	1626	2185	1976
	Зменшення помилок, %	0	-10	-11	-9	-10	+20	+9
CCT+ EATNet + MLP-Mixer	<i>Accuracy</i>	0.8401	0.8421	0.8425	0.8434	0.8460	0.7893	0.8056
	Кількість помилок	1599	1579	1575	1566	1540	2107	1944
	Зменшення помилок, %	0	-1,3	-1,5	-2,1	-3,7	+32	+22
CCT+ EATNet + MLP-Mixer+ FNet	<i>Accuracy</i>	0.8445	0.8461	0.8458	0.8465	0.8493	0.7894	0.8062
	Кількість помилок	1555	1539	1542	1535	1507	2106	1938
	Зменшення помилок, %	0	-1	-0,8	-1,3	-3,1	+35	+25
CCT+ EATNet + MLP-Mixer+ FNet+ gMLP	<i>Accuracy</i>	0.8457	0.8473	0.8471	0.8490	0.8496	0.7913	0.8067
	Кількість помилок	1543	1527	1529	1510	1504	2087	1933
	Зменшення помилок, %	0	-1	-1	-2,1	-2,5	+35	+25
CCT+ EATNet + MLP-Mixer+ FNet+ gMLP+ SwinTr	<i>Accuracy</i>	0.8468	0.8473	0.8482	0.8491	0.8500	0.7889	0.8034
	Кількість помилок	1 532	1527	1518	1509	1500	2111	1966
	Зменшення помилок, %	0	-0,33	-1	-1,5	-2,1	+38	+28

2.2.4 Дослідження залежності кількості збігів помилок базового класифікатора і моделі ансамблевого класифікатора з клонами від кута повороту зображення

Дослідження дозволить знайти найкращий кут повороту вхідного зображення при якому, з одного боку, буде найменша помилка самого клону, з іншого боку, буде найменша кількість помилок, що збігаються з помилками базового класифікатора.

Дані табл. 2.2 і 2.3 показують погіршення якості класифікації клонів зі збільшенням кута повороту зображення. З іншого боку, зменшення кута повороту асимптотично призведе до збігу повернутого зображення з неповернутим і, відповідно, до збігу помилок базового класифікатора та його клона. У зв'язку з цим становить інтерес дослідження показників клонів для невеликих кутів повороту зображення.

Оскільки з усіх базових класифікаторів ССТ і його клони (згідно з табл. 1.1 і 2.6) показують найкращі результати, дане дослідження проведено для ССТ і його клонів. Узагальнена архітектура ансамблевого класифікатора для цього випадку наведена на рис. 2.4.

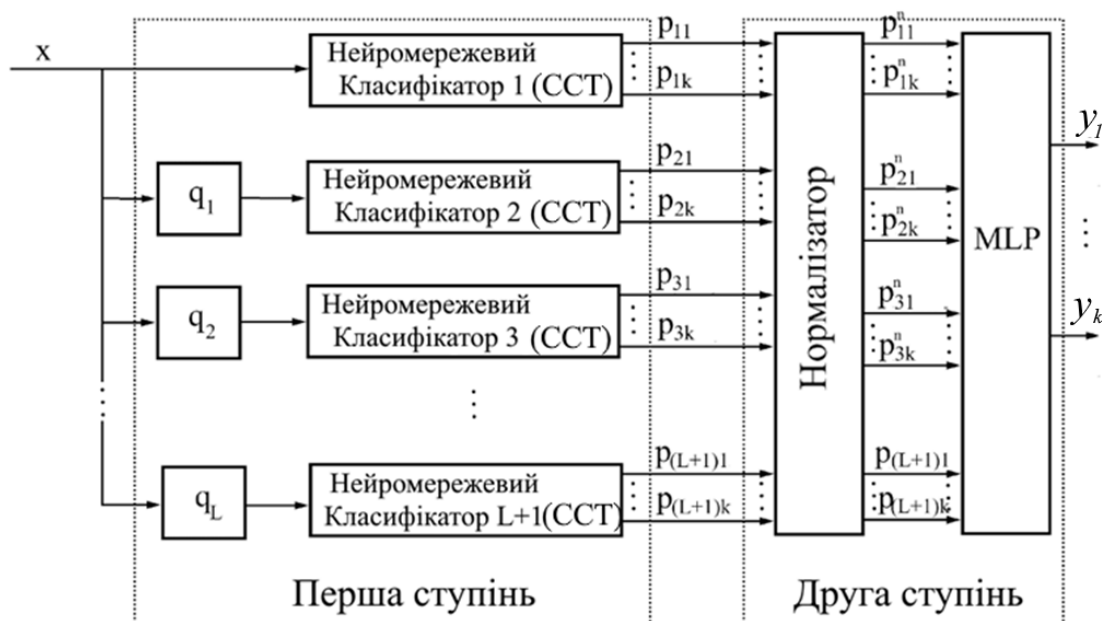


Рисунок 2.4 – Узагальнена архітектура ансамблевого класифікатора при використанні на першому ступені одного класифікатора ССТ та його клонів

За аналогією з пунктом 2.1.2, дослідження проведено для клонів ССТ у діапазоні кутів від 0 до 0,3 градуса. Результати для тестових 10 тис. зображень із набору даних CIFAR-10 наведені на рис. 2.5.

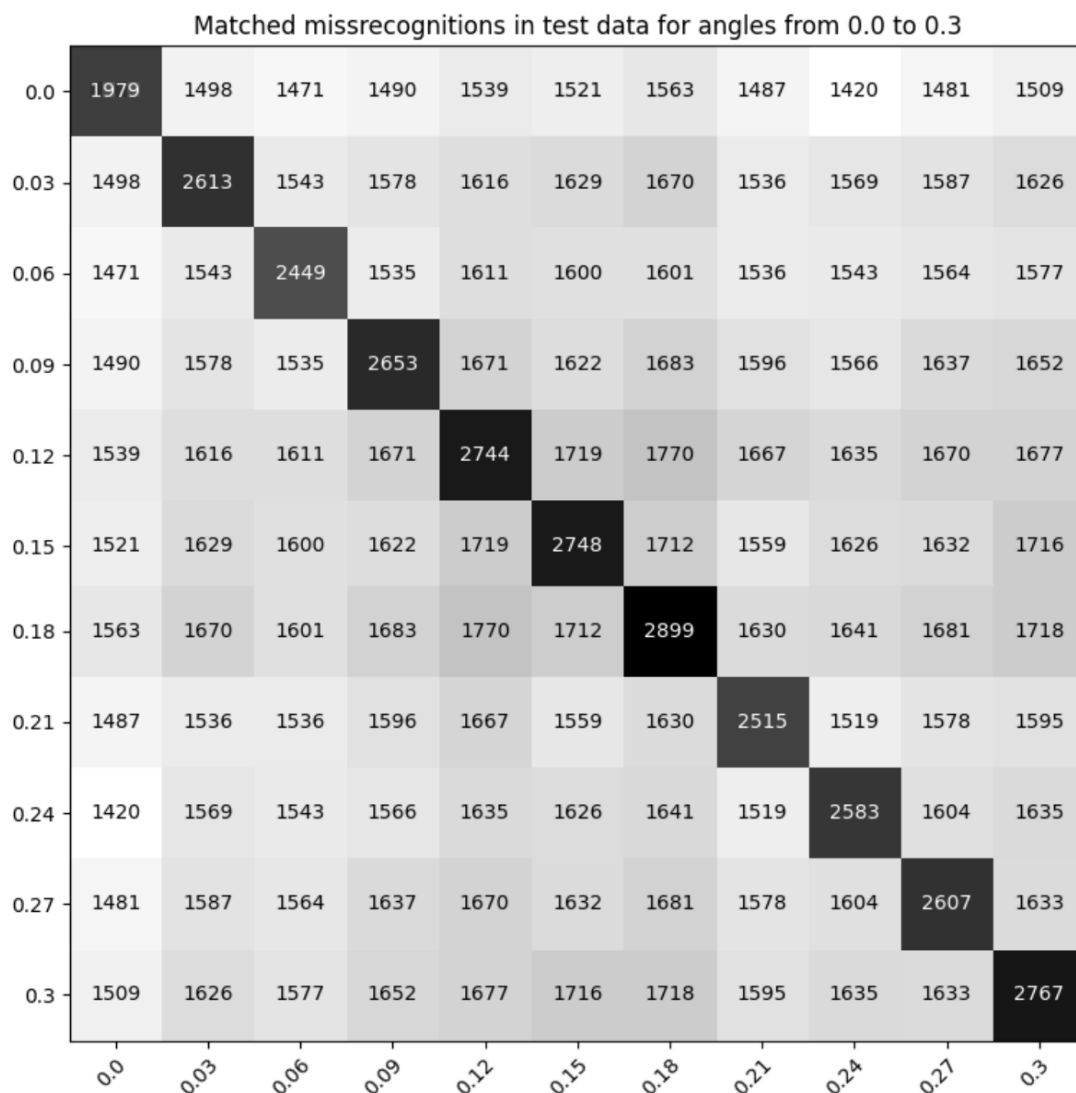


Рисунок 2.5 – Матриця збігів помилок на тестовому наборі між ССТ та його клонами, які використовують оператор повороту зображень T від 0° до 0.3° градуса з $\Delta = 0.03^\circ$

Для побудови цієї таблиці збігів помилок проводилися розрахунки помилок класифікації тестових зображень базовим класифікатором ССТ та його клонами і у клітини таблиці записувалася кількість зображень, на яких одночасно робили помилки базова нейронна мережа ССТ або клони, позначені відповідно по

горизонталі і вертикалі. По горизонталі та вертикалі вказані кути повороту в градусах. 0 градусів відповідає неповерненим зображенням, тобто базовому класифікатору ССТ. Відмінні від нуля кути повороту відповідають клонам, що використовують оператор повороту із зазначеним кутом.

Як видно з рис. 2.5, кількість збігів помилок практично не залежить від кута повороту і знаходиться в діапазоні від 1420 до 1770 помилок. Безпосередньо ССТ робить 1979 помилок, його клони роблять суттєво більше. Однак із даних рис. 2.5 з таблицею спільних помилок видно, що різні клони роблять помилки на різних зображеннях – кількість збігів помилок суттєво менша за кількість помилок, які робить ССТ. Також найменшу суму помилок, що збігаються з помилками ССТ має рядок для кута повороту 0.24 градуса. Тому цю величину доцільно обрати як крок кута поворотів Δ для побудови клонів ССТ. Необхідно зазначити, що у зв'язку з великим обсягом розрахунків обчислення параметрів клонів здійснювалося одноразово. При практичному використанні цього підходу може бути отримане інше найкраще значення кроку кута поворотів.

2.2.5 Порівняння якості класифікації ансамблевого класифікатора при використанні на першому ступені базових класифікаторів і ССТ та його клонів

Метою даного експерименту є експериментальне підтвердження того, що використання на першому ступені ансамблевого класифікатора зі стекінгом НМ найкращого базового класифікатора і заміна інших базових класифікаторів на клони найкращого базового класифікатора призводить до підвищення результуючої якості класифікації всього ансамблевого класифікатора.

Результати порівняння наведені в табл. 2.7. Клони моделі ССТ позначені як $CL_{\text{ССТ}}$, а в дужках указано кут повороту зображення у сотих частках градуса. Із таблиці видно, що використання ССТ та його клонів забезпечує кращу якість класифікації – від 0.07% до 1.2%.

Таблиця 2.7 – Порівняння якості класифікації ансамблевого класифікатора зі стекінгом базових моделей НМ та моделі CCT і її клонів на першому ступені

Результат використання різних моделей НМ на першому ступені		Результат використання CCT та її клонів CL_{CCT} на першому ступені	
1	2	3	4
Моделі НМ на першому ступені	<i>Accuracy</i>	Моделі НМ на першому ступені	<i>Accuracy</i> (% зміни)
CCT+ EATNet	0.8184	CCT+ $CL_{CCT}(0.24)$	0.8230 (+0.56%)
CCT+ EATNet+ MLP-Mixer	0.8401	CCT+ $CL_{CCT}(0.24)$ + $CL_{CCT}(-0.24)$	0.8407 (+0.07%)
CCT+ EATNet + MLP-Mixer+ FNet	0.8445	CCT+ $CL_{CCT}(0.24)$ + $CL_{CCT}(0.24)$ + $CL_{CCT}(0.48)$	0.8474 (+0.3%)
CCT+ EATNet + MLP-Mixer+ FNet+gMLP	0.8457	CCT+ $CL_{CCT}(0.24)$ + $CL_{CCT}(-0.24)$ + $CL_{CCT}(0.48)$ + $CL_{CCT}(-0.48)$	0.8538 (+1.0%)
CCT+ EATNet + MLP-Mixer+ FNet+gMLP+ SwinTr	0.8468	CCT+ $CL_{CCT}(0.24)$ + $CL_{CCT}(-0.24)$ + $CL_{CCT}(0.48)$ + $CL_{CCT}(-0.48)$ + $CL_{CCT}(0.72)$	0.8568 (+1.2%)

2.3 Використання оператора циклічного зсуву для створення клонів нейронних мереж

Ідея побудови клонів НМ з використанням оператора зсуву аналогічна до ідеї використання оператора повороту зображень. В її основі лежить базова НМ, на вхід до якої подається не повернуте зображення, а зображення, у якому деякі рядки або стовпці циклічно зсунуті. При цьому беремо модель ансамблевого класифікатора, що запропонована в цьому розділі (див. підрозділ 2.1).

Основна перевага такого способу полягає в тому, що він не потребує виконання операцій із рухомою комою, а лише операцій із пам'яттю. Додатковою перевагою є те, що кількість векторів, які задають зсуви, практично не обмежена, тобто кількість формованих клонів може бути дуже великою.

В даному випадку перетворення q являють собою циклічні зсуви S , рядків або стовпців пікселів вхідного зображення, відповідно до векторів зсувів:

$$s_r = (s_{r1}, s_{r2}, \dots, s_{rs}), \quad (2.7)$$

де s_{ri} – цілі числа в діапазоні від 0 до D , що формуються генератором випадкових чисел з рівномірним розподілом;

rs – кількість рядків або стовпців у зображенні.

Кожний рядок (або стовпець) вхідного зображення циклічно зсувається на s_{ri} позицій. Якщо $s_{ri} = 0$, то рядок не зсувається. Для кожного з клонів формується власний вектор зсувів, який залишається незмінним під час навчання та тестування.

Дослідженню підлягає залежність ефективності використання клонів від величини D , а також відмінності при зсуві по рядках і по стовпцях.

Крім того, досліджувалася ефективність використання розріджених векторів зсувів s_{poz} при $D = 1$: **Потрібно однакове позначення індексів**

$$s_{poz} = (s_{poz1}, s_{poz2}, \dots, s_{pozrs}), \quad (2.8)$$

де елементи цього вектору обчислюються за допомогою двох незалежних векторів $s_r^{(1)}$ та $s_r^{(2)}$ шляхом логічного множення відповідних компонентів:

$$s_{poz i} = \begin{cases} s_{ri}^{(1)}, & \text{якщо } s_{ri}^{(2)} = 1, \\ 0, & \text{якщо } s_{ri}^{(2)} = 0. \end{cases} \quad (2.9)$$

Кількість одиниць в векторі s_{poz} є випадковою величиною, а його розмірність дорівнює кількості рядків або стовпців у вхідному зображенні.

Відповідний розподіл, отриманий по 10 тис. реалізацій, наведено на рис. 2.6. Із рис. 2.6 видно, що найбільш ймовірна кількість одиниць у векторі s_{poz} дорівнює 8, тобто циклічно зсунути на 1 позицію будуть 8 рядків або стовпців. Не виключена ситуація, коли вектор s_{poz} не міститиме одиниць, тобто рядки або стовпці зображення не будуть зсунуті.

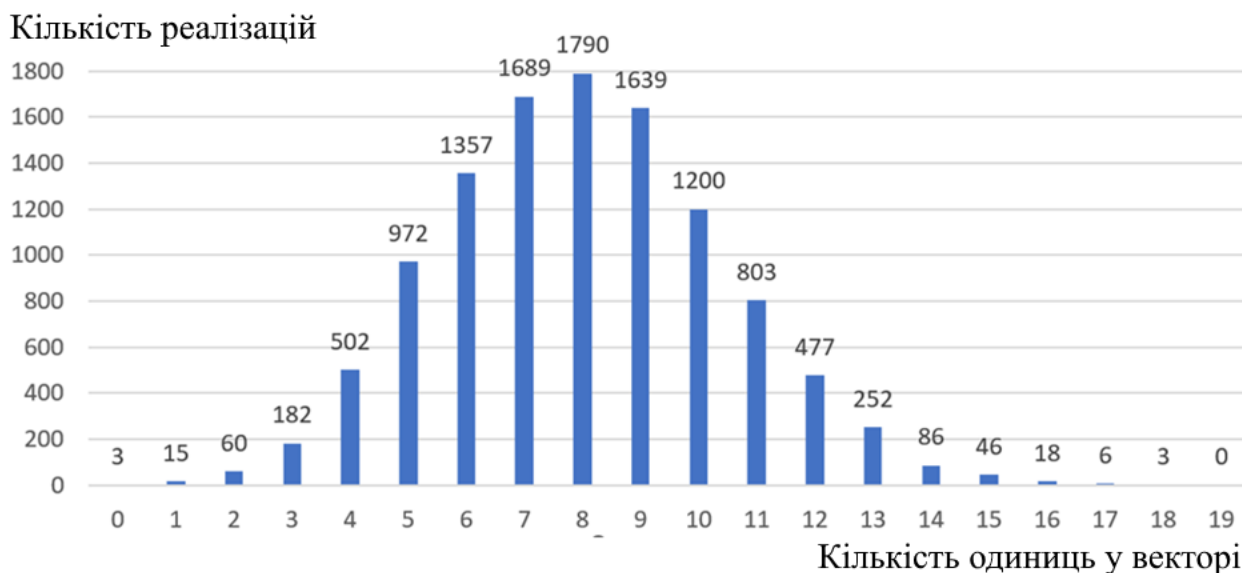


Рисунок 2.6 – Гістограма розподілу кількості одиниць у розрідженому векторі зсувів

Незважаючи на те, що ймовірність такої ситуації невелика, при практичній реалізації її необхідно виключити, оскільки це буде просто дублювання базової нейронної мережі.

Під час дослідження конкретного набору даних необхідно випробовувати використання різних векторів $s_{роз}$ і відбирати ті з них, які забезпечують найменшу кількість збігів помилок і потребують меншої кількості зсувів.

2.3.1 Порівняння кількості збігів помилок між клонами та базовою нейронною мережею на прикладі MLP-Mixer і CCT

Відповідно до методики оцінювання ефективності способу створення клонів, запропонованої в п. 2.2.1, у цьому підпункті вибрано два найкращі базові класифікатори MLP-Mixer і CCT. Для MLP-Mixer проведено дослідження кількості помилок, що збігаються між клонами і базовим класифікатором при формуванні клонів методом циклічних зрушень. Крім цього, для MLP-Mixer і CCT визначено мінімальні розміри груп, що складаються з базового класифікатора та його клонів, для яких немає зображень, на яких всі елементи групи роблять помилки одночасно.

У табл. 2.8 наведено кількість збігів помилок для базової нейронної мережі (позначення **0_1**) і десяти реалізацій її клонів (1_i , де i – номер клона базового класифікатора), які відрізняються своїм випадковим вектором зсуву по рядках S при $D = 1$. Перша цифра в позначенні – значення D , друга – номер реалізації. Кількість помилок наведено для набору даних для навчання CIFAR-10 (50 тис. зображень).

По діагоналі в таблиці зазначено кількість помилок, яких припускається сама нейронна мережа. Із табл. 2.8 видно, що кількість помилок, яких припускаються клони, співставна з кількістю помилок базової нейронної мережі й може бути навіть меншою. Кількість збігів помилок, яких припускаються різні мережі, становить 57%–65% від кількості помилок базової нейронної мережі.

Таблиця 2.8 – Кількість збігів помилок для MLP-Mixer (позначення **0_1**) і десяти реалізацій її клонів отриманих для різних векторів зсуву

НМ	0_1	1_1	1_2	1_3	1_4	1_5	1_6	1_7	1_8	1_9	1_10
0_1	8785	5229	5392	5117	5311	5292	5161	5433	5365	4967	5172
1_1	5229	8722	5425	5224	5257	5323	5407	5445	5429	5020	5381
1_2	5392	5425	8865	5133	5377	5327	5293	5539	5368	5097	5318
1_3	5117	5224	5133	8295	5179	5211	5123	5405	5317	4970	5285
1_4	5311	5257	5377	5179	8758	5435	5227	5699	5442	5177	5349
1_5	5292	5323	5327	5211	5435	8656	5277	5521	5327	5064	5315
1_6	5161	5407	5293	5123	5227	5277	8598	5546	5416	5031	5347
1_7	5433	5445	5539	5405	5699	5521	5546	9154	5742	5327	5612
1_8	5365	5429	5368	5317	5442	5327	5416	5742	9587	5172	5571
1_9	4967	5020	5097	4970	5177	5064	5031	5327	5172	7958	5074
1_10	5172	5381	5318	5285	5349	5315	5347	5612	5571	5074	8942

У табл. 2.9 наведено кількість збігів помилок, яких припускається базова НМ (**0_1**) та група її клонів, в якості базової НМ розглянуто MLP-Mixer і ССТ. Із табл. 2.9 видно, що клони припускаються значної кількості власних помилок на різних зображеннях. У групі, що складається з базової НМ MLP-Mixer та 7 клонів, немає збігів помилок, які б здійснювали всі 8 нейронних мереж. А для ССТ така група містить лише 5 клонів.

Таблиця 2.9 – Кількість збігів помилок у групі нейронних мереж для MLP-Mixer та CCT

Базова нейронна мережа	Склад класифікаторів першого ступеня (базового (0_1) і його клонів (1_i))							
	0_1	0_1 1_1	0_1 1_1 1_2	0_1 1_1 1_2 1_3	0_1 1_1 1_2 1_3 1_4	0_1 1_1 1_2 1_3 1_4 1_5	0_1 1_1 1_2 1_3 1_4 1_5 1_6	0_1 1_1 1_2 1_3 1_4 1_5 1_6 1_7
MLP-Mixer	8785	5229	892	140	25	2	1	0
CCT	7023	3541	558	69	8	0	0	0

2.3.2 Дослідження залежності результуючої якості класифікації ансамблевого класифікатора зі стекінгом від кількості клонів нейронної мережі

Для проведення експерименту було обрано дві базові НМ, а саме, MLP-Mixer та CCT.

Перший етап дослідження проводився за такими умовами експерименту:

- базовий класифікатор – MLP-Mixer;
- кількість клонів – L ;
- використовувані вектори s_r – випадкові;
- діапазон формування векторів зсуву – D ;
- зсув здійснювався по рядках;
- загальна структура ансамблевого класифікатора (див. рис. 2.4.)

Результати наведено для набору тестових зображень із набору даних CIFAR-10 (загальна кількість – 10 тис.).

У табл. 2.10 наведено значення якості класифікації ансамблевого класифікатора зі стекінгом НМ, що демонструють ефективність додавання до

першого ступеня ансамблевого класифікатора клонів MLP-Mixer. Із табл. 2.10 видно, що збільшення діапазону величин зсувів призводить до суттєвого погіршення якості класифікації. Для діапазону від 0 до $D=1$ якість класифікації зростає зі збільшенням кількості використаних клонів до $L=9$. При $L=10$ отримано те саме значення. При цьому якість класифікації зросла від 0.7674 (без клонів) до 0.7965 (для 9 клонів), тобто на 3.8%. Це відповідає зменшенню кількості помилок при класифікації на 291 або 12.5%.

В даному випадку перетворення q являють собою циклічні зсуви S , рядків або стовпців пікселів вхідного зображення, відповідно до векторів зсувів:

$$s_r = (s_{r1}, s_{r2}, \dots, s_{rs}), \quad (2.7)$$

де s_{ri} – цілі числа в діапазоні від 0 до D , що формуються генератором випадкових чисел з рівномірним розподілом;
 rs – кількість рядків s або стовпців у зображенні.

Таблиця 2.10 – Залежність якості класифікації *Accuracy* від кількості клонів та від діапазону зсувів під час формування векторів зсуву s_r

L	$D=1$	$D=2$	$D=3$	$D=4$	$D=5$	$D=6$
1	0.7828	0.7790	0.7818	0.7783	0.7755	0.7766
2	0.7856	0.7829	0.7829	0.7805	0.7794	0.7765
3	0.7910	0.7896	0.7840	0.7816	0.7823	0.7776
4	0.7920	0.7917	0.7863	0.7845	0.7833	0.7803
5	0.7936	0.7905	0.7852	0.7835	0.7841	0.7772
6	0.7953	0.7907	0.7842	0.7851	0.7824	0.7773
7	0.7957	0.7915	0.7859	0.7850	0.7840	0.7775
8	0.7954	0.7902	0.7861	0.7865	0.7838	0.7772
9	0.7965	0.7913	0.7860	0.7868	0.7841	0.7775
10	0.7965	0.7918	0.7873	0.7867	0.7835	0.7778

Аналогічні дані наведено в табл. 2.11, де порівнюються зсуви по рядках і стовпцях для діапазону від 0 до $D = 1$, а також використання розріджених векторів зсуву $S_{роз}$ (оператор зсуву $S_{роз}$, див.(2.11), (2.12)).

Таблиця 2.11 – Залежність якості класифікації *Accuracy* від кількості клонів (L) для MLP-Mixer при використанні різних операторів зсуву S

L	$S (D=1)$ Зсув по рядках	$S (D=1)$ Зсув по стовпцям	$S_{роз} (D=1)$ Зсув по стовпцям	$S_{роз} (D=1)$ Зсув по рядках
1	0.7828	0.7870	0.7828	0.7792
2	0.7856	0.7900	0.7902	0.7881
3	0.7910	0.7936	0.7943	0.7912
4	0.7920	0.7942	0.7949	0.7931
5	0.7936	0.7951	0.7969	0.7972
6	0.7953	0.7956	0.7987	0.8000
7	0.7957	0.7985	0.7995	0.8006
8	0.7954	0.7989	0.8004	0.8002
9	0.7965	0.7985	0.8012	0.8019
10	0.7965	0.7992	0.8015	0.8010
11	0.7975	0.7981	0.8011	0.8023
12	0.7972	0.7982	0.8007	0.8029

Із даних табл. 2.11 видно, що використання розріджених векторів зсуву та зсуву по рядках забезпечує певну перевагу за якістю класифікації й дозволяє отримати якість 0.8029 при використанні 12 клонів. Це становить 4.6% поліпшення якості або 15% зменшення кількості помилок класифікації.

На другому етапі дослідження аналогічні обчислення були проведені з використанням базового класифікатора ССТ (табл. 2.12).

У одиночного класифікатора ССТ (без клонів) якість класифікації на тестовому наборі даних становить 0.8021, що відповідає 1979 помилкам. При збільшенні кількості використаних клонів до 11 вдається підвищити якість класифікації до 0.8492 (покращення на 5.9%) і, відповідно, зменшити кількість помилок до 1508 (зменшення на 23.8%). Дані табл. 2.12 також свідчать про перевагу використання розріджених векторів зсуву та зсуву по рядках.

Таблиця 2.12 – Залежність якості класифікації *Accuracy* від кількості клонів (L) для ССТ при використанні різних операторів зсуву S

L	$S(D=1)$ Зсув по рядках	$S(D=1)$ Зсув по стовпцям	$S_{роз}$ Зсув по стовпцям	$S_{роз}$ Зсув по рядках
1	0.8242	0.8254	0.8223	0.8201
2	0.8295	0.8292	0.8286	0.8289
3	0.8340	0.8330	0.8342	0.8346
4	0.8372	0.8335	0.8352	0.8363
5	0.8381	0.8369	0.8377	0.8395
6	0.8400	0.8384	0.8413	0.8407
7	0.8414	0.8381	0.8424	0.8444
8	0.8420	0.8393	0.8417	0.8470
9	0.8428	0.8407	0.8446	0.8474
10	0.8425	0.8403	0.8439	0.8483
11	0.8434	0.8403	0.8435	0.8492
12	0.8422	0.8409	0.8449	0.8491

2.4 Порівняння якості класифікації при використанні на першому ступені моделі ансамблевого класифікатора зі стекінгом клонів НМ, отриманих за допомогою операторів зсуву і повороту зображень

Вище було показано (див. табл. 2.7), що додавання клонів MLP-Mixer, отриманих шляхом повороту зображень, завжди призводило до певного покращення якості класифікації всього ансамблевого класифікатора. На відміну від них, використання клонів MLP-Mixer, отриманих методом зсувів, для одних наборів базових класифікаторів давало кращі результати (–12% помилок), ніж при повороті (–11%), а для інших – гірші (+0.7% помилок).

Проведення такого порівняння для конкретного набору вхідних даних дозволяє визначити

- наскільки ефективно додавання на перший ступінь клонів базових нейронних мереж;
- який метод формування клонів дає найкращі результати для даного набору даних.

Результати порівняння якості використання клонів НМ, отриманих за допомогою операторів повороту зображень і зсуву, наведено в табл. 2.13. Для порівнянності результатів базові класифікатори та дані щодо клонів із поворотом узято з [пункту 2.1.4](#). Для зсувів використовувалися лише розріджені вектори.

Використання клонів ССТ завжди давало покращення якості класифікації, як у порівнянні із ситуацією без використання клонів, так і у порівнянні з використанням клонів MLP-Mixer, отриманих за допомогою повороту і зсуву. Діапазон зменшення кількості помилок у результаті додавання клонів ССТ, отриманих шляхом зсуву, становив від 3,1% до 13%.

Проведене дослідження показало, що зсув стовпців і рядків вхідних зображень дозволяє отримати клони базових НМ ансамблевого класифікатора, використання яких на першому ступені підвищує якість класифікації. Це пояснюється тим, що для використання на першому ступені ансамблевого класифікатора зі стекінгом необхідні НМ зі структурованою обробкою вхідних сигналів. Структурованість обробки призводить до того, що циклічний зсув рядків і стовпців змінює помилки, яких припускається НМ. Комбінування результатів НМ, що роблять помилки на різних наборах зображень, дозволяє підвищити на другому ступені результуючу якість класифікації.

Велика кількість клонів, які отримуються за допомогою зсувів, дозволяє виконувати відбір клонів із найкращою якістю класифікації. Це є додатковою перевагою запропонованого способу формування клонів.

Таблиця 2.13 – Якість класифікації ансамблевого класифікатора

Класифікатори першого ступеня	<i>Accuracy</i>	Кількість помилок	Зміна кількості помилок
ССТ+EATNet	0.8184	1816	0 %
ССТ+EAT+ 2 клони MLP-Mixer з поворотом	0.8390	1610	–11 %
ССТ+EATNet+	0.8402	1598	–12 %

2 клони MLP-Mixer зі зсувом			
CCT+EATNet+2 клони CCT зі зсувом	0.8429	1571	-13 %
CCT+EATNet+MLP-Mixer	0.8401	1599	0 %
CCT+EATNet+MLP-Mixer+ 2 клони MLP-Mixer з поворотом	0.8425	1575	-1,5 %
CCT+EATNet+MLP-Mixer+ 2 клони MLP-Mixer зі зсувом	0.8436	1564	-2,2 %
CCT+EATNet+MLP-Mixer+ 2 клони CCT зі зсувом	0.8482	1518	-5 %
CCT+EATNet+MLP-Mixer+FNet	0.8445	1555	0 %
CCT+EATNet+MLP-Mixer+FNet+ 2 клони MLP-Mixer з поворотом	0.8458	1542	-0,8 %
CCT+EATNet+MLP-Mixer+FNet+ 2 клони MLP-Mixer зі зсувом	0.8458	1542	-0,8 %
CCT+EATNet+MLP-Mixer+FNet+ 2 клони CCT зі зсувом	0.8496	1504	-3,3 %
CCT+EATNet+MLP-Mixer+FNet+gMLP	0.8457	1543	0 %
CCT+EATNet+MLP-Mixer+FNet+ gMLP+2 клони MLP-Mixer з поворотом	0.8471	1529	-1 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +2 клони MLP-Mixer зі зсувом	0.8457	1543	0 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +2 клони CCT зі зсувом	0.8514	1486	-3,7 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +SwinTr	0.8468	1 532	0 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +SwinTr+2 клони MLP-Mixer з поворотом	0.8482	1518	-1 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +SwinTr+2 клони MLP-Mixer зі зсувом	0.8457	1543	+0,7 %
CCT+EATNet+MLP-Mixer+FNet+gMLP +SwinTr+2 клони CCT зі зсувом	0.8515	1485	-3,1 %

Як впливає з отриманих результатів, найкращі результати забезпечують розріджені вектори зсувів (оператор циклічного зсуву S_{poz}). У середньому такі оператори зсувів вимагають виконати лише 8 циклічних зсувів рядків або стовпців на одну позицію (див. рис. 2.6).

Проведене порівняння клонів НМ MLP-Mixer і CCT показало, що їх клони суттєво відрізняються. Це свідчить про те, що при практичній реалізації ансамблевих класифікаторів зі стекінгом НМ необхідно здійснювати відбір клонів,

які забезпечують найвищу результуючу якість класифікації. Порівняння, що проведено у табл. 2.9 показує, що клони ССТ характеризуються більшим різноманіттям. Група, що складається з ССТ і п'яти клонів, уже не має таких помилок, які б здійснювали всі класифікатори першого ступеня. Для MLP-Mixer така група повинна містити вже 7 клонів. Із табл. 2.11 і табл. 2.12 видно, що збільшення кількості клонів MLP-Mixer на першому ступені до 12 дозволяє підвищити результуючу якість класифікації на 4,6%, тоді як для ССТ це підвищення становить 5,9%. Додавання двох клонів ССТ до різної кількості базових класифікаторів першого ступеня (див. табл. 2.13) забезпечувало вищу якість на 1–3,7% порівняно з двома клонами MLP-Mixer.

Як обмеження запропонованого підходу можна зазначити необхідність проведення додаткових досліджень для відбору клонів базових класифікаторів, що забезпечують найбільше підвищення результуючої якості класифікації всього ансамблевого класифікатора.

Недоліком запропонованого підходу є те, що зі збільшенням кількості клонів на першому ступені приріст результуючої якості класифікації зменшується і практично повністю зупиняється у розглянутому прикладі, наприклад, при кількості клонів 11–12 (див. табл. 2.11, 2.12). Подолання цього недоліку вимагає проведення додаткових досліджень і розроблення моделей нових НМ зі структурованою обробкою.

2.5 Метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зі стекінгом на основі клонів нейронних мереж

Ознаки моделі першого ступеня ансамблевого класифікатора зі стекінгом НМ (2.1) з додаванням клонів базових класифікаторів, побудованих з використанням операторів повороту і циклічного зсуву можна представити у вигляді кортежу параметрів:

$$V = \langle \mathbf{F}(\Theta), N, \mathbf{F}_{\square}^T(\Theta), \mathbf{F}_{\square}^S(\Theta), \mathbf{Q}(\mathbf{T}), \mathbf{Q}(\mathbf{S}_{\text{роз}}) \rangle, \quad (2.10)$$

де $\mathbf{F}(\Theta)$ – множина типів вибраних базових класифікаторів з їхніми власними параметрами Θ ;

$N = (M + L)$ – загальна кількість класифікаторів на першому ступені;

M – кількість базових класифікаторів на першому ступені;

L – кількість клонів на першому ступені;

$\mathbf{F}_{\square}^T(\Theta)$ – множина клонів, отриманих за допомогою оператора поворотів $\mathbf{T}$;

$\mathbf{F}_{\square}^S(\Theta)$ – множина клонів, отриманих за допомогою оператора циклічних зсувів з розрідженими векторами $\mathbf{S}_{\text{роз}}$;

$\mathbf{Q}(\mathbf{T})$ – множина типів і параметри операторів поворотів $\mathbf{T}$, що входять до складу клонів;

$\mathbf{Q}(\mathbf{S}_{\text{роз}})$ – множина типів і параметри операторів циклічних зсувів з розрідженими векторами $\mathbf{S}_{\text{роз}}$, що входять до складу клонів.

Враховуючи результати проведених комп'ютерних експериментів з оцінки якості класифікації зображень з використанням запропонованої моделі першого ступеня ансамблевого класифікатора зі стекінгом на основі клонів різних НМ, отриманих результатів використання операторів повороту і циклічного зсуву зображень запропоновано метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зображень зі стекінгом нейронних мереж, який складається з таких етапів.

Етап 1. Для заданого набору даних перевірити якість класифікації (*Accuracy*) відомих базових класифікаторів і вибрати найкращі для використання – $\mathbf{F}(\Theta)$. Враховуючи обчислювальну потужність процесора, на якому буде реалізовано ансамблевий класифікатор зі стекінгом НМ, і з необхідної частоти обробки вхідних зображень, визначаємо максимальну кількість класифікаторів N , які будуть реалізовані на першому ступені.

Етап 2. Для заданого набору даних побудувати клони вибраних базових класифікаторів з використанням оператора повороту вхідного зображення $\mathbf{T}$, для

кутів повороту вхідного зображення від 0 до t градусів, де $t < 1$, із кроком $\Delta = t/G$ градуса ($G > 20$ визначено в ході дослідження за експертними оцінками). Побудувати таблицю спільних помилок. У цій таблиці знайти рядок із найменшою сумою елементів (спільних помилок). Кут повороту, що відповідає цьому рядку, приймається за крок кута повороту Δ в операторі повороту $\mathbf{Q}(\mathbf{T}(\Delta))$ при побудові клонів.

Етап 3. Побудувати та навчити $N - 1$ клонів $\mathbf{F}_{\square}^T(\Theta)$ для кожного з вибраних базових класифікаторів з кутами повороту $\Delta, -\Delta, 2\Delta, -2\Delta, 3\Delta, -3\Delta$, тощо. Навчити повний ансамблевий класифікатор із використанням на першому ступені кожного вибраного базового класифікатора та одного клона з кутом Δ , потім вибраного базового класифікатора і двох клонів із кутами Δ і $-\Delta$, тощо.

Етап 4. Згенерувати C ($C > 50$, визначено в ході дослідження за експертними оцінками) розріджених векторів зсуву $\mathbf{S}_{\text{роз}}$, відповідно до оператору циклічного зсуву $\mathbf{Q}(\mathbf{S}_{\text{роз}})$, для кожного з вибраних базових класифікаторів. Побудувати та навчити для заданого набору даних із цими векторами зсуву C клонів для кожного з вибраних базових класифікаторів і вибрати з них $N - 1$ клонів $\mathbf{F}_{\square}^S(\Theta)$ із найбільшою якістю класифікації.

Етап 5. Навчити повний ансамблевий класифікатор із використанням на першому ступені кожного з обраних базових класифікаторів і одного клона, отриманого циклічним зсувом рядків, потім обраного базового класифікатора і двох клонів, тощо.

Етап 6. Порівняти результати етапів 3 і 5. Відповідно до результатів порівняння визначити наступні структурно-параметричні ознаки першого ступеня ансамблевого класифікатора зі стекінгом:

- типи використовуваних базових класифікаторів з їхніми власними параметрами $\mathbf{F}(\Theta)$;
- загальну кількість класифікаторів на першому ступені N ;
- типи базових класифікаторів, що входять до складу клонів $\mathbf{F}_{\square}^T(\Theta), \mathbf{F}_{\square}^S(\Theta)$;

– типи і параметри операторів перетворення, що входять до складу клонів $Q(T)$, $Q(S_{\text{роз}})$.

Модель ансамблевого класифікатора зі стекінгом на основі клонів НМ з використанням запропонованого методу побудови забезпечує:

- суттєве збільшення якості класифікації всього ансамблевого класифікатора за рахунок додавання на першу ступень необхідної кількості клонів;
- підвищення якості класифікації всього ансамблевого класифікатора за рахунок використання на першому ступені тільки найефективніших базових НМ та заміни інших клонами найбільш ефективних базових НМ.

Висновки до розділу 2

1. Набула подальшого розвитку модель першого ступеня ансамблевого класифікатора зі стекінгом НМ для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних. За рахунок додавання до заданої кількості базових класифікаторів їх клонів, навчених на трансформованих вибірках зображень, сформовано теоретичне обґрунтування нового методу визначення її структурно-параметричних параметрів.

2. Запропоновано принципово новий метод визначення структурно-параметричних ознак моделі першого ступеня ансамблевого класифікатора зображень зі стекінгом НМ. Завдяки інтеграції інваріантних операторів повороту та циклічного зсуву, метод дозволяє генерувати високоефективні ансамблі НМ в умовах низької роздільної здатності зображень та обмеженого обсягу навчальних вибірок.

3. На прикладі класифікаторів MLP-Mixer та CСТ показано, що поворот вхідних зображень при структурованій обробці в класифікаторі призводить до суттєвих відмінностей у тому, на яких зображеннях класифікатор робить помилки розпізнавання.

4. Показано, що збільшення класифікаторів у ширину шляхом підвищення кількості класифікаторів на першому ступені ансамблевого класифікатора забезпечує покращення якості класифікації порівняно зі збільшенням у глибину за

практично однакового обсягу обчислень. На прикладі алгоритму MLP-Mixer і набору даних CIFAR-10 таке покращення становило від 0.7% до 3.9%.

5. Показано, що при роботі з набором даних CIFAR-10 додавання до першого ступеня двох клонів алгоритму MLP-Mixer, отриманих оператором повороту вхідних зображень, забезпечило зменшення помилки від 1% до 11% залежно від кількості базових класифікаторів на першому ступені. Аналогічно для SST зменшення помилки становило від 2.1% до 10%.

6. Наведений у роботі підхід до вибору кроку в операторі повороту зображень при формуванні клонів для набору даних CIFAR-10 може бути використаний при роботі з будь-яким іншим набором даних.

7. На прикладі базових НМ MLP-Mixer, SST та набору даних CIFAR-10 показано, що запропонований метод формування клонів шляхом циклічних зсувів стовпців або рядків вхідних зображень забезпечує суттєво відмінні між собою множини зображень, на яких базова НМ та її клони роблять помилки. Це призводить до підвищення результуючої якості класифікації ансамблевого класифікатора зі стекінгом при додаванні клонів на першому ступені.

8. При використанні на першому ступені ансамблевого класифікатора клонів MLP-Mixer, побудованих шляхом циклічного зсуву рядків або стовпців вхідних зображень, збільшення кількості клонів від 0 до 12 забезпечило підвищення якості класифікації ансамблевого класифікатора з 0.7674 до 0.8029, тобто на 4.6%. Відповідно, кількість помилок при цьому зменшилася з 2326 до 1971 або на 15%. Для клонів SST підвищення якості класифікації становило з 0.8021 до 0.8492, тобто на 5.9%, а кількість помилок при цьому зменшилася з 1979 до 1508, або на 23.8%.

9. Порівняння з клонами, побудованими оператором повороту зображень, показало, що клони MLP-Mixer, побудовані за допомогою зсувів, дають кращу результуючу якість класифікації на 0.7% – 1% лише при невеликій кількості базових НМ на першому ступені. Водночас клони SST, побудовані з використанням зсувів, забезпечили перевагу порівняно з клонами MLP-Mixer (із поворотом) для всіх досліджених комбінацій базових НМ на 2% – 3.5%.

3 РОЗРОБКА МОДЕЛІ MLP-KAN ТА ЇЇ ВИКОРИСТАННЯ НА ДРУГОМУ СТУПЕНІ АНСАМБЛЕВОГО КЛАСИФІКАТОРА ЗІ СТЕКІНГОМ

Другий ступінь ансамблевого класифікатора зі стекінгом об'єднує результати використання НМ першого ступеня та формує вихідні значення. Використання моделі MLP на другому ступені забезпечує вищу якість класифікації порівняно з іншими методами [19]

Підвищення ефективності ансамблевого класифікатора зі стекінгом НМ передбачає вдосконалення моделей всіх його складових частин на першому і другому ступенях. У другому розділі було запропоновано модель та методи підвищення ефективності роботи першого ступеня ансамблевого класифікатора.

У третьому розділі розглядається можливість зменшення обсягу обчислень на другому ступені ансамблевого класифікатора зі стекінгом НМ без погіршення якості класифікації.

Представлені в цьому розділі результати базуються на працях [3], [4].

3.1 Модель нейронної мережі MLP-KAN для задач класифікації загального типу

Практика побудови НМ для задач класифікації загального типу показала, що, як правило, незалежно від моделі НМ, на її виході знаходиться MLP, який формує вихідні мітки класів. Це є наслідком того, що він є універсальним апроксиматором [35].

У роботі [40] була запропонована альтернативна архітектура KAN-мережі Колмогорова-Арнольда, яка також є універсальним апроксиматором [60]. Як правило, якість апроксимації певної функції та необхідний обсяг обчислень залежать від виду цієї функції й від використовуваного методу апроксимації.

MLP використовує багатовхідні суматори з навчальними ваговими коефіцієнтами на вході та нелінійною функцією на виході, які називаються

нейронами. Нелінійна функція на виході (функція активації) є однаковою для всіх нейронів шару в MLP. Усі функції активації в MLP не змінюються в процесі навчання.

У KAN на входах суматора замість навчальних вагових коефіцієнтів використовуються навчальні нелінійні функції (функції активації), і на виходах суматорів немає фіксованих функцій активації, як у звичайних нейронах.

Теореми, що стосуються апроксимальних властивостей універсальних апроксиматорів [35, 40], стверджують, що за їх допомогою можна апроксимувати будь-яку функцію з будь-якою заданою точністю. Однак при цьому можуть знадобитися дуже великі обчислювальні витрати. Малий обсяг обчислень досягається лише для функцій, які найкраще підходять для даного типу апроксимації. Оскільки, для НМ, цікавим є загальний випадок апроксимації довільних функцій, *доцільно одночасно використовувати властивості апроксимації нейронів і навчальних функцій активації.*

У класичному виді, при вирішенні задачі класифікації об'єктів на зображеннях, шар моделі KAN [40] передбачає, що кожен із V пікселів вхідного зображення перетворюється за допомогою U навчальних функцій активації. У результаті отримуються $V \times U$ результатів перетворення, які надходять на U суматорів. Виходи цих суматорів і є виходами шару KAN.

Для поєднання апроксимальних властивостей MLP і KAN в моделі MLP [85], запропоновано перший прихований шар замінити на шар KAN. *Заміна першого прихованого шару MLP на шар KAN дозволила запропонувати нову модель нейронної мережі MLP-KAN.* Загальний вигляд цієї моделі наведено (рис. 3.1).

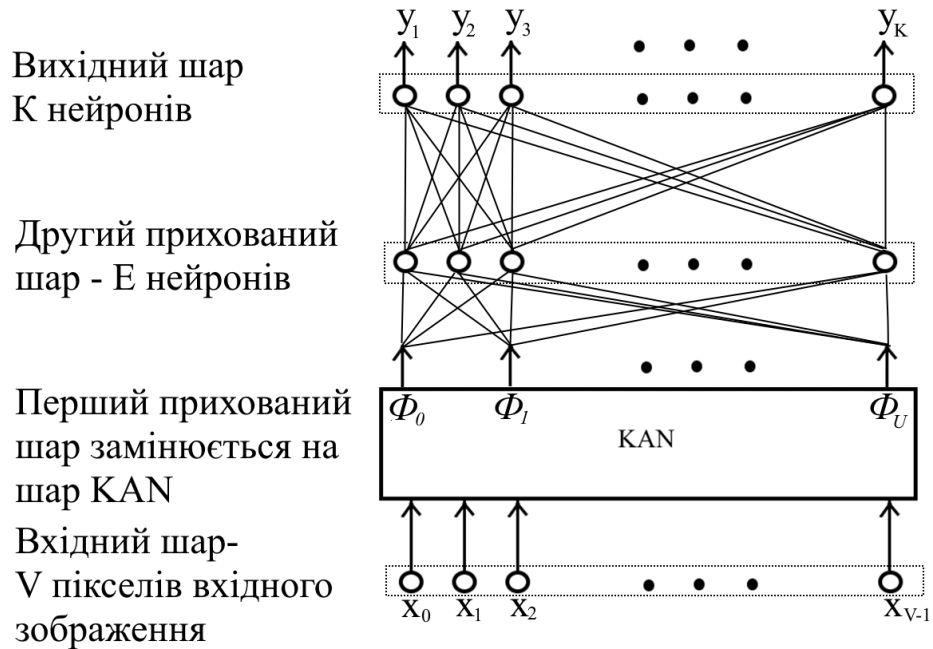
Розглянемо особливості побудови архітектури шару KAN.

В роботі [76] були розглянуті універсальні апроксимальні властивості RBF. Враховуючи простоту їх реалізації у вигляді заздалегідь обчисленої таблиці значень, радіальні базисні функції було використано в цієї роботі як базис для навчальних функцій активації шару KAN, який формується шляхом зсуву центрів RBF.

RBF мають такий вигляд [76]:

$$\varphi_i(x_l) = e^{-((x_l - x_{min} + i \times d)^2 / (2 \times b^2))}, \quad (3.1)$$

де $i=0, \dots, (m-1)$ – номер функції (m – загальне число використовуваних RBF);
 x_l – l -ий піксель вхідного зображення;
 x_{min} – мінімальне значення пікселів вхідного зображення;
 $d = (x_{max} - x_{min}) / (m - 1)$ – відстань між центрами сусідніх RBF;
 x_{max} – максимальне значення пікселів вхідного зображення, після нормалізації дорівнює 1;
 $b = (x_{max} - x_{min}) / 8.45$ – коефіцієнт, що задає ширину RBF.



V – кількість пікселів у вхідному зображенні;
 U – кількість виходів шару KAN;
 E – кількість нейронів у другому прихованому шарі;
 K – кількість класів об'єктів на вхідних зображеннях

Рисунок 3.1 – Модель MLP-KAN, що утворена заміною першого прихованого шару базової моделі MLP на шар KAN

Першою операцією під час реалізації шару KAN є обчислення для кожного пікселя x_l вхідного зображення m значень для відповідних функцій φ_i .

Типовий вигляд гаусових RBF – φ_i для $m = 6$ показано на рис. 3.2.

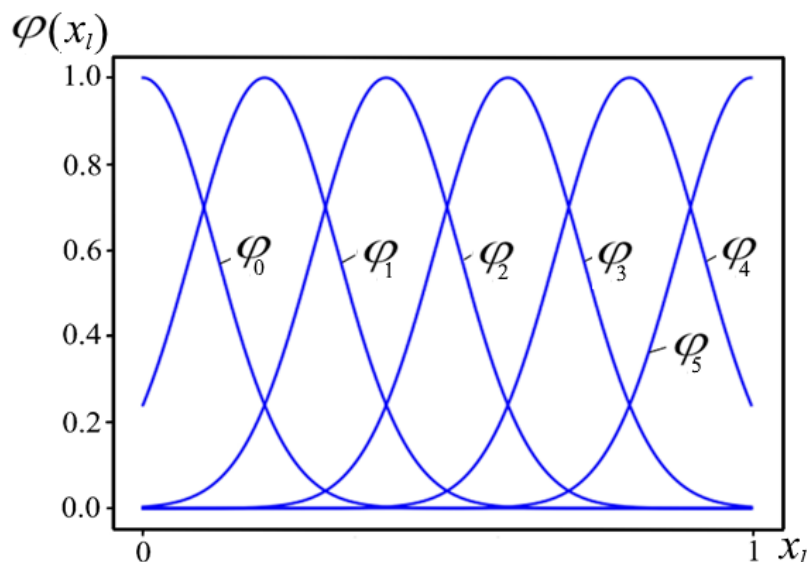


Рисунок 3.2 – Типовий вигляд функцій φ_i для $m = 6$

Для реалізації класичного шару KAN для кожного пікселя вхідного зображення необхідно сформувати E навчальних функцій активації (де E – кількість вузлів у наступному шарі), кожна з яких має вигляд:

$$\Phi_l(x_l) = \sum_{i=0}^{m-1} \varphi_i(x_l) \times \omega_{li}, \quad (3.2)$$

де Φ_l – функція активації для l -ого пікселя вхідного зображення;

x_l – l -ий піксель вхідного зображення (нормований);

ω_{li} – i -й навчальний ваговий коефіцієнт для формування функції активації для l -го пікселя вхідного зображення.

Відповідно до формули (3.2), елемент шару KAN, який формує навчальну функцію активації, представлений на рис. 3.3.

Оперування з ваговими коефіцієнтами ω_{li} шару KAN можна здійснювати так само, як і з ваговими коефіцієнтами в MLP. Це дозволяє уніфіковано навчати модель засобами бібліотеки Keras. В бібліотеці Keras для реалізації навчаємих вагових коефіцієнтів та суматору використовується шар *Dense*.

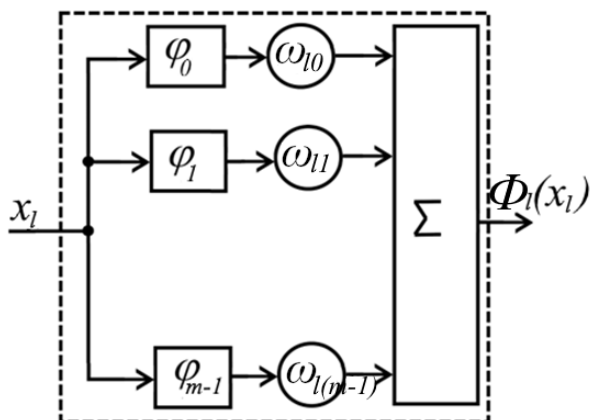


Рисунок 3.3 – Схема обчислення однієї навчальної функції активації шару KAN з використанням RBF для пікселя x_l вхідного зображення

Однак реалізація такої кількості функцій активації, як це потрібно для класичного шару KAN для зображення дорівнює $V \times E$, де E – кількість вузлів у наступному шарі, що є надмірним.

Запропонована архітектура шару KAN дозволяє об'єднувати довільну кількість функцій активації та мати в шарах *Dense* не лише один нейрон, а й більше. Кількість функцій RBF може змінюватися для досягнення кращої якості класифікації відповідних типів вхідних зображень.

З метою зменшення підсумкового обсягу обчислень, в даній роботі запропоновано обчислювати лише одну функцію активації для кожного пікселя, що потребує додаткового дослідження якості класифікації від кількості RBF у MLP-KAN та від параметра об'єднання функцій активації.

3.1.1 Побудова моделі нейронної мережі MLP-KAN для набору даних MNIST

Для набору даних MNIST з розміром зображень 28×28 , кількість пікселів у зображенні становить 784, після операції перетворення за допомогою функцій RBF отримуємо $784 \times m$ значень.

Для врахування взаємного впливу значень на виходах функцій активації будемо налаштовувати вагові коефіцієнти не для кожної функції активації окремо, а одразу для 2-х, 4-х, 8-ми тощо функцій активації.

При цьому вихід шару *Dense*, який формує об'єднану функцію активації для пікселів із номерами від n_1 до n_2 , має такий вигляд:

$$F_{n_1-n_2} = \left(\begin{aligned} &\sum_{i=0}^{m-1} \varphi_i(x_{n_1}) \cdot \omega_{n_1 i} + \\ &+ \sum_{i=0}^{m-1} \varphi_i(x_{n_1+1}) \cdot \omega_{(n_1+1)i} + \dots + \\ &+ \sum_{i=0}^{m-1} \varphi_i(x_{n_2}) \cdot \omega_{n_2 i} \end{aligned} \right). \quad (3.3)$$

Приклад архітектури KAN-шару для набору MNIST з об'єднанням по дві функції активації, використанням 6 функцій RBF та одного нейрона в шарі *Dense* наведено на рис. 3.4.

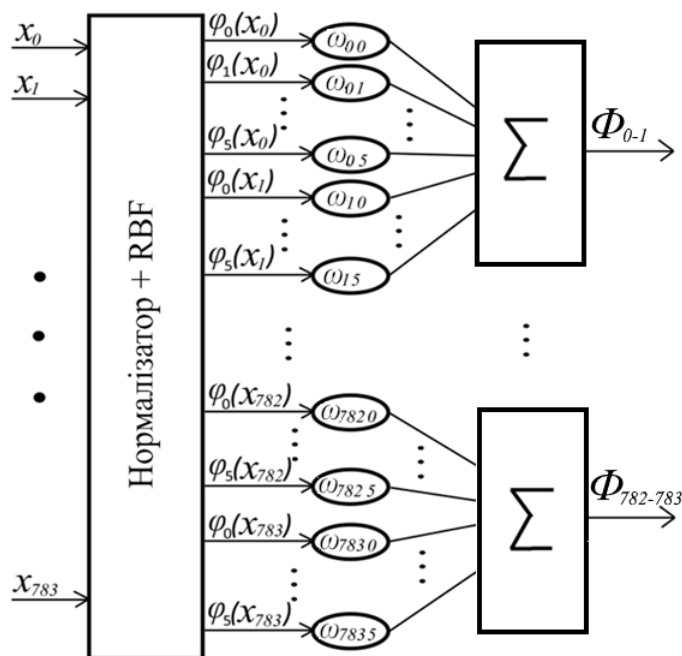


Рисунок 3.4 – Архітектура KAN-шару з об'єднанням по дві функції активації в моделі MLP-KAN для набору даних MNIST

На рис. 3.4 на вхід шару KAN надходять 784 пікселі вхідного зображення $x_0, \dots, x_{783}$. Далі вони нормалізуються та перетворюються функціями RBF. Це дає

784×6 значень, які множаться на вагові коефіцієнти ω_{li} і підсумовуються у відповідних суматорах. Сукупність вагових коефіцієнтів, об'єднаних із суматором, становить шар *Dense*, в реалізації якого для шару KAN не використовується функція активації.

В цьому прикладі в шарі KAN використано 392 шари *Dense*, кожен із яких містить один нейрон і має 12 входів із відповідними навчальними вагами. 392 виходи шару KAN далі надходять на другий прихований шар MLP, а з нього – на вихідний шар (рис. 3.5). На кожний нейрон другого прихованого шару надходять 392 вихідних сигналів шару KAN зі своїми ваговими коефіцієнтами.

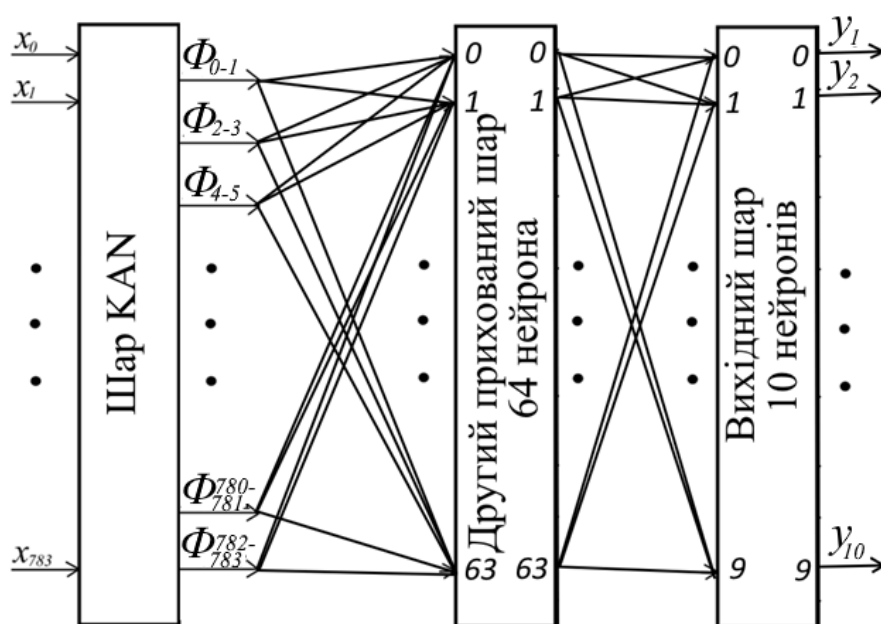


Рисунок 3.5 – Модель MLP-KAN з об'єднанням по 2 функції активації шару KAN для набору зображень MNIST

Обчислення значень RBF від вхідних зображень не потребує виконання арифметичних операцій, і в багатьох практичних ситуаціях відпадає необхідність у нормалізації вхідних зображень. Достатньо попередньо обчислити та зберегти таблиці значень, після чого просто здійснювати пошук потрібного значення в таблиці.

Наприклад, для набору MNIST значення пікселів задаються 8 бітами, тому для знаходження однієї функції RBF достатньо мати в пам'яті таблицю із 256

значень. Якщо для апроксимації використовується 6 функцій RBF, то необхідно мати 6 таблиць, у кожній з яких по 256 значень.

Можлива заміна функцій RBF іншим видом базисних функцій, із збереженням принципу реалізації у вигляді таблиць, але потрібна додаткова оцінка якості класифікації для підтвердження раціональності такої заміни.

3.1.2 Дослідження залежності якості класифікації зображень моделлю MLP-KAN від кількості RBF

Кількість можливих комбінацій параметрів НМ MLP-KAN є достатньо великою. Як було підкреслено вище, кількість функцій RBF, що використовуються для апроксимації однієї функції активації, може впливати на якість класифікації відповідних типів вхідних зображень. Для підтвердження цього проведемо обчислювальні експерименти.

Дослідження проведено на наборі даних MNIST для трьох різних варіантів. Результати наведено в табл. 3.1–3.3. При цьому використано такі позначення:

- 1) k_o – коефіцієнт об'єднання функцій активації в шарі KAN;
- 2) k_n – кількість нейронів у шарах *Dense* в реалізації шару KAN;
- 3) E – кількість нейронів у прихованому шарі мережі MLP-KAN (другий прихований шар MLP, який залишився в структурі MLP-KAN);
- 4) m – кількість використаних функцій RBF.

Таблиця 3.1 – Результати оцінювання якості класифікації моделлю MLP-KAN об'єктів на зображеннях та кількості навчальних вагових коефіцієнтів при $k_o = 2, k_n = 2, E = 64$

m	<i>Accuracy</i>	Кількість вагових коефіцієнтів, що навчаються
3	0.9712	55594
4	0.9741	57162
5	0.9739	58730
6	0.9749	60298
7	0.9764	61866
8	0.9750	63434

Таблиця 3.2 – Результати оцінювання якості класифікації моделлю MLP-KAN об'єктів на зображеннях та кількості навчальних вагових коефіцієнтів при $k_o = 4, k_n = 1, E = 64$

m	<i>Accuracy</i>	Кількість вагових коефіцієнтів, що навчаються
3	0.9758	15610
4	0.9767	16394
5	0.9753	17178
6	0.9774	17962
7	0.9756	18746
8	0.9752	19530

Таблиця 3.3 – Результати оцінювання якості класифікації моделлю MLP-KAN об'єктів на зображеннях та кількості навчальних вагових коефіцієнтів при $k_o = 8, k_n = 1, E = 64$

m	<i>Accuracy</i>	Кількість вагових коефіцієнтів, що навчаються
2	0.9729	8554
3	0.9750	9338
4	0.9755	10122
5	0.9751	10906
6	0.9741	11960
7	0.9737	12474

Кількість вагових коефіцієнтів моделі, що навчаються, обчислюється при виконанні програми за допомогою стандартної бібліотечної функції `model.summary()` бібліотеки Keras.

Як видно з табл. 3.1–3.3, при класифікації зображень набору даних MNIST, у першому варіанті найкраща якість класифікації була отримана при кількості використаних RBF – $m=7$, у другому варіанті – при $m=6$, у третьому варіанті – при $m=4$. Тому для кращого розуміння залежностей у наступному пункті продовжимо дослідження використання MLP-KAN з параметром $m=6$.

3.1.3 Дослідження залежності якості класифікації зображень моделлю MLP-KAN від параметра об'єднання функцій активації

Вище у п. 3.1.1 було наведено приклад реалізації моделі MLP-KAN з об'єднанням по 2 функції активації шару KAN для набору зображень MNIST.

Доцільно проаналізувати залежність якості класифікації об'єктів на зображеннях від кількості функцій активації, що об'єднуються.

Проведено комп'ютерний експеримент з класифікації зображень набору MNIST з різними коефіцієнтами об'єднання функцій активації в шарі KAN. Результати дослідження якості класифікації та обсягу обчислень залежно від коефіцієнта об'єднання функцій активації наведено в табл. 3.4.

Кількість вагових коефіцієнтів моделі, що навчаються, обчислюється при виконанні програми за допомогою стандартної бібліотечної функції `model.summary()` бібліотеки Keras.

У табл. 3.4 використано такі позначення:

- 1) k_o – коефіцієнт об'єднання функцій активації в шарі KAN;
- 2) k_n – кількість нейронів у шарах *Dense* шару KAN (1 або 2);
- 3) E – кількість нейронів у другому прихованому шарі (64 або 96).

Усі результати були отримані при кількості використаних функцій RBF $m = 6$. Крім цього, у табл. 3.4 наведено співвідношення кількості навчальних коефіцієнтів для MLP (55050) і для MLP-KAN k_ω (виграш в кількості обчислень), а також співвідношення якості класифікації MLP-KAN і MLP (0.9782) k_{qc} (виграш в якості класифікації).

$$k_\omega = R_{MLP} / R_{MLP-KAN}, \quad (3.4)$$

де $R_{MLP}=55050$ - кількості навчальних вагових коефіцієнтів для MLP,
 $R_{MLP-KAN}$ - кількості навчальних вагових коефіцієнтів для MLP-KAN.

$$k_{qc} = Accuracy_{MLP-KAN} / Accuracy_{MLP}, \quad (3.5)$$

де $Accuracy_{MLP} = 0.9782$ - якість класифікації MLP,

$Accuracy_{MLP-KAN}$ - якість класифікації MLP-KAN .

Таблиця 3.4 – Результати оцінювання якості класифікації моделлю MLP-KAN об'єктів на зображеннях та обсягу обчислень

k_o	k_n	E	Кількість вагових коефіцієнтів, що навчаються ($R_{MLP-KAN}$)	k_ω	$Accuracy_{MLP-KAN}$	k_{qc}
2	1	64	30506	1.80	0,9763	0.9981
2	1	96	43402	1.27	0.9784	1.0002
2	2	64	60298	0.91	0.9749	0.9966
2	2	96	87738	0.64	0.9776	0.9994
4	1	64	17962	3.06	0.9774	0.9992
4	1	96	24586	2.24	0.9793	1.0011
4	2	64	35210	1.56	0.9769	0.9987
4	2	96	48106	1.14	0.9805	1.0024
8	1	64	11690	4.70	0.9741	0.9958
8	1	96	15178	3.63	0.9782	1.0000
8	2	64	22666	2.43	0.9780	0.9998
8	2	96	29290	1.88	0.9782	1.0000
16	1	64	8554	6.44	0.9728	0.9945
16	1	96	10474	5.26	0.9744	0.9961
16	2	64	16394	3.36	0.9752	0.9969
16	2	96	19882	2.77	0.9786	1.0004

Із табл. 3.4 видно, що при тій самій якості класифікації 0.9782, що й у MLP, мережа MLP-KAN може мати у 3.63 рази менший обсяг обчислень. Крім того, може бути досягнута вища якість класифікації на 0.24 % при меншому на 14 % обсязі обчислень. Найкраще співвідношення кількості навчальних коефіцієнтів і якості класифікації було отримано при коефіцієнті об'єднання функцій активації $k_o = 8$. Збільшення кількості нейронів у шарах *Dense* шару KAN при малих k_o призводило до погіршення якості класифікації, а при великих k_o – до її покращення. Збільшення кількості нейронів у прихованому шарі завжди призводило до підвищення якості класифікації.

Для порівняння швидкості навчання на рис. 3.6 наведено криві навчання для моделей MLP та MLP-KAN.

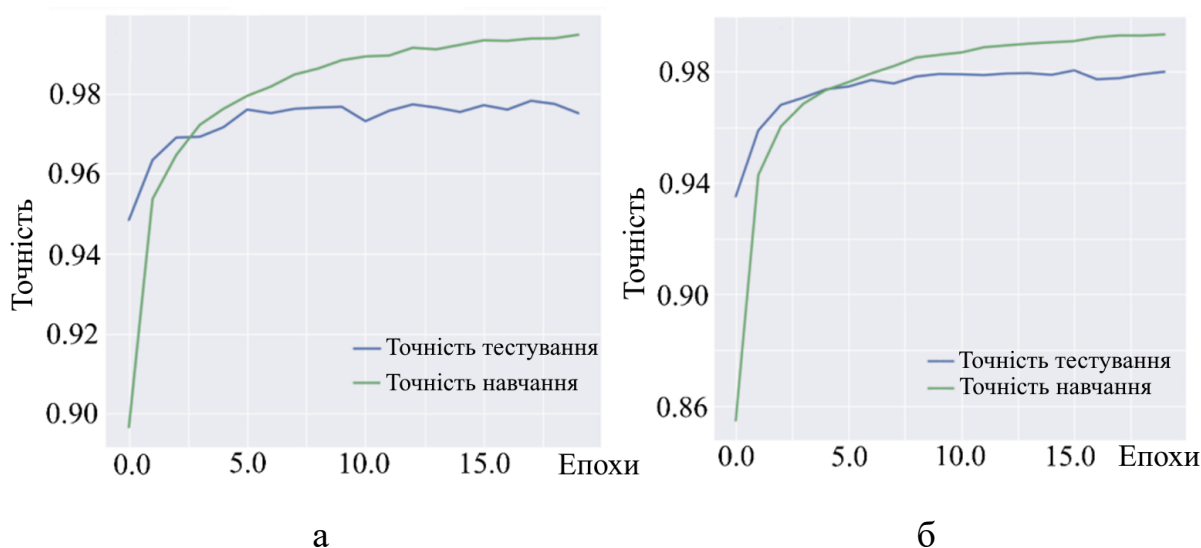


Рисунок 3.6 – Криві навчання: а – для MLP; б – для MLP-KAN

Вісь Y на рис. 3.6 показує точність класифікації (*Accuracy*), вісь X – кількість епох навчання (*Epochs*). Навчання проводилося за навчальними даними (*Training Accuracy*). За тестовими даними виконувалося лише обчислення результату при поточних вагових коефіцієнтах. Із рис. 3.6 видно, що швидкості навчання MLP та MLP-KAN практично ідентичні. Максимальна точність класифікації для MLP (рис. 3.5, а) становить 0.9782, для MLP-KAN (рис. 3.5, б) – 0.9805.

3.1.4 Обговорення результатів дослідження використання моделі MLP-KAN при класифікації об'єктів на зображеннях

Проведене дослідження використання моделі MLP-KAN для класифікації об'єктів на зображеннях дозволяє зробити висновок, що запропонована архітектура реалізації KAN-шару відрізняється від відомих суттєвою гнучкістю, а саме:

1) використання RBF (див. (3.1) та рис. 3.2) є універсальним, але не обмежувальним варіантом. Замість RBF може бути використаний будь-який інший базис, якщо він краще відповідає характеристикам практичної задачі;

2) кількість об'єднаних функцій активації (див. (3.3) і рис. 3.3) може бути підібрана для досягнення найкращого співвідношення між якістю класифікації та обсягом обчислень для конкретного застосування;

3) реалізація навчальних функцій активації у вигляді виразів (3.1) і (3.2) дозволяє легко реалізовувати їх за допомогою традиційних програмних бібліотек, що використовуються для побудови наступних шарів. Таке проектування моделі MLP-KAN дозволяє ефективно визначати найрезультативніші варіанти побудови при розв'язанні практичних задач.

Дослідження ефективності використання різної кількості функцій під час апроксимації функцій активації (див. табл. 3.1 – 3.3) виявило складний характер залежності якості класифікації від кількості функцій для різних параметрів KAN-шару. Для набору даних MNIST та використання RBF як раціональний варіант, що спрощує інтерпретацію подальших експериментів, було прийнято $m = 6$. Проте будь-яке практичне застосування MLP-KAN повинно включати додаткові дослідження для визначення раціональної кількості базисних функцій.

Результати експериментального дослідження (див. табл. 3.4) показують, що заміна першого прихованого шару в MLP на KAN-шар при відповідному виборі параметрів дозволяє отримати вищу якість класифікації при суттєво меншому обсязі обчислень.

Зокрема, при однаковій якості класифікації 0.9782 обсяг обчислень для MLP-KAN був у 3.63 рази менший, ніж для MLP. При більш високій точності класифікації MLP-KAN (0.9786) виграв в обсязі обчислень становив 2.77 рази. Крім того, слід відзначити, що запропонована модифікація MLP-KAN навчається з такою ж швидкістю, як і MLP (див. рис. 3.6), на відміну від сповільнення, отриманого в роботі [86].

Отримані результати пояснюються тим, що НМ MLP-KAN одночасно використовує два типи апроксимації, а саме, використання вагових коефіцієнтів

нейронів і навчальних функцій активації. Необхідно підкреслити, що результати отримані на наборі даних загального типу (MNIST), що дозволяє сподіватися на аналогічну ефективність і на інших наборах даних подібного типу.

Додатковою перевагою запропонованої НМ MLP-KAN є те, що вона може бути реалізована за допомогою широко використовуваної бібліотеки Keras, на відміну від спеціальної бібліотеки, запропонованої в [40].

Особливістю моделі MLP-KAN є використання спільної обробки невеликої кількості функцій активації. Найкращі результати були досягнуті при спільній обробці 8 функцій активації. Водночас класична реалізація KAN, запропонована в [40], передбачає спільну обробку у вузлах наступного шару власних функцій активації від кожного вузла поточного шару. Модель MLP-KAN при відповідному виборі параметрів забезпечує суттєво менший обсяг обчислень у порівнянні з MLP при збереженні якості класифікації.

Таким чином, запропонована модель MLP-KAN успішно вирішує задачу спільного використання апроксимальних властивостей MLP і KAN для набору даних загального типу MNIST. У цій роботі використовувалася апроксимація функцій активації за допомогою RBF. Однак у межах запропонованої моделі можуть бути використані й інші функції апроксимування, якщо вони краще відображають властивості конкретного набору даних. Виконувати обчислення цих функцій на кожній ітерації не потрібно. Достатньо попередньо розрахувати таблиці значень із достатньою кількістю відліків і надалі просто вибирати значення з таблиць.

Обмеженням запропонованої архітектури є те, що вона побудована на базі MLP. MLP використовується в різних застосуваннях як самостійно, так і як вихідний елемент практично всіх складних НМ. Проте існують моделі, ефективніші за MLP, тому цікавить розвиток запропонованого підходу і в цих моделях.

3.2 Вдосконалення моделі нейронної мережі для другого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж

У підпункті 3.1 була запропонована архітектура MLP-KAN, що поєднує апроксимальні властивості нейронів і навчальних функцій активації. Порівняння з MLP здійснювалося на прикладі набору даних загального типу – MNIST.

Особливістю роботи другого ступеня ансамблевого класифікатора зі стекінгом НМ є те, що на його входи надходять вихідні вектори НМ першого ступеня. Особливістю цих векторів є те, що вони складаються з елементів, що містять оцінки ймовірності належності вхідного зображення до кожного з класів.

Максимальне значення елементу вектору вказує на найімовірніший клас, якому належить зображення. Інші компоненти містять менші, але ненульові ймовірності належності зображень до відповідних класів. Тому становить інтерес адаптація розробленої моделі MLP-KAN, спеціально пристосованої для використання на другому ступені ансамблевого класифікатора зі стекінгом НМ та дослідження її ефективності.

3.2.1 Використання нейронної мережі MLP-KAN для другого ступеня ансамблевого класифікатора зі стекінгом НМ

Для використання MLP-KAN на другому ступені ансамблевого класифікатора зі стекінгом НМ виконаємо такі кроки адаптації:

- кожен вхідний вектор елементів другого ступеня перетворюється за допомогою m функцій RBF;
- результати зважуються навчальними ваговими коефіцієнтами та подаються на суматор;
- вилучається другий прихований шар MLP;
- не виконується об'єднання функцій активації.

Така адаптація здійснена також з урахуванням особливостей вихідного вектору елементів першого ступеня ансамблевого класифікатора, а саме того, що вихідні вектори НМ першого ступеня ансамблевого класифікатора зі стекінгом НМ

мають явно виражений максимум на найбільш ймовірному класі, інші значення елементів векторів суттєво менші.

Як було визначено у розділі 2 на вхід класифікатору другого ступеня надходять вихідні елементи (ймовірності належності зображення кожному з класів K) у вигляді вектору $\mathbf{P}$, який складається з елементів на виході першої НМ першого ступеня – $p_{11}, \dots, p_{1K}$, другої НМ – $p_{21}, \dots, p_{2K}$, і так далі – до вихідних елементів $(M+L)$ -ої НМ першого ступеня – $p_{(M+L)1}^{\square}, \dots, p_{(M+L)K}^{\square}$. Кожен із цих елементів на другому ступені ансамблевого класифікатору зі стекінгом НМ перетворюється власною навчальною функцією активації та подається на кожен із K нейронів вихідного шару $n_1, \dots, n_K$.

Звичайно на другому ступені ансамблевого класифікатору зі стекінгом НМ використовується мета-класифікатор $\mathbf{H}$ (1.3), зазвичай це повнозв'язкова НМ MLP, яка виконує нелінійне перетворення мета-ознак і має такий вигляд

$$\hat{y} = \sigma(W_2 \times \text{ReLU}(W_1 \times \mathbf{P} + b_1) + b_2), \quad (3.6)$$

де W_2 – вагові коефіцієнти вихідного шару, що навчається;
 W_1 – вагові коефіцієнти прихованого шару, що навчається;
 b_1, b_2 – зміщення прихованого та вихідного шарів, що навчаються;
 σ – функція активації *Softmax*;
 ReLU – функція активації прихованого шару.

Пропонується замінити в моделі другого ступеня ансамблевого класифікатора за стекінгом НМ класифікатор MLP на адаптовану модель MLP-KAN, тоді вихідний вектор ймовірностей належності зображення кожному з K класів визначається таким чином

$$\hat{y} = \sigma(W_2 \times \Phi(\mathbf{P}) + b_2), \quad (3.7)$$

де $\Phi = \{\Phi_{11}, \dots, \Phi_{(M+L)K}\}$ – оператор перетворення кожного елементу вектору ймовірностей P .

Відповідно до введених кроків адаптації в моделі формується лише одна навчальна функція активації для кожного елемента вхідного вектору P другого ступеня ансамблевого класифікатора зі стекінгом НМ. Із виходів суматорів вектори ознак надходять на вихідний шар НМ, який складається зі звичайних нейронів із ваговими коефіцієнтами на входах, суматором і функцією активації *softmax*. Навчальна функція активації шару KAN класифікатора другого ступеня має вигляд:

$$\Phi_{jk}(p_{jl}) = \sum_{i=0}^{m-1} \varphi_{ji}(p_{jl}) \times \omega_{jli}, \quad (3.8)$$

де Φ_{jl} – функція активації для l -го вихідного елементу j -ої НМ першого ступеня ансамблевого класифікатора;

p_{jl} – l -ий вихідний елемент j -ої НМ першого ступеня ансамблевого класифікатора (нормований);

ω_{jli} – i -й навчальний ваговий коефіцієнт для формування функції активації для l -го вихідного елементу j -ої НМ першого ступеня ансамблевого класифікатора.

Відповідно до формули (3.8), елемент моделі НМ MLP-KAN другого ступеня ансамблевого класифікатора зі стекінгом НМ, який формує навчальну функцію активації, має вигляд, поданий на рис. 3.7.

Модель НМ другого ступеня ансамблевого класифікатора для K класів зображень наведена на рис. 3.8. У кожному нейроні вихідного шару всі вхідні вектори зважуються навчальними ваговими коефіцієнтами, підсумовуються, перетворюються функцією активації *softmax* і подаються на вихід. На суматор у нейронах також подається значення зсуву (англ. *bias*), зважене власним навчальним коефіцієнтом. Вихідний елемент того нейрона, який має максимальне значення, приймається як відповідь НМ другого ступеня, і, відповідно – як вихідний клас усього ансамблевого класифікатора зі стекінгом НМ.

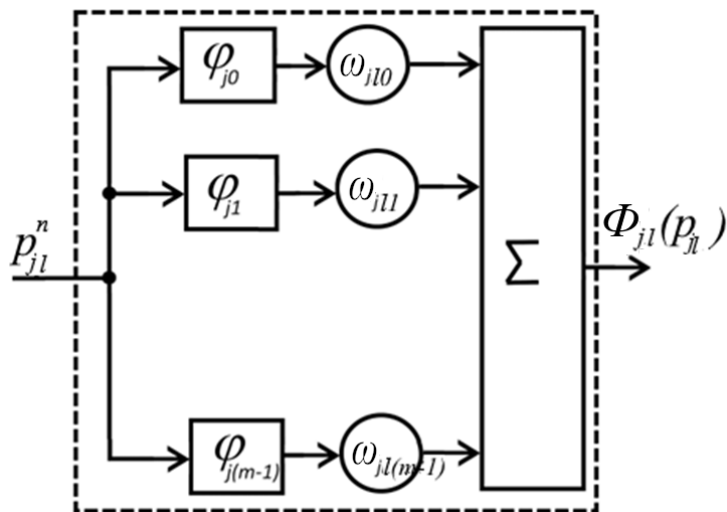


Рисунок 3.7 – Схема формування навчальної функції активації для вихідного елемента p_{jl} нейронної мережі першого ступеня ансамблевого класифікатора зі стекінгом НМ

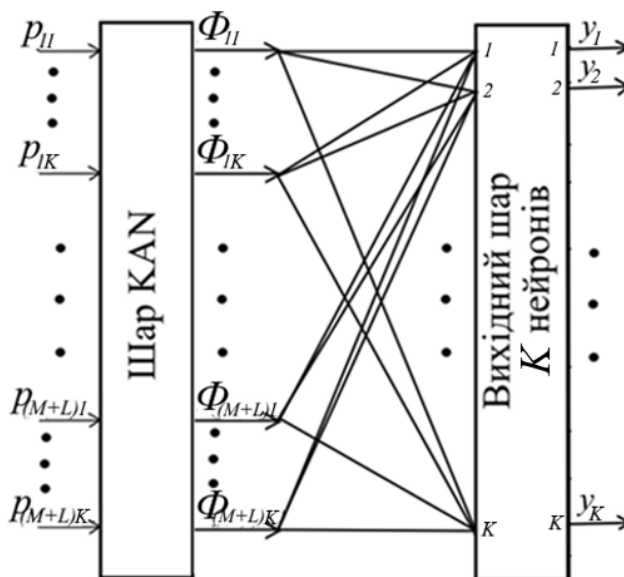


Рисунок 3.8 – Модель нейронної мережі на другому ступені ансамблевого класифікатора зі стекінгом НМ

Загальний вид вдосконаленої моделі ансамблевого класифікатора зі стекінгом НМ, що використовує на другому ступені MLP-KAN наведено на рис. 3.9. На відміну від моделі ансамблевого класифікатора зі стекінгом НМ, наведеної на рис.2.1, у другому ступені MLP замінено на MLP-KAN, що дозволило суттєво скоротити обсяг обчислень на другому ступені без погіршення якості класифікації об'єктів на зображеннях.

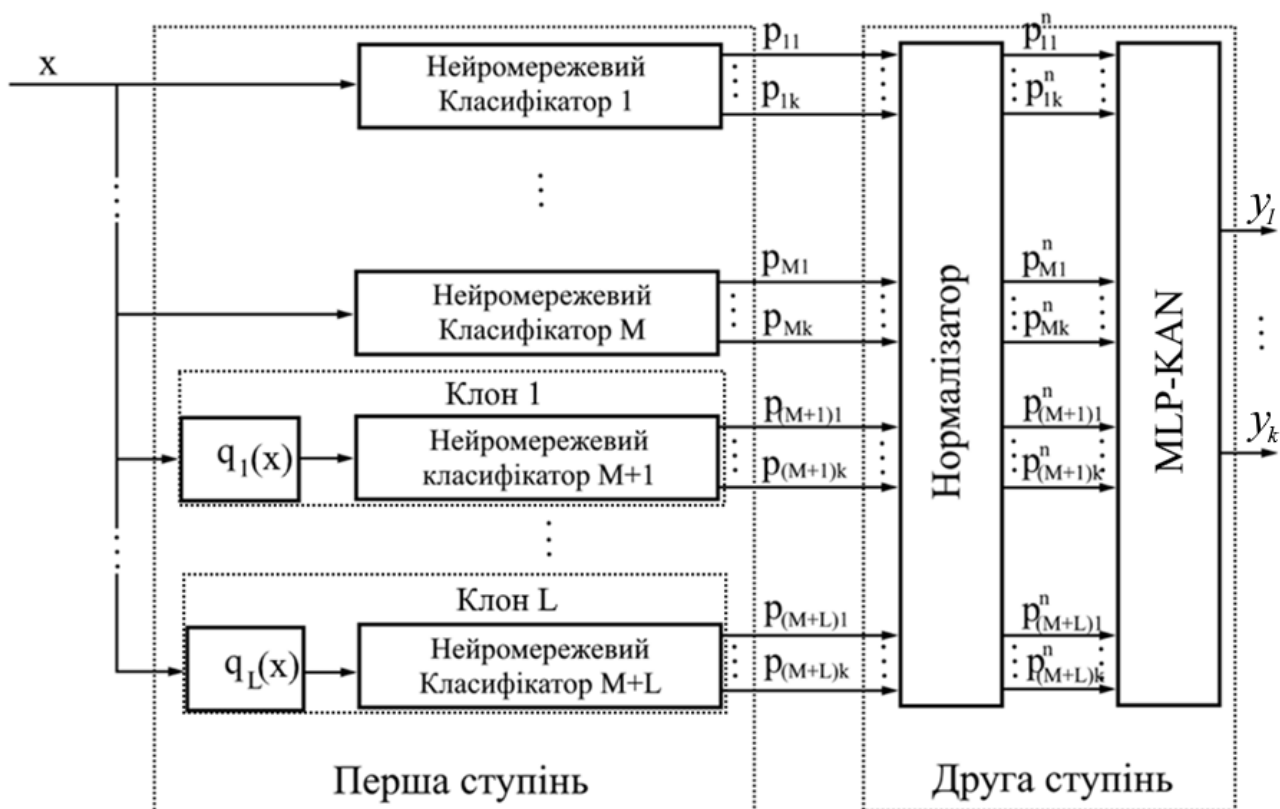


Рисунок 3.9 – Модель ансамблевого класифікатора зі стекінгом НМ, що використовує на другому ступені MLP-KAN

3.2.2 Дослідження обсягу обчислень при використанні моделі MLP-KAN на другому ступені ансамблевого класифікатора зі стекінгом НМ

Ефективність використання навчальних функцій активації в НМ другого ступеня порівнювалася з результатами роботи [40].

У цій роботі НМ другого ступеня містила три шари:

- вхідний шар розмірності $K \times (M+L)$ (K – кількість класів у наборі даних, $(M+L)$ – кількість НМ (базових класифікаторів та клонів) на першому ступені класифікатора);
- прихований шар розмірності $K \times (M+L - 1)$, функція активації – *ReLU*;
- вихідний шар розмірності K , функція активації – *softmax*.

У даному підрозділі замість прихованого шару НМ другого ступеня використовувався шар навчальних функцій активації. Кількість навчальних

коефіцієнтів у цьому шарі становила $(M+L) \times K \times m$. Порівняльна таблиця кількості навчальних вагових коефіцієнтів наведена в табл. 3.5.

Таблиця 3.5 – Кількість вагових коефіцієнтів, що навчаються у НМ другого ступеня ансамблевого класифікатора зі стекінгом НМ

Параметр	Значення				
Кількість класифікаторів на першому ступені ансамблевого класифікатора (M+L)	2	3	4	5	6
Кількість вагових коефіцієнтів, що навчаються у MLP [40] на другому ступені ансамблевого класифікатора	630	830	1540	2450	3560
Кількість вагових коефіцієнтів, що навчаються у MLP-KAN на другому ступені ансамблевого класифікатора при $m = 6$	321	481	641	801	961
Співвідношення кількості вагових коефіцієнтів	1.96	1.73	2.4	3.06	3.7

Як видно з табл. 3.5, використання моделі MLP-KAN на другому ступені ансамблевого класифікатора зі стекінгом НМ забезпечує зменшення кількості навчальних коефіцієнтів в 1.96 – 3.7 рази, залежно від кількості НМ (базових класифікаторів та клонів) на першому ступені.

3.2.3 Дослідження залежності якості класифікації від кількості RBF, що використовуються для апроксимації однієї функції активації в шарі KAN

Для дослідження було використано ансамблевий класифікатор із трьома НМ на першому ступені: CCT, EATNet і MLP-Mixer, на входи яких подавалися зображення з набору даних CIFAR-10 [39]. Результуюча якість класифікації для всього ансамблевого класифікатора після навчання наведена в табл. 3.6. Найкращі результати були отримані при $m = 6$, тому саме це значення кількості RBF було використано для подальшого дослідження.

Таблиця 3.6 – Залежність якості класифікації (*Accuracy*) від кількості RBF, що використовуються для апроксимації однієї функції активації в шарі KAN

Параметр	Значення				
Кількість RBF (m), що використовуються для апроксимації	4	5	6	7	8
<i>Accuracy</i>	0.8402	0.8423	0.8431	0.8430	0.8418

Для прикладу на рис. 3.10 наведено функції активації для вихідних елементів НМ ССТ після навчання другого ступеня ансамблевого класифікатора.

За формою навчених кривих активації видно, що після навчання на вихідний шар НМ другого ступеня подаються лише великі позитивні значення, а від’ємні та малі додатні значення обнуляються.

3.2.4 Порівняння ефективності використання моделі MLP-KAN з моделлю MLP на другому ступені класифікатора

Порівняння ефективності заміни прихованого шару MLP другого ступеня ансамблевого класифікатора зі стекінгом НМ на шар KAN з навчальними функціями активації проводилося для кількості НМ на першому ступені від 3 до 6. На вхід першого ступеня подавалися зображення з набору CIFAR-10. Кількість використаних RBF в усіх випадках становила $m = 6$. Результати наведено в табл. 3.7. Показники для використання MLP взяті зі статті [19].

Із результатів табл. 3.7 видно, що заміна прихованого шару в MLP на шар навчальних функцій активації призвела до невеликого зменшення кількості помилок класифікації – від 0,45% до 1,9%. При цьому в НМ другого ступеня, згідно з табл. 3.5, було використано суттєво меншу кількість навчальних коефіцієнтів, яка відрізнялася у 1,69 – 3,7 рази.

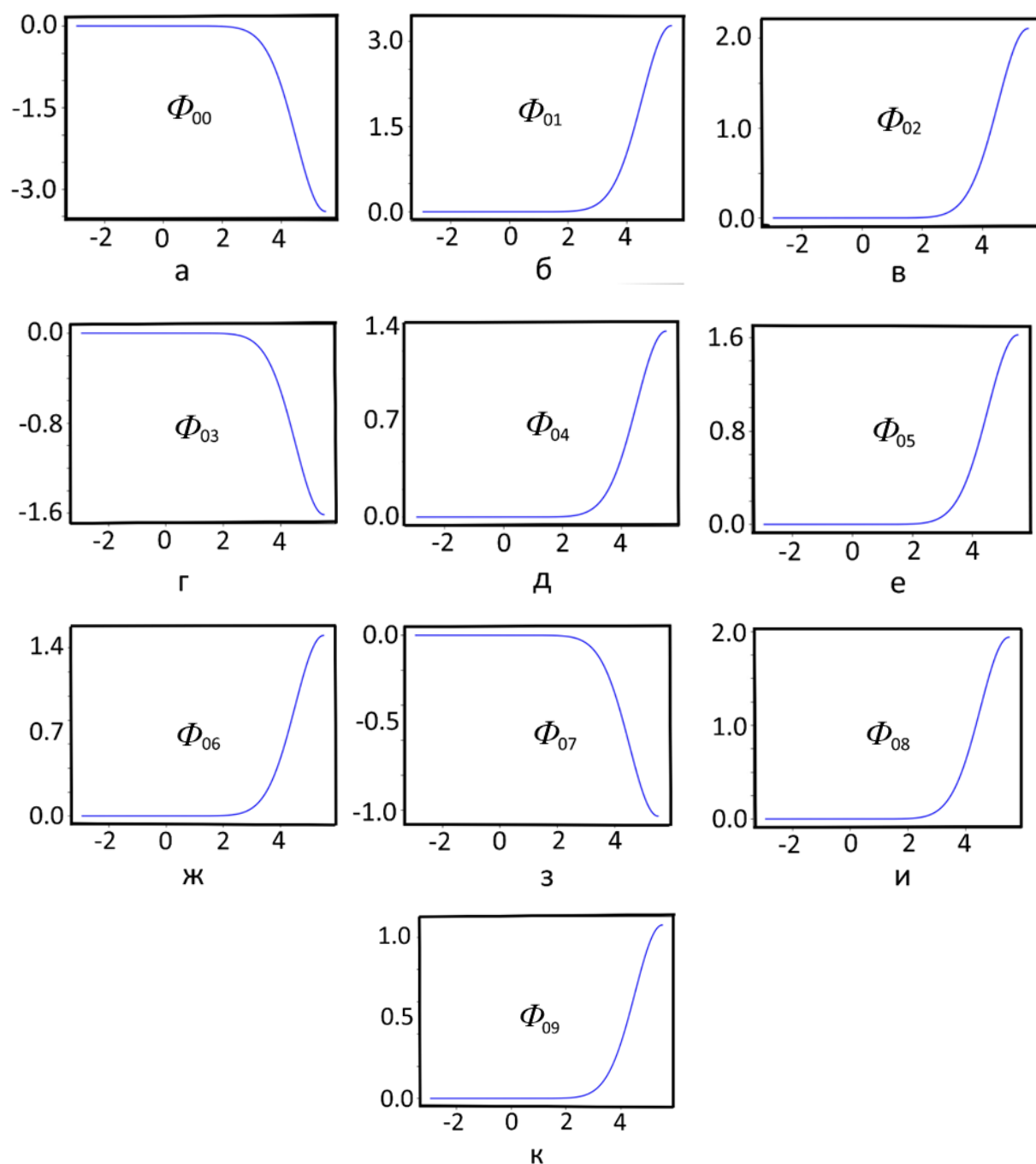


Рисунок 3.10 – Функції активації для вихідних сигналів нейронної мережі ССТ після навчання другого ступеня ансамблевого класифікатора:
а – функція Φ_{00} ; б – функція Φ_{01} ; в – функція Φ_{02} ; г – функція Φ_{03} ;
д – функція Φ_{04} ; е – функція Φ_{05} ; ж – функція Φ_{06} ; з – функція Φ_{07} ;
и – функція Φ_{08} ; к – функція Φ_{09} . [3-25]

Відмінними особливостями запропонованої моделі НМ другого ступеня ансамблевого класифікатора зі стекінгом є такі. Заміна прихованого шару MLP, використаного в роботі [19], на шар навчальних функцій активації дозволила для

того самого набору даних і тих самих НМ на першому ступені ансамблевого класифікатора зменшити кількість помилок класифікації на 0.45% – 1.9%. При використанні шару навчальних функцій активації кількість навчальних вагових коефіцієнтів на другому ступені зменшилася у 1,69 – 3,7 раз. Таке скорочення кількості навчальних параметрів досягнуто завдяки тому, що, на відміну від робіт [40,73,86-94] кожен вихідний елемент першого ступеня оброблявся лише однією навчальною функцією активації. Крім того, НМ другого ступеня містила не тільки навчальні функції активації, а й вихідний шар з класичними нейронами. Необхідно також відзначити, що відмінності запропонованої моделі другого ступеня ансамблевого класифікатора зі стекінгом НМ від інших моделей з навчальними функціями активації, представленими у вище згаданих роботах, дозволили використовувати стандартну бібліотеку Keras [82], а не спеціально розроблену PyKan [40].

Таблиця 3.7 – Порівняння ефективності використання моделі MLP-KAN та моделі MLP на другому ступені ансамблевого класифікатора зі стекінгом НМ

Класифікатори першого ступеня	<i>M</i>	<i>Accuracy</i> , при використанні MLP	<i>Accuracy</i> , при використанні MLP-KAN	Зменшення кількості помилок при використанні MLP-KAN, %
CCT+EATNet+ MLP-Mixer	3	0.8401	0.8431	1.9
CCT+EATNet+ MLP-Mixer+FNet	4	0.8445	0.8452	0.45
CCT+EATNet+ MLP-Mixer+FNet +gMLP	5	0.8457	0.8467	0.65
CCT+EATNet+ MLP-Mixer+FNet +gMLP+SwinTr	6	0.8468	0.8477	0.59

Отриманий у цій роботі позитивний результат є наслідком одночасної наявності в моделі другого ступеня двох різних типів апроксиматорів, які доповнюють один одного:

- шару з навчальними функціями активації;
- вихідного шару з класичними нейронами.

Переваги запропонованої моделі другого ступеня ансамблевого класифікатора зі стекінгом НМ полягають у покращенні якості класифікації при меншій кількості навчальних вагових коефіцієнтів. Додатковою перевагою є можливість використання високорівневої бібліотеки Keras, що спрощує реалізацію.

Результати, наведені в табл. 3.6, показали, що для використаного набору даних CIFAR-10 і дослідження моделі лише другого ступеня класифікатора найкращі результати забезпечує використання шести функцій RBF для апроксимації. Очевидно, що для інших наборів даних і для іншого застосування навчальної НМ необхідно проводити додаткові дослідження, щоб визначити найбільш придатний тип апроксимальних функцій і їх раціональну кількість. Наведені на рис. 3.10 результуючі функції активації підтвердили загальну перевагу використання навчальних функцій активації – вищу інтерпретованість результатів навчання нейронної мережі [40]. На рис. 3.10 чітко видно, що функції активації пропускають до подальшої обробки лише значущі вихідні сигнали, відсікаючи від’ємні та малі позитивні значення. Результати табл. 3.5 та табл. 3.7 показують, що для всіх варіантів ансамблевих класифікаторів, розглянутих у роботі [40], запропонована модель НМ MLP-KAN другого ступеня ансамблевого класифікатора зі стекінгом НМ забезпечує вищу якість класифікації при меншій кількості вагових коефіцієнтів. Необхідно зазначити, що слід очікувати позитивних результатів і в інших моделях нейронних мереж, де можуть бути спільно використані апроксиматори різних типів.

Одним із перспективних напрямів подальших досліджень є створення нових моделей НМ, що містять навчальні функції активації, для першого ступеня ансамблевого класифікатора зі стекінгом. Цей напрям є перспективним для подальшого зменшення обсягу обчислень і підвищення якості класифікації об’єктів на зображеннях.

Висновки до розділу 3

1. Уперше розроблено гібридну модель НМ MLP-KAN, концептуальна відмінність якої полягає в інтеграції шару Колмогорова-Арнольда замість першого прихованого шару MLP. Таке рішення забезпечило синергетичний ефект за рахунок одночасного використання двох типів апроксимації: через навчання функцій активації та навчання вагових коефіцієнтів нейронів. У результаті досягнуто суттєвого скорочення простору параметрів моделі (параметрів, що навчаються) без втрати генералізуючої здатності та точності класифікації на тестових зображеннях.

2. Удосконалено модель другого рівня ансамблевого класифікатора зі стекінгом НМ імплементації запропонованої моделі MLP-KAN як мета-класифікатора замість стандартного MLP. Модифікацію реалізовано з урахуванням специфіки розподілу вихідного вектору ознак першого ступеня ансамблевого класифікатора, що дозволило знизити обчислювальну складність процесу мета-навчання за умов класифікації низькороздільних зображень та дефіциту навчальної вибірки.

3. Проведене дослідження залежності якості класифікації від кількості використаних RBF показало, що оптимальна кількість залежить від параметрів усієї НМ. Середнім найкращим значенням для наборів даних MNIST і CIFAR-10 є 6.

4. Результати експериментальних досліджень показали, що при тій самій якості класифікації, що й у MLP, НМ MLP-KAN (за відповідного вибору параметрів) потребує у 3,63 рази менше обчислень. При цьому MLP-KAN може забезпечити й вищу якість класифікації, ніж MLP, однак виграш у обсязі обчислень при цьому зменшується.

5. Розроблено та досліджено модель НМ для другого ступеня ансамблевого класифікатора. Кожен вихідний сигнал першого ступеня класифікатора обробляється лише однією навчальною функцією активації. Показано, що така архітектура забезпечує вищу якість класифікації при меншій кількості навчальних коефіцієнтів – у 1,69 – 3,7 разів менше, ніж у MLP.

6. Порівняння використання запропонованої архітектури НМ з MLP на другому ступені ансамблевого класифікатора (при однакових класифікаторах на першому ступені) показало зменшення похибки класифікації на 0,45 % – 1,9 %.

4 ДОСЛІДЖЕННЯ ЗАПРОПОНОВАНИХ МОДЕЛЕЙ, МЕТОДУ, ПРОГРАМНИХ ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ В АВТОНОМНИХ РОБОТОТЕХНІЧНИХ КОМПЛЕКСАХ

4.1 Особливості класифікації зображень з використання нейронних мереж в автономних робототехнічних комплексах

Актуальність впровадження автономних робототехнічних комплексів зумовлена інтенсивним розвитком технологій штучного інтелекту, комп'ютерного зору та систем розпізнавання, що забезпечують виконання складних завдань у динамічних середовищах без втручання людини [95]. Такі системи забезпечують підвищення продуктивності та мінімізацію ризиків у небезпечних умовах і є критично важливими для промислової автоматизації, прецизійного землеробства, медичної діагностики, систем безпеки, тощо. [96]. Ключовим аспектом автономності є здатність робота до аналізу сцен та прийняття рішень на основі отриманих сенсорних візуальних даних, серед яких класифікація об'єктів на зображеннях посідає центральне місце [97].

Сучасні методи класифікації зображень базуються на моделях глибокого навчання (англ. Deep Learning), зокрема згорткових нейронних мереж, таких як ResNet, EfficientNet та YOLO, які дозволяють виконувати детекцію та сегментацію зображень в реальному часі [98][99]. Попри високу ефективність цих моделей за наявності великих розмічених наборів даних, їх застосування обмежується в умовах дефіциту обчислювальних ресурсів, обмеженості навчальних вибірок або специфічних характеристик вхідних даних (наприклад, малого розміру об'єктів) [47].

Існує широкий клас прикладних задач, де використання класичних підходів до класифікації або ансамблевих класифікаторів зі стекінгом НМ є доцільнішими ніж використання глибоких моделей, що характеризуються надлишковістю. Характеристиками такого класу задач є:

- вбудовані системи (англ. Edge Computing). Обмеження пам'яті та енергоспоживання у вузлах IoT потребують легких моделей для класифікації [100];
- малі вибірки (англ. Small-Sample Learning). Схильність глибоких моделей до перенавчання у спеціалізованих галузях (медицина, мікроскопія та ін.) потребує методів регуляризації або розробки спеціалізованих моделей [101];
- аналіз об'єктів малого розміру. Класичні CNN втрачають інформативність ознак, коли об'єкт займає мінімальну кількість пікселів [102];
- системи реального часу. Промисловий контроль вимагає детермінованості обробки, що часто суперечить використанню масивних НМ[103].

У другому розділі проведено порівняльний аналіз нейромережових класифікаторів для першого рівня ансамблевого класифікатора зі стекінгом на наборі даних CIFAR-10. Доведено, що модель ССТ забезпечує оптимальний компроміс між точністю та обчислювальною складністю [19].

Одним із перспективних прикладів застосування ССТ є гуманітарне розмінування за допомогою БПЛА. Традиційні методи розмінування характеризуються низькою швидкістю та високою загрозою для життя персоналу, що зумовлює використання дронів-міношукачів [104]. Специфіка цієї задачі полягає у виявленні малорозмірних, замаскованих об'єктів на зображеннях ґрунту при малих висотах зйомки, а також необхідності виконання розрахунків на борту БПЛА при суворих обмеженнях енергоспоживання [105]. Аналіз ознак, таких як порушення текстури ґрунту, наявність сторонніх об'єктів або аномалії рослинності, вимагає адаптивних класифікаторів, здатних працювати в умовах дефіциту навчальних даних [106].

4.2 Опис використовуваного набору даних із зображеннями вибухових пристроїв

Для проведення дослідження був побудований спеціальний набір даних, що містить 1826 зображень для навчання, 186 – для валідації та 274 – для тестування [6]. Усі зображення кольорові та мають розмірність (64×64×3). Усі групи зображень

суворо збалансовані. Половина містить зображення вибухових пристроїв – протитанкових мін на різному тлі. Друга половина містить лише фон. Зображення мін були взяті з набору даних Landmines Detection Dataset (by Northumbria) [107]. У цьому наборі даних зображення зберігаються у форматі YOLOv11 із зазначенням координат центру об'єкта та його розмірів на зображенні, розміром 640×640 пікселів. При формуванні набору даних для дослідження із зображення вирізалася ділянка розміром 128×128 пікселів, що містить зображення міни, зменшувалася до розміру 64×64 пікселів і перетворювалася у формат 64×64×3. Приклади таких зображень наведено на рис. 4.1.



Рисунок 4.1 – Приклади зображень з мінами

Фонові зображення формувалися аналогічно з тих самих зображень, тільки вирізалася ділянка, що не містить зображення міни (рис. 4.2).

Для ускладнення завдання класифікації 120 фонових зображень для навчання, 24 валідаційних фонових зображень і 20 тестових фонових зображень були замінені фоном, що містить каміння, яке можна прийняти за міни [108]. Приклади таких зображень наведено на рис. 4.3. Усі зображення промарковані: 1 – на зображенні є міна, 0 – на зображенні немає міни. Повний сформований набір даних збережено на GitHub [109].

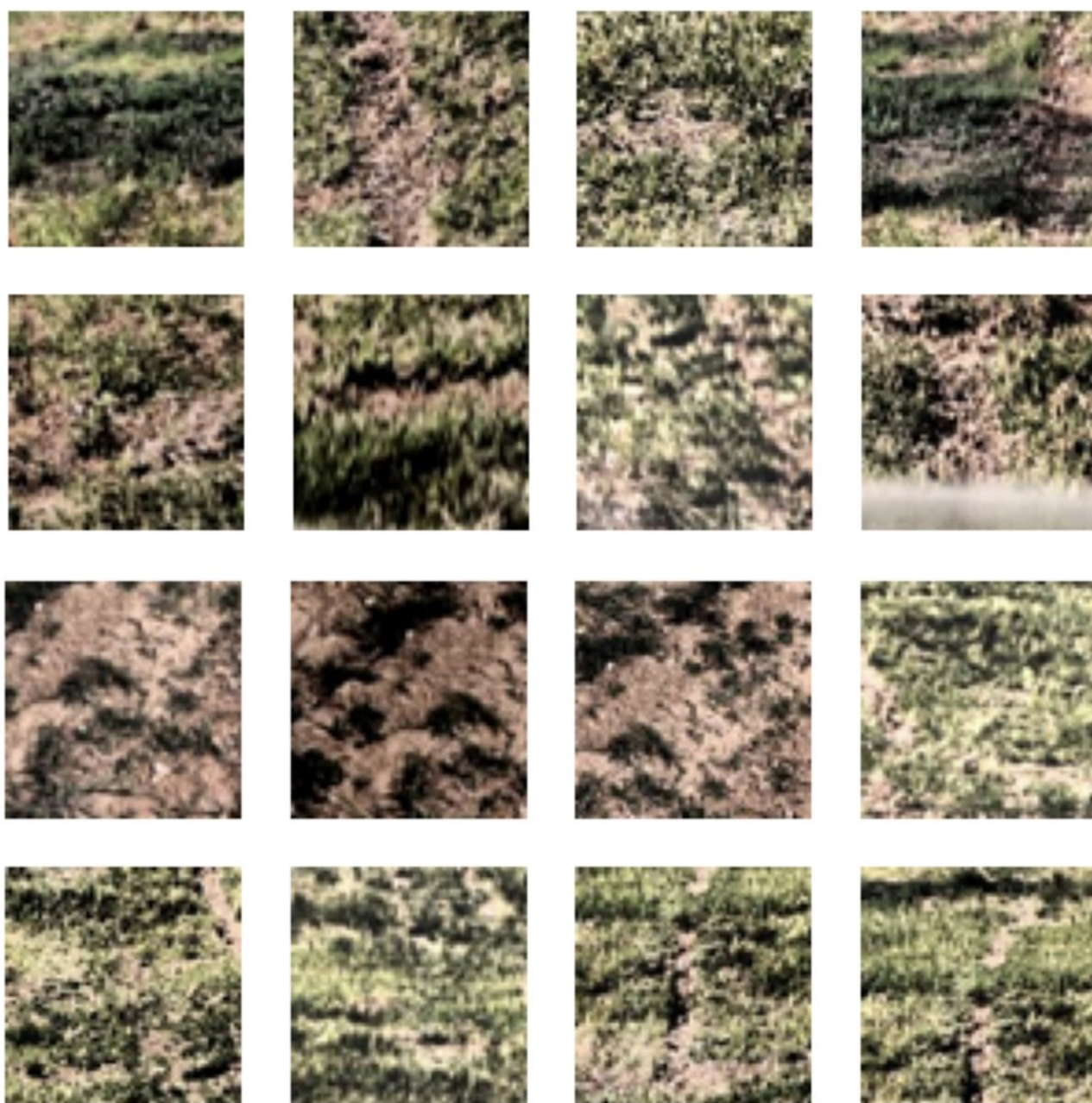


Рисунок 4.2 – Приклади фонових зображень

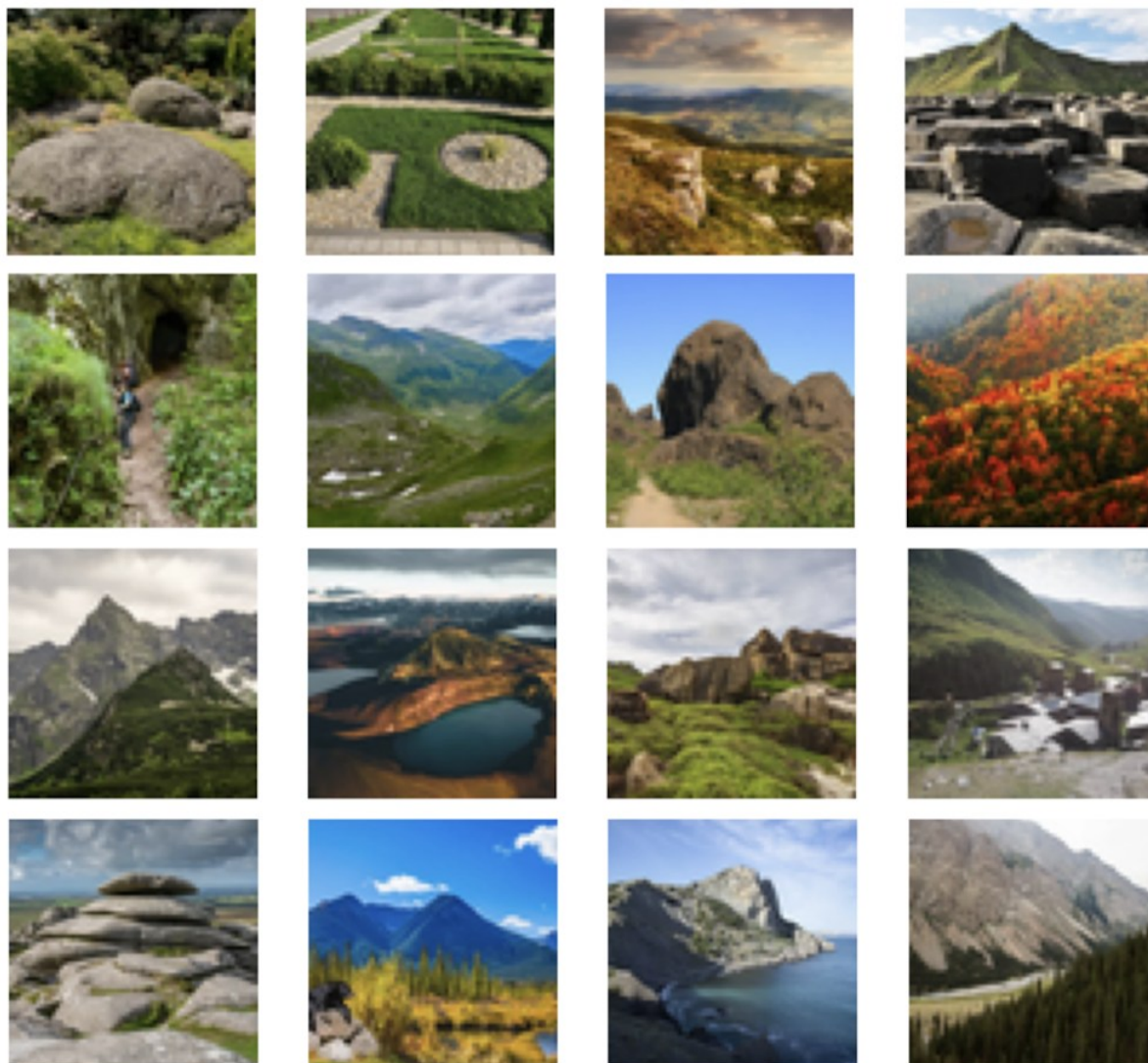


Рисунок 4.3 – Приклади фонових зображень з камінням

Використовувані метрики:

– точність (*Accuracy*), як частка правильно розпізнаних зображень до загальної кількості зображень

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

де TP (True Positive) – істинно позитивні, модель правильно класифікувала наявність міні;

TN (True Negative) – істинно негативні, модель правильно класифікувала відсутність мін;

FP (False Positive) – хибно позитивн, модель класифікувала наявність мін для зображень, де їх реально немає;

FN (False Negative) – хибно негативні, модель класифікувала відсутність мін для зображень, де вони реально є;

– повнота (*Recall*), кількість розпізнаних мін із загальної кількості мін на зображеннях

$$Recall = \frac{TP}{TP + FN}.$$

Тексти всіх програм для НМ, які досліджувалися, наведено в GitHub [109]. Під час навчання запам'ятовувався вектор вагових коефіцієнтів, що забезпечив максимальну *Accuracy* для валідаційних даних. Цей вектор використовувався під час розрахунку результуючих характеристик на тестових даних.

4.3 Класифікація зображень на наявність мін з використанням базових моделей нейронних мереж

На цей час розроблено велику кількість НМ для класифікації зображень. Однак особливістю їх використання для завдань гуманітарного розмінування є відсутність великих наборів даних для навчання та тестування. Тому виникає завдання пошуку НМ, що відрізняються високим співвідношенням точності класифікації та необхідного обсягу обчислень і водночас здатних навчатися на дуже малих обсягах даних. У цьому розділі проведено порівняння згорткових нейронних мереж і мереж, побудованих на базі трансформерів. Навчання мереж проводилося на невеликому наборі даних, що містить зображення протитанкових мін, каміння та різний фон.

Для спеціальних застосувань найчастішим підходом є розроблення нових нейронних мереж. Зазвичай як базис використовується модель ViT. Вона дає змогу змінювати обсяг обчислень від вельми малого до мільярдів операцій із рухомою

комою. Спочатку модель Трансформер була запропонована для обробки текстових даних [30]. На відміну від рекурентних нейронних мереж, Трансформер обробляє все речення/весь контекст одразу. Вбудований у ньому механізм уваги фокусується на релевантних словах. Для кожного слова формується запит (Q – що шукається), ключ (K – що являє собою слово), значення (V – яку інформацію несе слово). Запит цього слова зіставляється з ключами інших слів, чим більший збіг, тим більше уваги (ваги) отримує це слово. Результируюча увага слова визначається як зважена сума значень інших слів. При обробці зображень воно розбивається на патчі, еквіваленти слів у реченні. Потім патчі перетворюються на плоскі вектори (англ. Linear Projection), які далі використовуються в обчисленнях. Якщо для слів кодується їхнє місце в реченні/контексті, то для патчів кодуються їхні координати на зображенні. Особливістю Трансформерів є те, що вони дуже добре масштабуються і потребують максимально можливого обсягу даних для навчання. Тому ефективність відео-трансформерів для обробки невеликих зображень і малої вибірки даних для навчання потребує додаткового дослідження для кожного конкретного набору даних.

Для забезпечення високого співвідношення якості класифікації та обсягу обчислень розроблено низку спеціальних модифікацій моделі Трансформер. У роботі [110] запропоновано модифікацію класичного трансформера, в якій ділянки, на які розбивається вихідне зображення (патчі), зроблені такими, що перекриваються. Це дає змогу враховувати локальні структури зображень, розташовані на межах патчів. У [111] покращено гнучкість до розташування об'єктів на зображеннях за рахунок введення в архітектуру трансформера додаткової позиційної самоуваги. У [112] запропоновано багаторівневий механізм перехресної уваги, що витягує ознаки нижчого рівня. У [113] лінійні проекції на етапі токенизації замінені на згорткові проекції. У [25] механізм внутрішньої самоуваги, що має квадратичну складність, замінено на механізм зовнішньої уваги з лінійною складністю. Таким чином, модель трансформера має велику гнучкість і допускає численні модифікації. Однак ефективність використання тих чи інших модифікацій потребує дослідження та адаптації для конкретних цільових наборів

даних. У [19] було проведено таке дослідження для набору даних CIFAR-10 [39]. Були розглянуті такі модифікації трансформерів (див. п. 1.3):

- CCT [24] – у трансформер додано згорткову нейронну мережу на вході для зменшення розмірів патчів і підвищення ефективності навчання на невеликих наборах даних;
- EATNet [25] – використано додаткові блоки багатошарового перцептрона (MLP) для реалізації зовнішньої уваги;
- FNet [26] – внутрішній блок самоуваги замінено на простіше перемішування токенів за допомогою швидкого перетворення Фур'є;
- SwinTr [31] – забезпечено гнучкість при обробці об'єктів із різними масштабами за рахунок обчислення самоуваги не по всьому зображенню, а тільки всередині зсунутих вікон.

У [19] показано, що на наборі даних CIFAR-10 найкраще співвідношення якості класифікації та обсягу обчислень має CCT. Однак для набору зображень із мінами ефективність використання CCT потребує окремого дослідження.

На відміну від архітектур на базі трансформера, які є чутливими до величини навчального набору даних, простіші згорткові мережі [85] мають меншу чутливість до обсягу даних для навчання.

З проведеного аналізу літератури випливає, що для виявлення та класифікації мін з урахуванням реалізації на обчислювачі, встановленому на БПЛА, доцільно порівняти ефективність використання згорткових мереж типу CNN і ResNet-18 [85] із мережами на базі трансформерів – CCT та варіантами ViT із різною розмірністю.

4.3.1 Класифікатор на основі згорткової нейронної мережі CNN

Мережа містить два згорткові шари з функцією активації Relu. Після кожного з них застосовано шар MaxPool для зменшення розмірності. Після цього – повнозв'язний шар Dense з функцією активації Relu, розмірністю 64, і вихідний шар Dense, розмірністю 1 і функцією активації sigmoid (див. Лістинг 4.1).

Лістинг 4.1 – Фрагмент коду налаштування параметрів моделі CNN

```
# Блок 1
x = layers.Conv2D(32, (3, 3), activation="relu")(x)
x = layers.MaxPool2D((2, 2))(x)

# Блок 2
x = layers.Conv2D(64, (3, 3), activation="relu")(x)
x = layers.MaxPool2D((2, 2))(x)

# Повнозв'язкові шари
x = layers.Flatten()(x)
x = layers.Dense(64, activation="relu")(x)
x = layers.Dropout(0.5)(x)
outputs = layers.Dense(1, activation="sigmoid")(x)
```

Кількість вагових коефіцієнтів, що навчаються, у цій НМ 822337. Криві навчання для точності, помилки та ROC-криві для даних для навчання та валідаційних даних наведено відповідно на рис. 4.4, 4.5 і 4.6. Аналіз кривих навчання показує, що мережа навчилася приблизно за 40 епох. Площа під ROC-кривою AUC=0.99, що свідчить про те, що нейронна мережа добре розрізняє класи (є міна на зображенні / немає міни на зображенні).

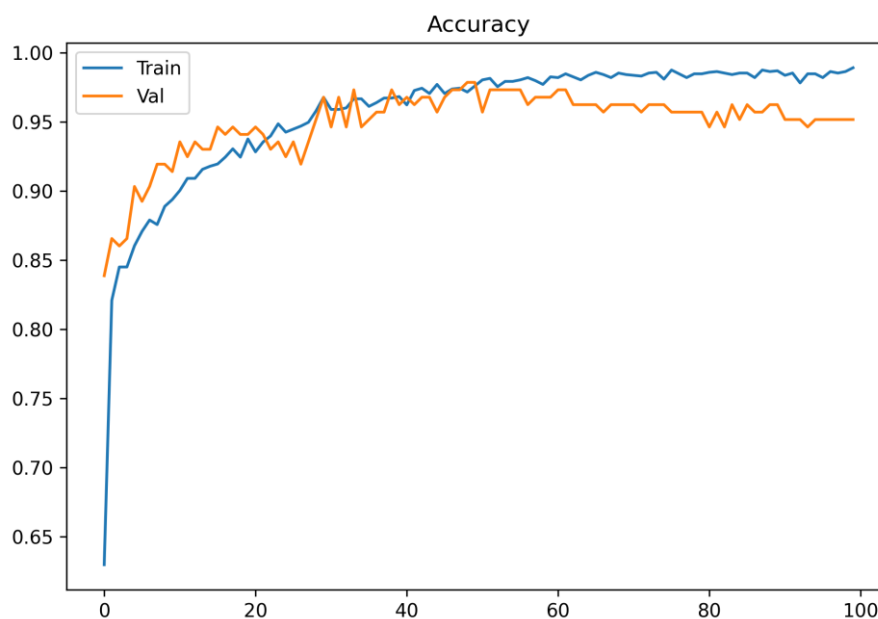


Рисунок 4.4 – Криві навчання точності класифікатора CNN

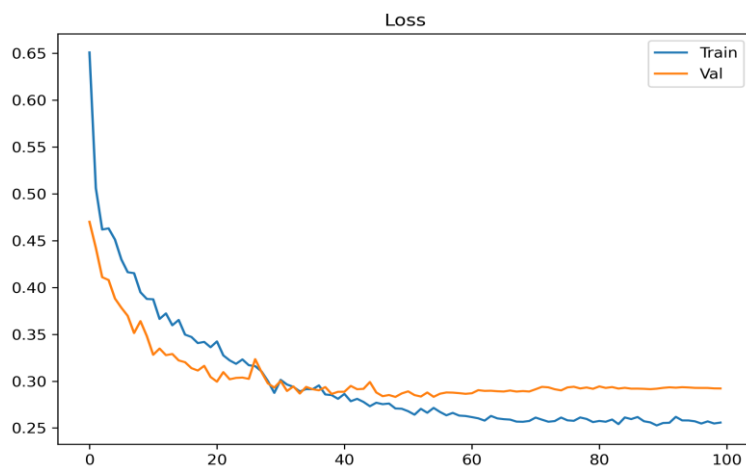


Рисунок 4.5 – Криві навчання помилки класифікатора CNN

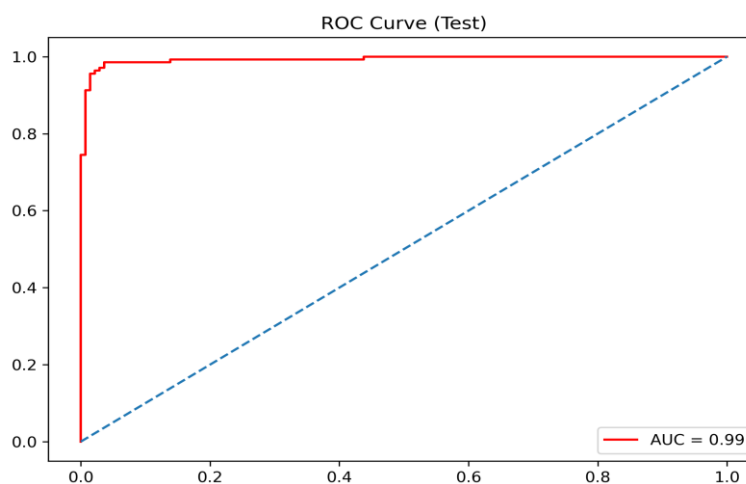


Рисунок 4.6 – ROC крива класифікатора CNN

Матриця помилок для тестових даних наведена на рис. 4.7.

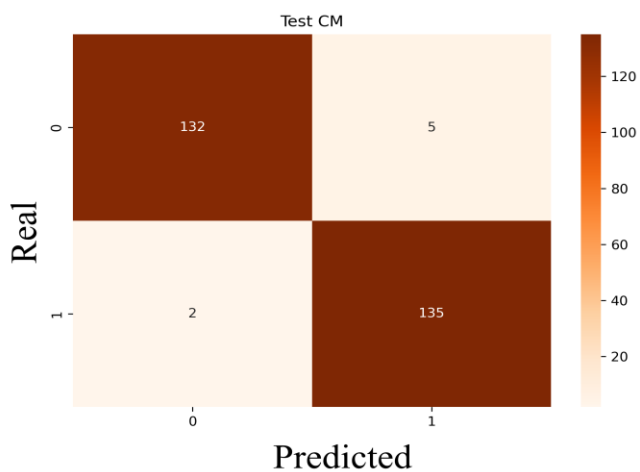


Рисунок 4.7 – Матриця помилок на тестових даних класифікатора CNN

Матриця помилок свідчить про те, що із загальної кількості зображень мін 137 два були неправильно класифіковані, тобто дві міни були пропущені. Хибних вказівок було 5.

Розподіл упевненості моделі наведено на рис. 4.8.

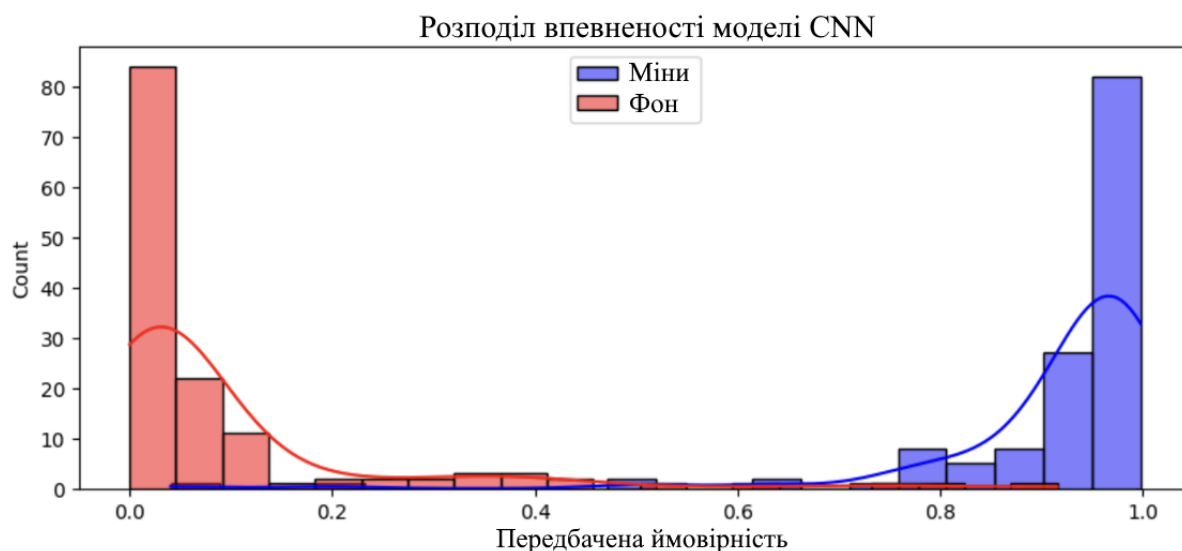


Рисунок 4.8 – Розподіл впевненості для класифікатора CNN

З розподілу впевненості видно, що класи досить добре розділяються, вгадування практично відсутні, однак невелика їх кількість має місце. Необхідно також відзначити, що краї розподілу трохи розмиті, що свідчить про схильність до помилок при погіршенні якості вхідних зображень.

Зображення, на яких були зроблені помилки, наведено на рис. 4.9.

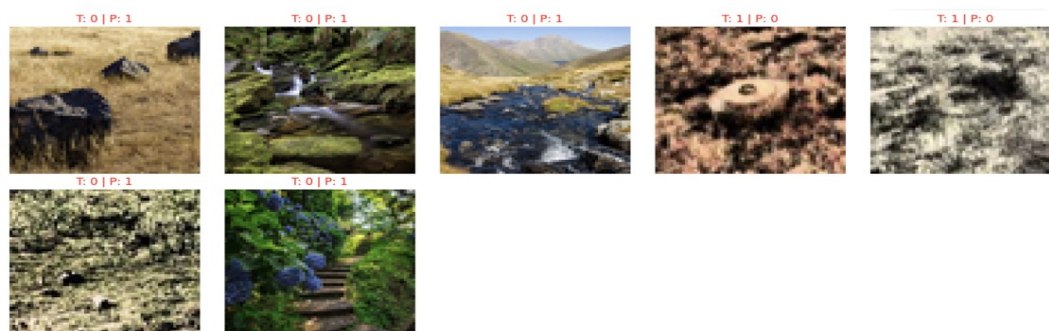


Рисунок 4.9 – Приклади зображень, на яких класифікатор CNN зробив помилки (T – True, P – Predicted)

З рис. 4.9 видно, що один пропуск міні був зроблений на зображенні з рідкісним розподілом освітленості – від міні були відблиски, другий пропуск – міна сильно зливалася з фоном.

4.3.2 Класифікатор на основі згорткової нейронної мережі ResNet-18

На вході згорткової мережі ResNet-18 стоїть вхідний згортковий шар, на який надходить вхідне зображення. Далі стоять 8 блоків, кожен із яких містить 2 згорткові шари. Функція активації – Relu. Вихідний сигнал другої згорткової мережі підсумовується з вхідним сигналом першої згорткової мережі (skip-connections) (див. лістинг 4.2).

Лістинг 4.2 – Код налаштування параметрів моделі ResNet-18

```
def residual_block(x, filters, kernel_size=3, stride=1,
downsample=False):
    shortcut = x
    if downsample:
        shortcut = layers.Conv2D(filters, 1, strides=stride,
padding="same")(shortcut)
        shortcut = layers.BatchNormalization()(shortcut)

    x = layers.Conv2D(filters, kernel_size, strides=stride,
padding="same", use_bias=False)(x)
    x = layers.BatchNormalization()(x)
    x = layers.Activation("relu")(x)

    x = layers.Conv2D(filters, kernel_size, strides=1, padding="same",
use_bias=False)(x)
    x = layers.BatchNormalization()(x)

    x = layers.Add()([x, shortcut])
    x = layers.Activation("relu")(x)
    return x
```

На виході стоїть повнозв'язний шар Dense з функцією активації sigmoid. У сумі НМ ResNet-18 містить 18 шарів із ваговими коефіцієнтами, що навчаються. Кількість вагових коефіцієнтів, що навчаються, у цій НМ 11177921. Криві навчання для точності, помилки та ROC-криві для даних для навчання та валідаційних даних

наведено відповідно на рис. 4.10, 4.11 і 4.12. Аналіз кривих навчання показує, що мережа навчилася приблизно за 100 епох. Площа під ROC-кривою $AUC=0.99$, що свідчить про те, що НМ добре розрізняє класи (є міна на зображенні / немає міни на зображенні).

Матриці помилок для тестових даних наведено на рис. 4.13.

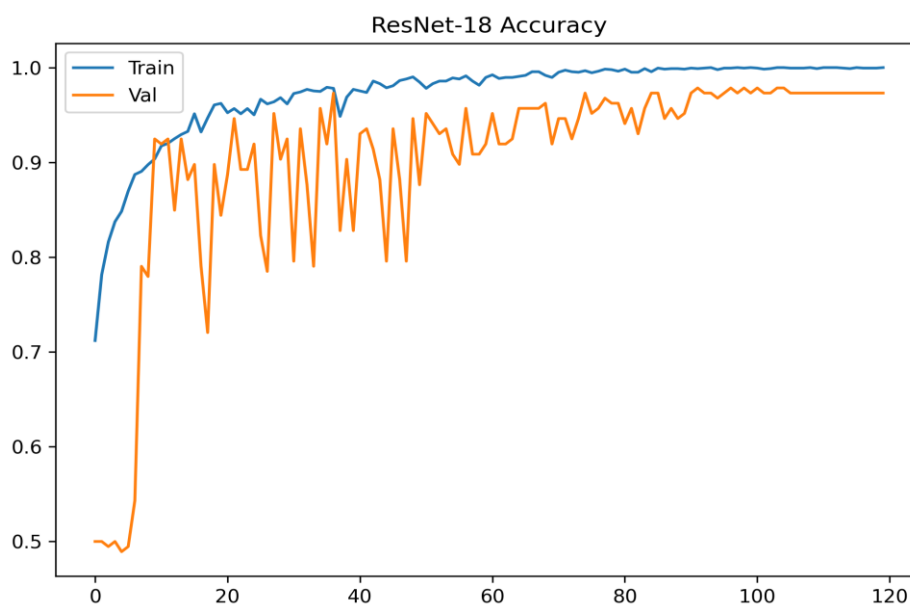


Рисунок 4.10 – Криві навчання точності класифікатора ResNet-18

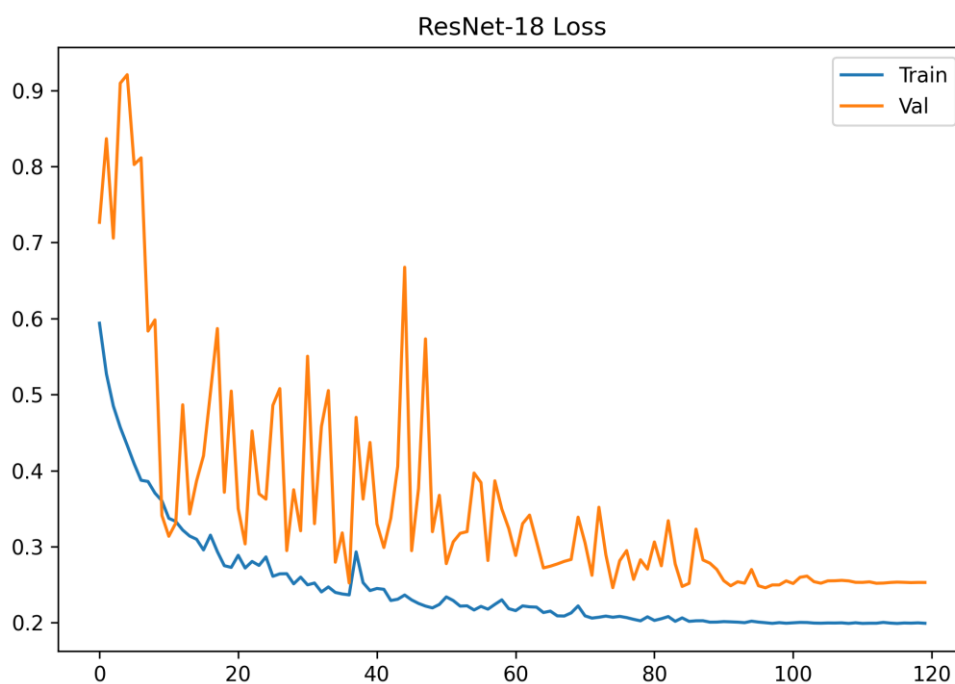


Рисунок 4.11 – Криві навчання помилки класифікатора ResNet-18

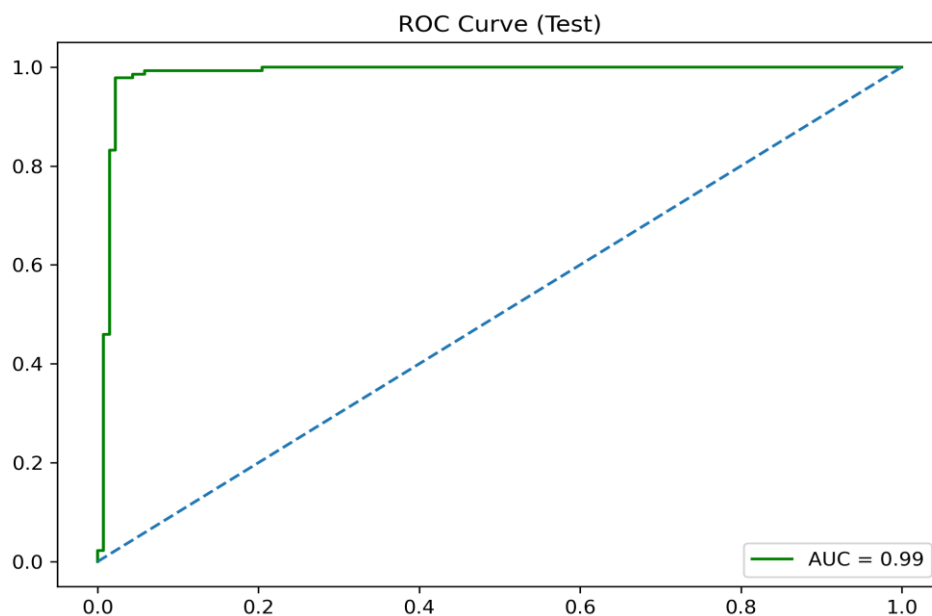


Рисунок 4.12 – ROC крива класифікатора ResNet-18

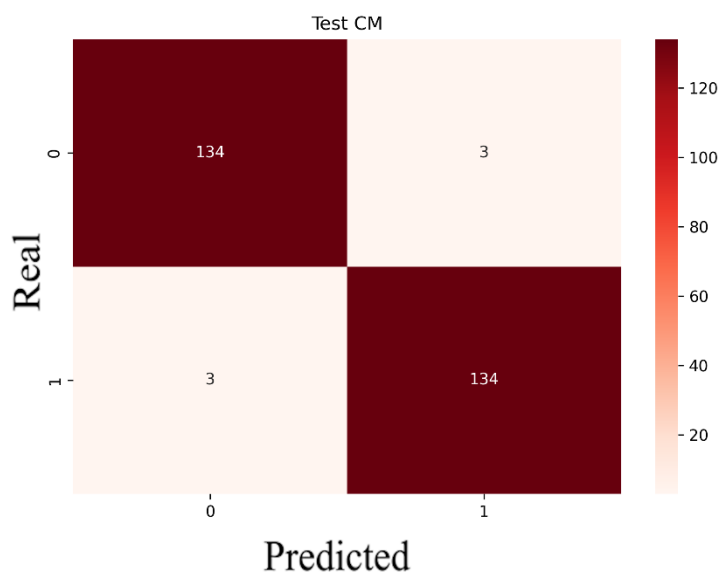


Рисунок 4.13 – Матриця помилок на тестових даних класифікатора ResNet-18

Матриця помилок свідчить про те, що із загальної кількості зображень мін 137 три були неправильно класифіковані. Хибних вказівок було 3. Незважаючи на те, що ResNet-18 має більш ніж на порядок більше вагових коефіцієнтів і точність Accuracy=97,81% (зроблено 6 помилок), ніж проста CNN, пропущено 3 міни. У CNN пропущено тільки 2 міни.

Розподіл упевненості моделі наведено на рис.4.14.

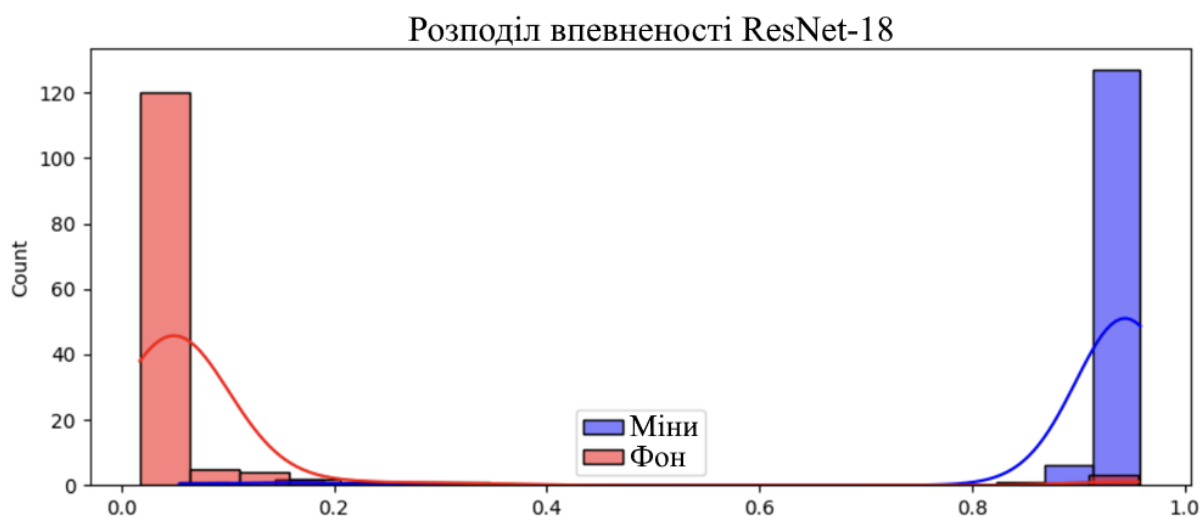


Рисунок 4.14 – Розподіл впевненості для класифікатора ResNet-18

З розподілу впевненості видно, що класи досить добре розділяються, вгадування відсутні. Необхідно також відзначити, що краї розподілу суттєво менше розмиті, ніж у CNN.

Зображення, на яких були зроблені помилки, наведено на рис. 4.15.



T – True, P – Predicted

Рисунок 4.15 – Зображення, на яких класифікатор ResNet-18 зробив помилки

З рис. 4.15 видно, що ResNet-18 пропустила міни на тих самих двох зображеннях, що й CNN. Крім цього, ResNet-18 пропустила міну на зображенні, де міна була розташована за огорожею. Дві з трьох хибних вказівок на міни зроблені на тих самих зображеннях, що й у CNN.

4.3.3 Класифікатор на основі класичного трансформеру ViT

У зв'язку з тим, що класичний ViT потребує великого обсягу даних для навчання, а дане дослідження проводиться на дуже малому обсязі даних, були досліджені кілька варіантів ViT із невеликими обсягами обчислень. З огляду на невеликий розмір зображень 64×64 пікселів, були використані невеликі розміри патчів 8×8 і 4×4 . Результати наведено в табл. 4.1.

З табл. 4.1 видно, що найкраще співвідношення точності до обсягу обчислень має варіант ViT-Tiny, однак на тестових даних він пропустив 6 мін. Другим за співвідношенням точності до обсягу обчислень є ViT-Lite. Кількість пропущених мін – 4. Решта два варіанти ViT-Robust і ViT-Deer забезпечують вищу точність, меншу кількість пропущених мін, але й мають суттєво більший обсяг обчислень. Тому, за сукупністю параметрів, найкращим результатом, який можна очікувати від класичного ViT для досліджуваного набору даних, є ViT-Lite.

Загальна закономірність між параметрами і характеристиками видна з табл. 4.1. Збільшення точності ViT досягається:

- збільшенням розмірності проекції (довжиною вектора, на який перетворюється патч);
- збільшенням кількості голів;
- збільшенням кількості шарів Трансформера;
- зменшенням розмірності патчів.

Однак необхідно зазначити, що параметри Трансформерів не змінюються довільно, а з певними кроками. Наприклад, розмірності векторів, на які перетворюються патчі, повинні націло ділитися на кількість використовуваних голів.

Криві навчання ViT-Lite для точності, помилки та ROC-криві для даних для навчання та валідаційних даних наведено відповідно на рис. 4.16, 4.17 і 4.18. Аналіз кривих навчання показує, що мережа навчилася приблизно за 110 епох. Площа під

ROC-кривою AUC=0.98, що свідчить про те, що НМ добре розрізняє класи (є міна на зображенні / немає міни на зображенні).

Таблиця 4.1 – Параметри модифікацій моделі ViT

Параметр	ViT-Tiny (базовий)	ViT-Lite	ViT-Robust (посилений)	ViT-Deep (глибокий)
Projection Dim	64	128	192	128
Heads	4	4	6	8
Transformer Layers	2	4	6	12
Patch Size	8×8	8×8	8×8	4×4
Label Smoothing	0.05	0.1	0.15	0.1
Перевага	Дуже швидкий, не перенавчається на малих даних	Збалансований за швидкістю та точністю	Найкраща стійкість до шуму та туману	Максимальне вилучення дрібних ознак
Ризик	Може не вловити складні текстури міни	Стандартний варіант	Потребує більше пам'яті та часу навчання	Високий ризик перенавчання (overfitting)
Trainable Parameters	328705	563201	1856705	3667521
Accuracy	0.9635	0.9416	0.9526	0.9708
Recall	0.9562	0.9708	0.9854	0.9854
FN	6	4	2	2
Accuracy / Trainable Parameters	2.93×10^{-6}	1.67×10^{-6}	0.51×10^{-6}	0.26×10^{-6}

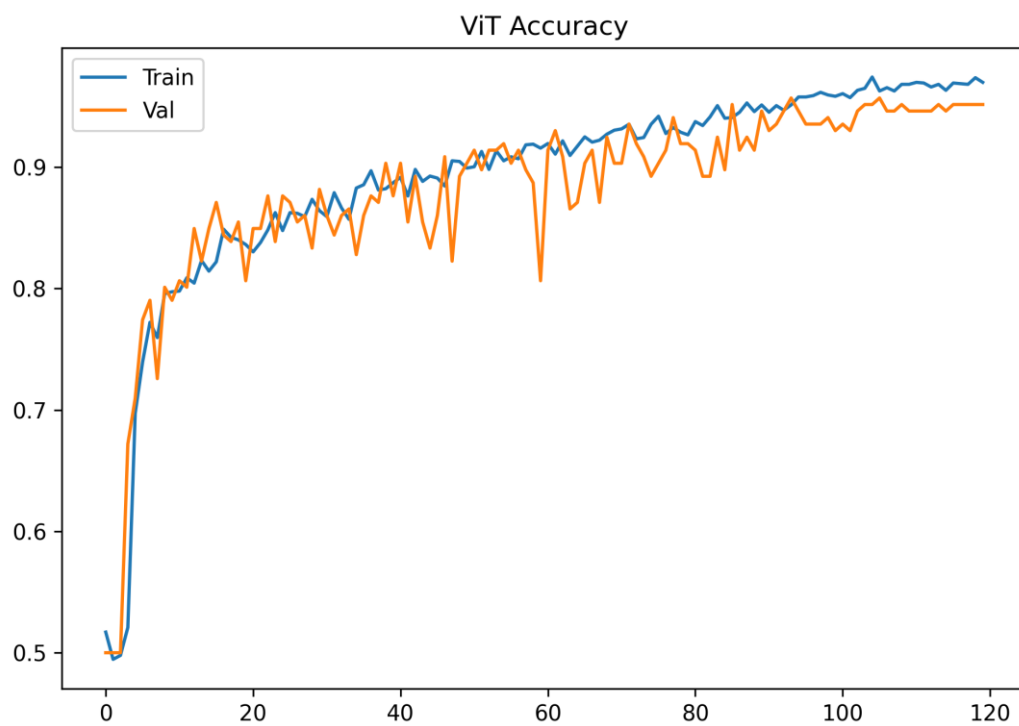


Рисунок 4.16 – Криві навчання точності класифікатора ViT-Lite

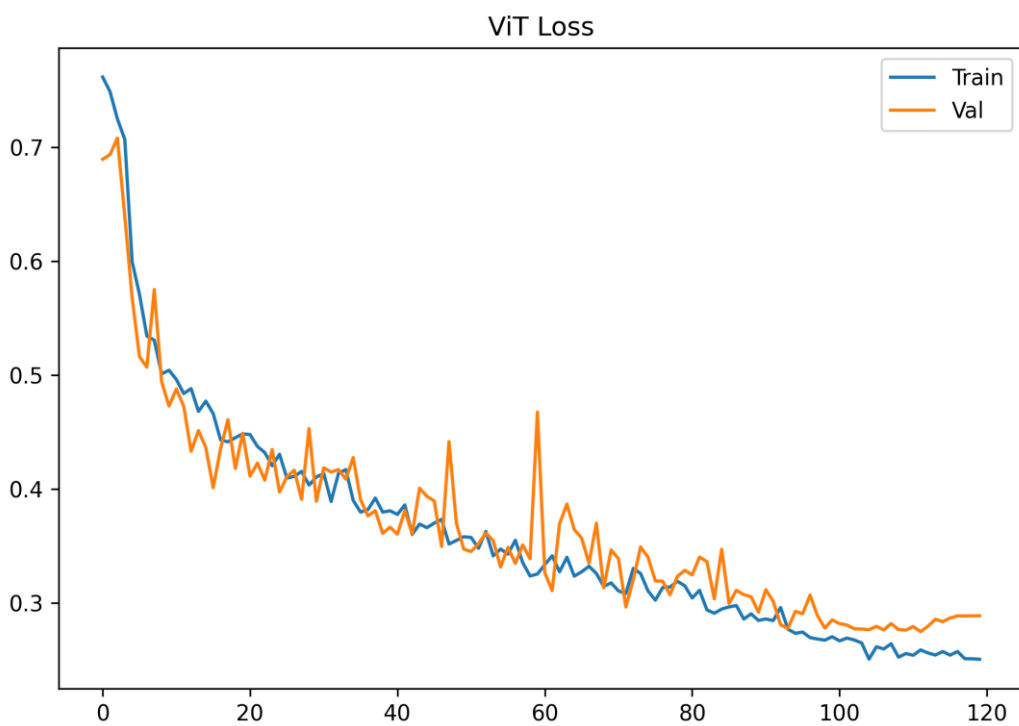


Рисунок 4.17 – Криві навчання помилки класифікатора ViT-Lite

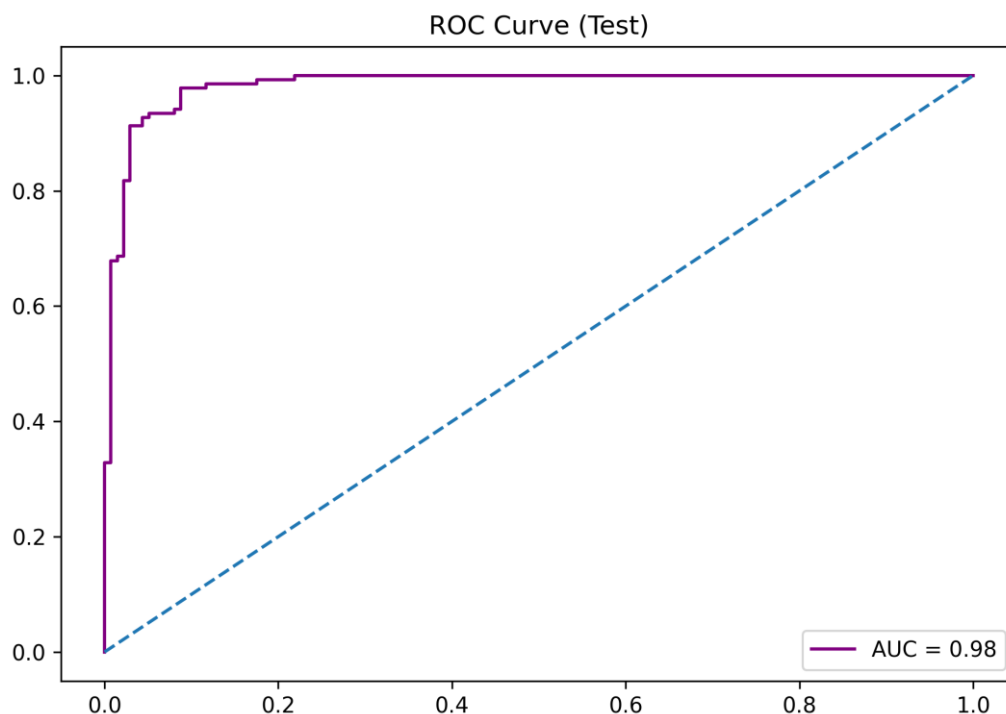


Рисунок 4.18 – ROC крива класифікатора ViT-Lite

Матриця помилок для тестових даних наведена на рис. 4.19.



Рисунок 4.19 – Матриця помилок на тестових даних класифікатора ViT-Lite

Матриця помилок свідчить про те, що із загальної кількості зображень мін 137 чотири були неправильно класифіковані. Хибних вказівок було 12. Розподіл впевненості моделі наведено на рис. 4.20.

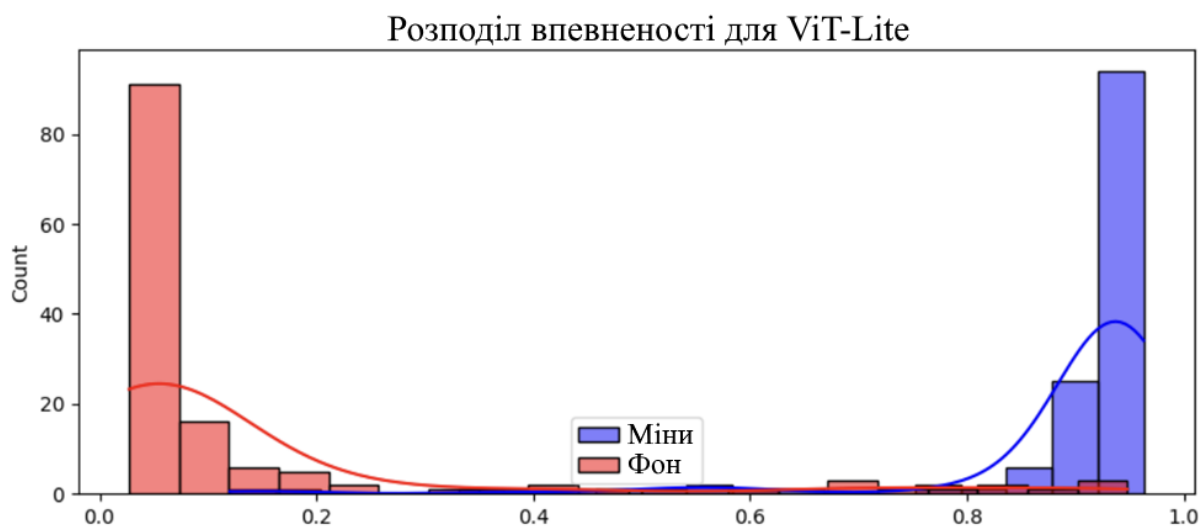


Рисунок 4.20 – Розподіл впевненості для класифікатора ViT-Lite

З розподілу впевненості видно, що класи досить добре розділяються, однак невелика кількість вгадувань має місце. Необхідно також відзначити, що межі розподілу трохи розмиті, що свідчить про схильність до помилок при погіршенні якості вхідних зображень.

15 зображень, на яких були зроблені помилки, наведено на рис. 4.21.



Рисунок 4.21 – Приклади зображень, на яких ViT-Lite припустився помилки
(T – True, P – Predicted)

4.3.4 Класифікатор на основі ССТ

Для забезпечення вищої точності класифікації при одночасному зменшенні обсягу обчислень в модель ССТ внесено такі зміни порівняно з класичним ViT:

- розбиття зображення на патчі, що не перекриваються (Patch-ембеддінг), замінено на згортковий токенизатор. Токенизатор використовує два згорткові шари з фільтрами, що перекриваються, функцією активації GELU і нормалізацією BatchNormalization. Це дає змогу ефективніше вилучати низькорівневі ознаки та враховувати локальні зв'язки між пікселями;

- механізм Global Average Pooling замінено на Sequence Pooling. У ViT усі токени мають однакову вагу. ССТ, використовуючи увагу (Attention), сам навчається визначати, які саме частини зображення (токени) найважливіші для класифікації (наприклад, де саме знаходиться міна), і надає їм більшої ваги;

- збільшено коефіцієнт розширення MLP-блоку з 2 до 3. Це збільшує кількість вагових коефіцієнтів у кожному шарі, що обробляє ознаки, вилучені токенизатором;

- введено плавний розігрів (Warmup) – плавне збільшення швидкості навчання і подальше її зменшення за косинусом (Cosine Decay). Це дає змогу моделі краще навчатися на малих наборах даних.

Криві навчання для точності, помилки та ROC-криві для даних для навчання та валідаційних даних наведено відповідно на рис. 4.22, 4.23 і 4.24. Аналіз кривих навчання показує, що мережа навчилася приблизно за 90 епох. Площа під ROC-кривою $AUC=0.98$, що свідчить про те, що нейронна мережа добре розрізняє класи (є міна на зображенні / немає міни на зображенні). Матриця помилок для тестових даних наведена на рис. 4.25.

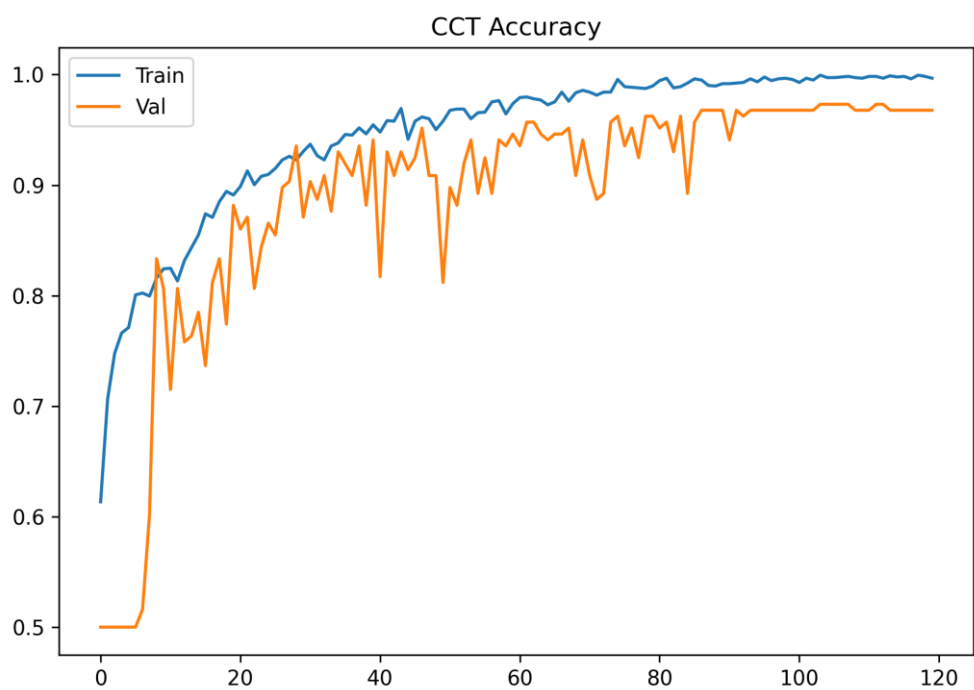


Рисунок 4.22 – Криві навчання точності класифікатора CCT

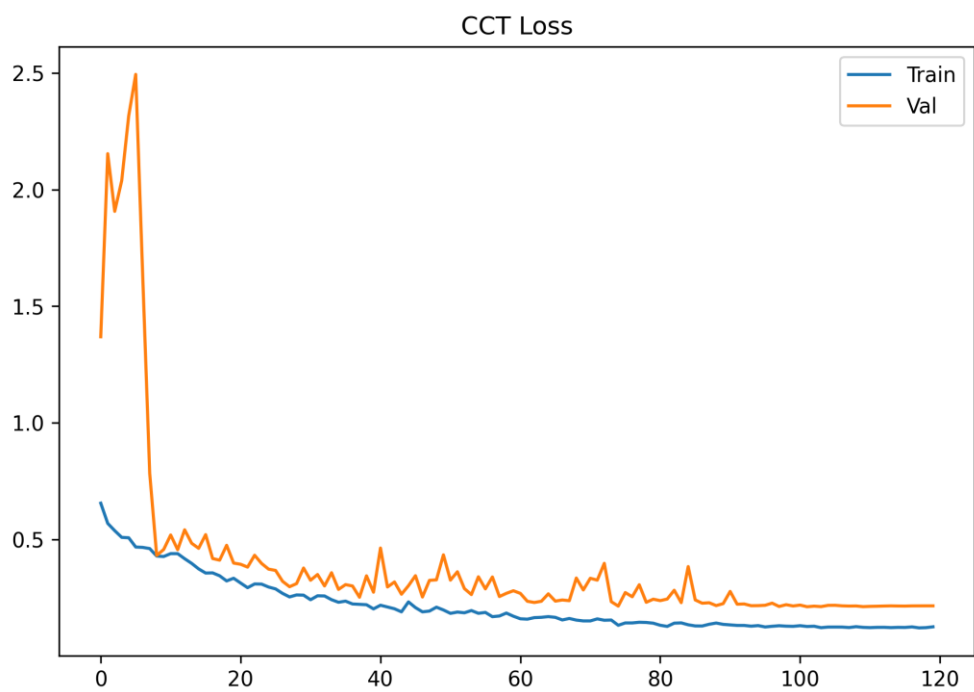


Рисунок 4.23 – Криві навчання помилок класифікатора CCT

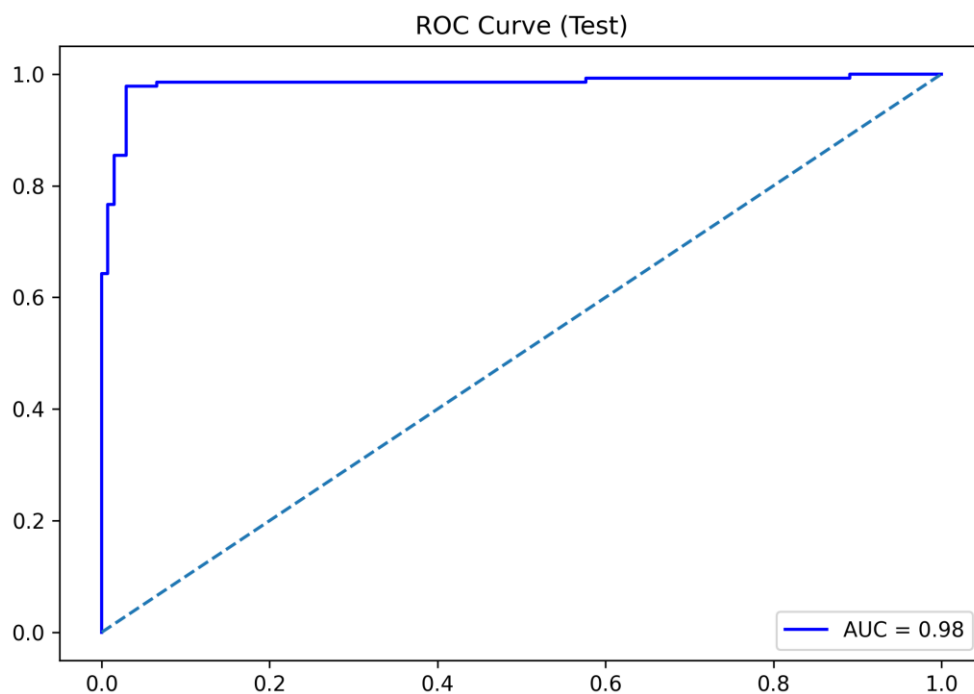


Рисунок 4.24 – ROC крива класифікатора CCT

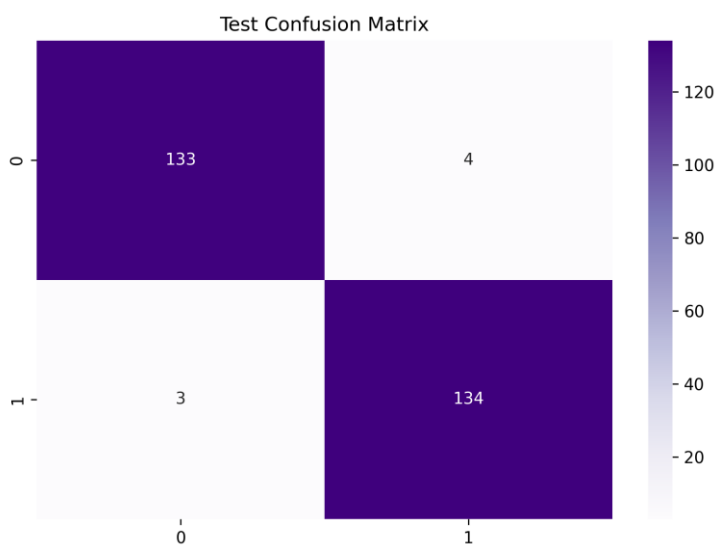


Рисунок 4.25 – Матриця помилок на тестових даних класифікатора CCT

Матриця помилок свідчить про те, що із загальної кількості зображень мін 137 три були неправильно класифіковані. Хибних вказівок було 4.

Розподіл упевненості моделі CCT наведено на рис. 4.26.

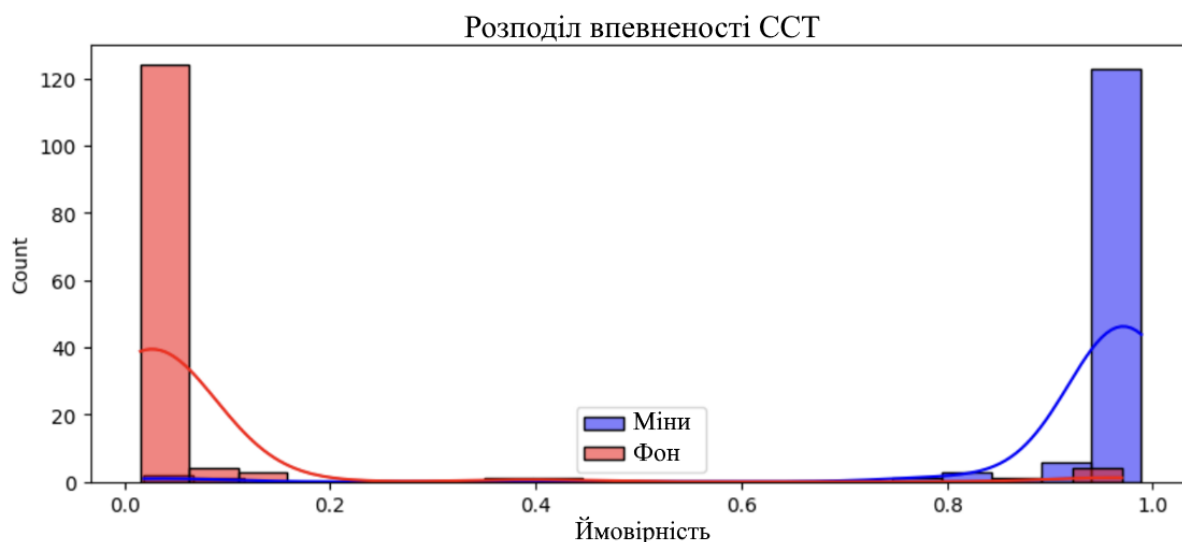


Рисунок 4.26 – Розподіл впевненості для класифікатора SST

З розподілу впевненості видно, що класи досить добре розділяються, однак невелика кількість вгадувань має місце. Необхідно також відзначити, що краї розподілу для SST суттєво менш розмиті, ніж у ViT-Lite.

7 зображень, на яких були зроблені помилки, наведено на рис. 4.27.



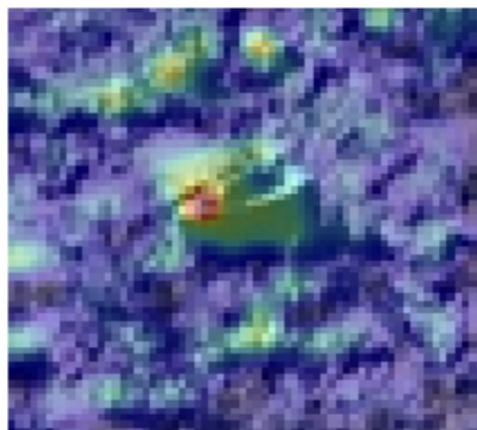
Рисунок 4.27 – Зображення, на яких класифікатор SST припустився помилки
(T – True; P – Predicted)

Дві з трьох помилок щодо пропуску мін зроблені на тих самих зображеннях, на яких робили помилки й інші НМ.

Для ілюстрації роботи класифікатора ССТ на рис. 4.28 наведено приклад вхідного зображення з міною та карти уваги для цього зображення в першому шарі ССТ. На рис. 4.29 наведено те саме, тільки на вхідному зображенні замість мін присутнє каміння.



а



б

Рисунок 4.28 – Приклади зображень: а – вхідне зображення з міною; б – карта уваги в першому шарі класифікатора ССТ



а



б

Рисунок 4.29 – Приклади зображень: а – вхідне зображення з камінням; б – карта уваги в першому шарі класифікатора ССТ

Світлим кольором на карті уваги вказані місця на зображенні, де найбільш імовірно розташування міни. Чим світліший колір, тим більша ймовірність. Відповідно, ці місця обробляються з більшими ваговими коефіцієнтами, ніж інші.

4.3.5 Аналіз стійкості нейронних мереж до погіршення якості вхідних зображень

Для перевірки стійкості НМ до погіршення якості вхідних зображень для тестових зображень були змодельовані:

- шум із рівнем 15 %;
- похмурий день (зменшення яскравості на 40 %);
- туман (легкий серпанок, частка білого кольору зменшена до 20 %).

Приклад із вихідним зображенням і погіршеними наведено на рис. 4.30, значення точності на тестових даних наведено в табл. 4.2.

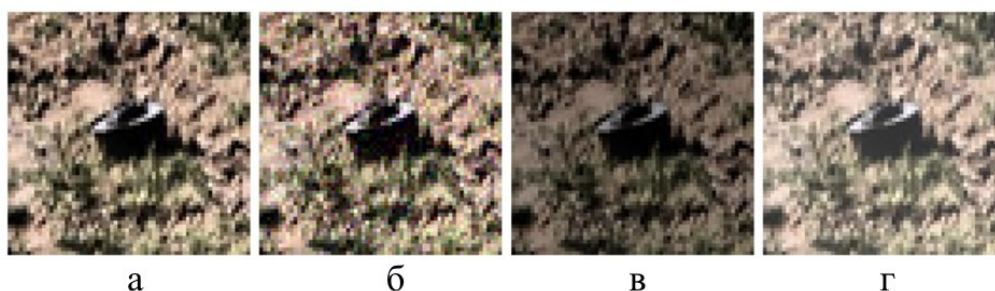


Рисунок 4.30 – Приклади зображень: а – початкове зображення; б – зображення з шумом; в – зображення з шумом «похмурий день»; г – зображення з шумом «туман»

Таблиця 4.2 – Результат оцінювання точності класифікації мін на зображеннях для моделей базових нейромережових класифікаторів

Умови експерименту на зображеннях	Точність класифікації моделі (<i>Accuracy</i>)			
	CNN	ResNet-18	ViT-Lite	CCT
без погіршення	0,9745	0,978	0,9416	0,9745
з додаванням шуму	0,960	0,971	0,920	0,909
з додаванням шуму «похмурий день»	0,832	0,898	0,803	0,832
з додаванням шуму «туман»	0,916	0,942	0,843	0,836

З даних таблиці 4.2 видно, що згорткові мережі мають вищу стійкість, порівняно з мережами на базі Трансформера. Це можна пояснити великою кількістю вагових коефіцієнтів, що навчаються. Однак необхідно зауважити, що робота в погіршених умовах не є основним режимом, оскільки ймовірності пропуску мін суттєво збільшуються.

4.3.6 Результуюче порівняння нейронних мереж

Результуючі параметри порівнюваних НМ на тестовому наборі даних наведено в табл. 4.3.

Таблиця 4.3 – Результати оцінювання параметрів базових НМ при класифікації мін на зображеннях

Параметри НМ	CNN	ResNet-18	ViT-Lite	CCT
Точність (Accuracy)	0,9745	0,978	0,9416	0,9745
Кількість вагових коефіцієнтів, що навчаються	822337	11177921	563201	737858
Відношення точності та кількості навчальних ваг	$1,18 \times 10^{-6}$	$0,087 \times 10^{-6}$	$1,67 \times 10^{-6}$	$1,32 \times 10^{-6}$
Кількість пропущених мін (FN)	2	3	4	3
Впевненість у прийнятих рішеннях	знижена, можливі вгадування	підвищена	знижена	підвищена
Стійкість до погіршення якості зображення	підвищена	підвищена	знижена	знижена

З даних табл. 4.3 видно, що вище співвідношення точності до обсягу обчислень мають НМ на базі трансформерів. Виходячи з того, що ViT-Lite має знижену точність і більшу кількість пропущених мін, найкращим варіантом є використання CCT. Порівняння показників CNN і ResNet-18 свідчить про те, що

потенціал згорткових мереж у цьому застосуванні обмежений, оскільки радикальне збільшення кількості вагових коефіцієнтів, що навчаються, не призвело до суттєвого підвищення точності, а тільки до підвищення впевненості в прийнятих рішеннях.

4.4 Обговорення результатів дослідження ефективності використання різних нейронних мереж для виявлення мін дроном

Проведене дослідження показало, що потенційні можливості простих згорткових мереж обмежені. Суттєве збільшення використовуваних шарів не призводить до великого підвищення точності. Це пояснюється тим, що в структурі згорткових нейронних мереж закладено механізм пошуку деталей на зображенні за допомогою невеликих фільтрів. Такий механізм дає їм змогу швидко навчатися на невеликих наборах даних. Це добре видно з отриманих кривих навчання. Але коли потенційна межа близька, то додавання додаткових шарів, наприклад у класифікатора ResNet-18 порівняно з класифікатором CNN, який здійснює пошук ще більших деталей, не дає суттєвого приросту якості класифікації.

НМ, побудовані на базі трансформерів, відрізняються від згорткових мереж тим, що обробляють одразу всі патчі зображення і намагаються знайти зв'язки між ними незалежно від того, чи стоять ці патчі поруч на зображенні, чи знаходяться на різних сторонах. Така побудова архітектури призводить до того, що такі НМ навчаються повільніше і потребують великого обсягу даних для навчання. Якщо CNN вивчає всі компоненти зображення без пріоритетів будь-яким областям, то Трансформери формують карти уваги, дозволяючи виділяти області найбільш імовірного розташування об'єктів.

Під час обробки невеликих наборів даних це призводить до того, що як CNN, так і Трансформери не забезпечують достатньо високе співвідношення якості класифікації та обсягу обчислень. Однак, якщо в CNN, здебільшого, можна тільки збільшувати кількість використовуваних шарів, то в Трансформерах можливостей модифікації моделі значно більше. Це видно з характеристик класифікатора на

основі ССТ. Введення двох згорткових шарів у токенізатор і ділення на патчі, що перекриваються, дало змогу забезпечити високе співвідношення якості класифікації до обсягу обчислень при високому рівні точності.

Аналіз зображень, на яких НМ робили помилки, показує, що деякі зображення були помилково класифіковані усіма НМ. Це свідчить про те, що в наборі даних не всі можливі варіації зображень мін на різному фоні представлені досить повно. Розширення набору даних подібними зображеннями має призвести до збільшення точності розпізнавання мін.

Аналіз стійкості до погіршення якості тестових зображень показує вищу залежність Трансформерів від набору даних, на якому вони навчалися. Згорткові мережі менш чутливі до шуму, освітленості та туману завдяки архітектурі, що виділяє окремі ознаки. При розширенні набору даних за рахунок включення туди зображень із погіршеною якістю можна очікувати на підвищення точності у всіх нейронних мереж.

Обмеженням проведеного дослідження є маленька величина використаного набору даних. Проте, це дозволило виявити найперспективнішу модель для створення класифікатора – ССТ. Варіюючи номенклатуру та кількість додаткових архітектурних елементів, що вводяться в Трансформер, можна суттєво підвищувати як точність, так і співвідношення точності класифікації та необхідного обсягу обчислень. Разом із розширенням кількості та різноманітності зображень мін на різному тлі це дозволить отримати максимально високі характеристики в такій важливій галузі, як гуманітарне розмінування.

4.5 Дослідження якості класифікації зображень на наявність мін з використанням запропонованого ансамблевого класифікатора

Як показали результати дослідження, найкращим базовим класифікатором є ССТ. Однак у нього на аналізованому наборі даних 3 міні пропущено і дано 4 помилкові вказівки на міні. Тому представляє інтерес використання ансамблевого класифікатора зі стекінгом НМ на основі моделей і методу, запропонованих в

розділах 2 і 3. Цей класифікатор використовує на першому ступені базовий класифікатор ССТ та його клони (див. рис. 2.4), що були отримані з застосуванням оператора циклічного зсуву, а на другому ступені – модель MLP-KAN (див. рис. 3.9). Результати класифікації наведено у табл. 4.4.

З табл. 4.4 видно, що використання ансамблевого класифікатора замість звичайної нейронної мережі дозволяє суттєво збільшити точність. Слід зазначити, що використаний набір даних з мінами містив незамасковані об'єкти, тому навіть одиночна нейронна мережа показала високі результати. У випадку, коли міни були б значною мірою замасковані, то результати одиночної НМ були б істотно нижчими, а використання ансамблевого класифікатора давало б більший приріст точності.

Таблиця 4.4 – Результати використання ансамблевого класифікатора зі стекінгом на основі клонів ССТ і моделі MLP-KAN на наборі даних з мінами

Класифікатори на першому ступені	<i>Accuracy</i>	<i>Recall</i>	<i>FN</i>	<i>FP</i>
ССТ+1 клон	0,9854	0,9854	2	2
ССТ+2клона	0,9890	0,9927	1	2
ССТ+3 клона	0,9964	1,00	0	1

4.6 Дослідження швидкодії запропонованого ансамблевого класифікатора зображень при обробці відеопослідовності

При практичній реалізації окрім якості класифікації об'єктів на зображеннях, важливе значення має обсяг обчислень або еквівалентний інтервал часу, необхідний для обчислення вихідного сигналу навченого класифікатора [7], [8].

Одним із варіантів процесорів, які можуть бути використані на дроні або іншому роботизованому комплексі, є процесор STM32MP157 [114]. Він має два

ядра Cortex-A7 з тактовою частотою 650 або 800 МГц і мікропроцесорне ядро Cortex-M4 з тактовою частотою 209 МГц. Ядра Cortex-A7 працюють під керуванням операційної системи Linux і мають блок для операцій над 32- та 64-розрядними числами з рухомою комою VFPv4. Чисті операції додавання та множення з рухомою комою з 32-розрядною точністю виконуються за 1 такт роботи ядра. Операції додавання та множення виконуються одночасно на різних вузлах. Вважаємо, що в кожному процесорі обробкою кадрів із камери займається лише одне ядро Cortex-A7 (800 млн операцій додавання і стільки ж операцій множення з рухомою комою за секунду). Друге ядро Cortex-A7 і ядро Cortex-M4 займаються обробкою сигналів інших датчиків та керуванням дрону роботизованого комплексу. У табл. 4.5 наведено час обробки одного кадру та видачі результату класифікації на наборі з зображеннями мін для двох варіантів реалізації:

- 1) увесь класифікатор реалізовано на одному ядрі;
- 2) кількість ядер Cortex-A7 дорівнює кількості НМ на першому ступені класифікатора.

У другому варіанті кожна НМ першого ступеня обробляється на своєму ядрі, а загальний час обчислення для першого ступеня ансамблевого класифікатора зі стекінгом НМ дорівнює часу обробки НМ з максимальною кількістю вагових коефіцієнтів.

З табл. 4.5 видно, що при використанні на першому ступені ансамблевого класифікатора зі стекінгом різних базових НМ час обробки одного кадру при використанні одного ядра Cortex-A7 може бути дещо більшим або меншим за час обробки одного кадру при використанні ССТ та його клонів. Проте відмінності становлять не більше ніж 26%. Однак при використанні для кожної НМ першого ступеня окремого ядра Cortex-A7 застосування класифікатора ССТ та його клонів забезпечує вищі показники, починаючи з 4 мереж на першому ступені. Це пояснюється тим, що при використанні окремого ядра для кожної НМ загальний час визначається НМ з найбільшою кількістю вагових коефіцієнтів.

Таблиця 4.5 – Порівняння часу обробки одного кадру при класифікації

Різні алгоритми на першому ступені			ССТ та його клони на першому ступені		
Алгоритми на першому ступені (Загальна кількість вагових коефіцієнтів)	Час на 1 кадр(мс) (кількість ядер Cortex A7)		Алгоритми на першому ступені (Загальна кількість вагових коефіцієнтів)	Час на 1 кадр (мс) (кількість ядер Cortex A7)	
CCT+EATNet (764299)	0.96 (1)	0.51 (2)	CCT+CL(0.24) (828908)	1.04 (1)	0.51 (2)
CCT+EATNet +MLP-Mixer (984157)	1.23 (1)	0.51 (3)	CCT +CL(0.24)+CL(-0.24) (1249247)	1.56 (1)	0.51 (3)
CCT+EATNet +MLP-Mixer+FNet (1567277)	1.96 (1)	0.73 (4)	CCT +CL(0.24)+CL(-0.24) +CL(0.48) (1670096)	2.09 (1)	0.51 (4)
CCT+EATNet +MLP-Mixer+FNet +gMLP (2430405)	3.04 (1)	1.08 (5)	CCT +CL(0.24)+CL(-0.24) +CL(0.48)+CL(-0.48) (2091145)	2.61 (1)	0.51 (5)
CCT+EATNet +MLP-Mixer+FNet +gMLP+SwinTr (2582901)	3.23 (1)	1.08 (6)	CCT +CL(0.24)+CL(-0.24) +CL(0.48)+CL(-0.48) +CL(0.72) (2512394)	3.14 (1)	0.51 (6)

Таким чином, використання на першому ступені ансамблевого класифікатора зі стекінгом моделі ССТ та його клонів при реалізації на великій кількості обчислювальних ядер забезпечує як вищу якість класифікації, так і менший час обчислення вихідного сигналу всього ансамблевого класифікатора.

Висновки до розділу 4

1. Порівняння згорткових нейронних мереж і мереж на базі трансформерів показало, що найперспективнішою архітектурою для використання на дронах-

міношукачах є архітектура ССТ. Вона дозволяє забезпечувати як високу точність класифікації мін, так і високе співвідношення точності до необхідного обсягу обчислень. Це є наслідком використання як самої архітектури Трансформера, так і додаткових архітектурних компонентів.

2. Аналіз стійкості до погіршення якості тестових зображень показав перевагу згорткових нейронних мереж. Це висуває вимогу до набору даних збільшити кількість і номенклатуру зображень для навчання, що має привести до підвищення стійкості всіх досліджуваних НМ.

3. На наборі даних з мінами класифікатор ССТ забезпечив якість бінарної класифікації (є міна на зображенні/немає міни) 0,9745.

4. Ансамблевий класифікатор, який використовує як базовий класифікатор на першому ступені ССТ і два його клони, дозволив збільшити точність до 0,9964 а повноту до 1,00 (всі міни були виявлені, була одна помилкова вказівка на міну).

5. Архітектура ансамблевого класифікатора зі стекінгом дозволяє ефективно розпаралелювати обчислення. При використанні по одному ядру процесора для реалізації кожного з класифікаторів першого ступеня використання на першому ступені класифікатора ССТ і 5 і більше його клонів забезпечує швидкодію 0,51 мс для обробки одного кольорового вхідного зображення розміром 32×32 пікселів, на відміну від використання базових класифікаторів, що вимагають 1,08 мс. Тобто запропонована архітектура забезпечує більш ніж у 2 рази вищу швидкодію.

ВИСНОВКИ

Дисертаційна робота містить нові наукові результати, що є наслідком розв'язання актуального наукового завдання:

- розвитку нових моделей і методу, які дозволяють збільшувати кількість НМ на першому ступені в ансамблевих класифікаторах зі стекінгом;
- розвитку нових моделей для об'єднання результатів первинних НМ на другому ступені ансамблевого класифікатора зі стекінгом.

У ході дослідження отримано такі наукові та практичні результати.

1. Проведений аналіз архітектур ансамблевих класифікаторів показав, що для оброблення неструктурованих даних, якими є зображення, найвищими потенційними характеристиками для класифікації зображень володіють ансамблеві класифікатори зі стекінгом. Основною проблемою таких класифікаторів є забезпечення на першому ступені великої кількості первинних класифікаторів, що відрізняються один від одного різним ступенем структурованості оброблення вхідних сигналів і мають високе співвідношення якості класифікації до обсягу обчислень. Проведений аналіз літератури дозволив виявити лише сім відомих НМ, які за своїми параметрами підходять для використання на першому ступені ансамблевого класифікатора зі стекінгом.

2. Проведений аналіз підходів, що є перспективними для зменшення обсягу обчислень на другому ступені ансамблевого класифікатора зі стекінгом, дозволив визначити два основні напрями: ітераційне проріджування та використання НМ KAN. Основною проблемою для першого напрямку є відсутність конструктивної методики, яка б враховувала взаємозв'язок швидкості навчання НМ, частоти проріджування та обсягу одночасно видалених зв'язків для досягнення максимальних результуючих характеристик. Основною проблемою для другого напрямку є відсутність розроблених архітектур для ефективного комплексування MLP і KAN з метою зменшення обсягу обчислень на другому ступені ансамблевого класифікатора при збереженні якості класифікації. У якості найбільш перспективного було обрано комплексування MLP і KAN.

3. Вдосконалено модель першого ступеня ансамблевого класифікатора зі стекінгом нейронних мереж для зображень з низькою роздільною здатністю в умовах дефіциту навчальних даних, шляхом додавання до заданої кількості базових класифікаторів їх клонів, отриманих навчанням на вибірках зображень, які створено за визначеними операторами перетворення, що дозволило запропонувати метод визначення структурно-параметричних ознак моделі.

4. На прикладі алгоритму MLP-Mixer і набору даних CIFAR-10 показано, що збільшення кількості класифікаторів у ширину (шляхом збільшення кількості класифікаторів на першому ступені ансамблевого класифікатора) забезпечує покращення якості класифікації порівняно зі збільшенням у глибину за практично однакового обсягу обчислень. Експериментальні результати показали підвищення якості класифікації від 0,7% до 3,9%.

5. Показано, що при роботі з набором даних CIFAR-10 додавання до першого ступеня двох клонів алгоритму MLP-Mixer, отриманих методом повороту вхідних зображень, забезпечує зменшення помилки від 1% до 11 % залежно від кількості базових класифікаторів на першому ступені. Аналогічно для SST зменшення помилки становило від 2.1% до 10%.

6. Показано, що при використанні на першому ступені ансамблевого класифікатора клонів MLP-Mixer, побудованих шляхом циклічного зсуву рядків або стовпців вхідних зображень, збільшення кількості клонів від 0 до 12 забезпечило підвищення якості класифікації ансамблевого класифікатора з 0.7674 до 0.8029, тобто на 4.6%. Відповідно, кількість помилок зменшилася з 2326 до 1971, або на 15%. Для клонів SST підвищення якості класифікації становило з 0.8021 до 0.8492, або на 5.9%, при цьому кількість помилок зменшилася з 1979 до 1508, або на 23.8%.

7. Вперше розроблено модель НМ MLP-KAN, що поєднує апроксимальні властивості MLP і KAN. Модель призначена для задач класифікації зображень загального типу. Результати експериментальних досліджень показали, що при тій самій якості класифікації, що й у MLP, модель MLP-KAN за належних параметрів потребує у 3,63 рази менше обсягу обчислень.

8. Запропоновано метод синтезу другого ступеня ансамблевого класифікатора, який відрізняється від існуючих використанням розробленої нейронної мережі MLP-KAN з урахуванням особливостей вихідного вектора ознак першого ступеня ансамблевого класифікатора. Показано, що така модель забезпечує зменшення помилки класифікації на 0.45%-1.9% при меншій кількості навчених коефіцієнтів у 1.69-3.7 рази порівняно з MLP.

6. На прикладі малого набору даних із зображеннями мін, що складається всього з 2000 зображень, показано, що ансамблевий класифікатор зі стекінгом, що використовує на першому ступені SST та його клони, може бути успішно навчений і дозволяє отримати більш високі значення точності порівняно з одиночним класифікатором SST.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Galchonkov, O., Baranov, O., Babych, M., Kuvaieva, V., Babych, Y. Improving the quality of object classification in images by ensemble classifiers with stacking. *Eastern-European Journal of Enterprise Technologies*, 2023, 3 (9 (123)), 70–77.
DOI: <https://doi.org/10.15587/1729-4061.2023.279372> (Indexed in SCOPUS)
2. Galchonkov, O., Baranov, O., Chervonenko, P., Babilunga, O. Reducing the volume of computations when building analogue of neural networks for the first stage of an ensemble classifier with stacking. *Eastern-European Journal of Enterprise Technologies*, 2024, 2 (9 (128)), 27–35.
DOI: <https://doi.org/10.15587/1729-4061.2024.299734> (Indexed in SCOPUS)
3. Galchonkov, O., Baranov, O., Antoshchuk, S., Maslov, O., Babych, M. Development of a neural network with a layer of trainable activation functions for the second stage of the ensemble classifier with stacking. *Eastern-European Journal of Enterprise Technologies*, 2024, 5 (9 (131)), 6–13.
DOI: <https://doi.org/10.15587/1729-4061.2024.311778> (Indexed in SCOPUS)
4. Galchonkov, O., Baranov, O., Maslov, O., Babych, M., Baskov, I. (2025). MLP-KAN: implementation of the Kolmogorov-Arnold layer in a multilayer perceptron. *Eastern-European Journal of Enterprise Technologies*, 3 (4 (135)), 34–41.
DOI: <https://doi.org/10.15587/1729-4061.2025.328928> (Indexed in SCOPUS)
5. Galchonkov O. N., Baranov A. M., Baskov I. O. Reducing the amount of computations at the second stage of an ensemble classifier with stacking on mine-detection drones. *Applied Aspects of Information Technology*. 2024; Vol.7 No.4: 370–377.
DOI: <https://doi.org/10.15276/aait.07.2024.26> (Indexed in Copernicus)
6. Galchonkov O., Baranov O., Baskov I. Improving the quality of mine detection by a mine-detecting drone by selecting an architecture for classifying objects in visible-light camera images. *Applied Aspects of Information Technology* 2026; Vol.9 No.1: 9–24.
DOI: <https://doi.org/10.15276/aait.09.2026.01> (Indexed in Copernicus)

7. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babilunga O., Shunin O. Using Stacked Ensemble Classifiers to Speed Up Object Recognition in Images for Mine Detection Drones. Proceedings of 2024 IEEE 7th International Conference on ACTUAL PROBLEMS OF UNMANNED AERIAL VEHICLES DEVELOPMENTS (APUAVD). October 22-24, 2024 Kyiv, Ukraine. P.262-265.
DOI: <https://ieeexplore.ieee.org/xpl/conhome/10765843/proceeding> (Indexed in SCOPUS)
8. Galchonkov O., Baranov O., Bascov I. Building a Stacked Ensemble Object Classifier on Images Using the Base CCT Classifier and Its Clones. Proceedings of 2025 IEEE 8th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC). October 21-24, 2025 Kyiv, Ukraine. P.280-284.
DOI: 10.1109/MSNMC61130.2025.11399700 (Indexed . in SCOPUS)
9. Yang, J., Shi, R., Wei, D., Liu, Z., Zhao, L., Ke, B., Ni, B. MedMNIST v2-A large-scale lightweight benchmark for 2D and 3D biomedical image classification. // Scientific Data, 2023, 10(1), 41.
10. Islam M.R., Nahiduzzaman M., Goni, M.O.F. Sayeed, A., Anower M.S., Ahsan M., Haider J. Explainable Transformer-Based Deep Learning Model for the Detection of Malaria Parasites from Blood Cell Images. // Sensors 2022, 22, 4358.
<https://doi.org/10.3390/s22124358>
11. Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., Xie, P. COVID-CTdataset: A CT scan dataset about COVID-19. // arXiv:2003.13865 doi:10.48550/arXiv.2003.13865
12. Zhang, K., Liu, X., Shen, J., Li, Z., Sang, Y., Wu, X., et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of Covid-19 pneumonia using computed tomography. // Cell 181, 1360–1433.
doi:10.1016/j.cell.2020.08.029
13. Zhang, Z.; Zhu, L. A Review on Unmanned Aerial Vehicle Remote Sensing: Platforms, Sensors, Data Processing Methods, and Applications. Drones 2023, 7, 398, 42p. <https://doi.org/10.3390/drones7060398>
14. Zhou, Z.-H. (2012). Ensemble Methods: Foundations and Algorithms. CRC Press, Taylor & Francis Group. 280p.

15. Breiman, Leo. Bagging Predictors (англ.) // :en:Machine Learning (journal)|Machine Learning : journal. — 1996. — Vol. 24, no. 2. — P. 123—140. — doi:10.1007/BF00058655
16. Friedman J.H. Greedy Function Approximation: A Gradient Boosting Mashine. IMC 1999 Reitz Lecture , February 24, 1999. / <https://web.archive.org/web/20191101082737/http://statweb.stanford.edu/~jhf/ftp/trebst.pdf>
17. Tianqi Chen, Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. ArXiv: 1603.02754v3 [cs.LG] 10 Jun 2016.
18. Guolin Ke¹, Qi Meng², Thomas Finley³, Taifeng Wang¹, Wei Chen¹, Weidong Ma¹, Qiwei Ye¹, Tie-Yan Liu¹. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
19. Galchonkov. O., Babych, M., Zasidko, A., Poberezhnyi, S. (2022). Using a neural network in the second stage of the ensemble classifier to improve the quality of classification of objects in images. Eastern-European Journal of Enterprise Technologies, 3 (9 (117)), 15–21. doi: <https://doi.org/10.15587/1729-4061.2022.258187>
20. Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4):541–551, 1989.
21. LeCun Y., Cortes C., Burges C. J.C. // THE MNIST DATABASE of handwritten digits. URL: <http://yann.lecun.com/exdb/mnist/>.
22. Mariya Ouaisa, Mariyam Ouaisa, Zakaria Boulouad, Inam Ullah Khan. Machine Intelligence for Internet of Medical Things: Applications and Future Trends. Bentham Science Publishers, 2023, 656 p.
23. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE.// ArXiv: 2010.11929v2 [cs.CV] 3 Jun 2021.

24. Hassani A., Walton S., Shah N., Abuduweili A., Li J., Shi H. Escaping the Big Data Paradigm with Compact Transformers.// ArXiv: 2104.05704v3 [cs.CV] 13 Aug 2021.
25. Guo, M.-H., Liu, Z.-N., Mu, T.-J., Hu, S.-M. (2021). Beyond Self-attention: External Attention using Two Linear Layers for Visual Tasks. arXiv. doi: <https://doi.org/10.48550/arXiv.2105.02358>
26. Lee-Thorp, J., Ainslie, J., Eckstein, I., Ontañón, S. (2021). FNet: Mixing Tokens with Fourier Transforms. arXiv. doi: <https://doi.org/10.48550/arXiv.2105.03824>
27. Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.
28. Bharat Singh and Larry S Davis. An analysis of scale invariance in object detection snip. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3578–3587, 2018.
29. Bharat Singh, Mahyar Najibi, and Larry S Davis. Sniper: Efficient multi-scale training. In Advances in Neural Information Processing Systems, volume 31. Curran Associates, Inc., 2018.
30. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. Advances in Neural Information Processing Systems, 30:5998–6008, 2017.
31. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z. et. al. (2021). Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. arXiv. doi: <https://doi.org/10.48550/arXiv.2103.14030>
32. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.
33. Liu, H., Dai, Z., So, D. R., Le, Q. V. (2021). Pay Attention to MLPs. arXiv. doi: <https://doi.org/10.48550/arXiv.2105.08050>
34. Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. In ICLR, 2015

35. Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.
36. Tolstikhin, I., Houlsby, N., Kolesnikov, A., Beyer, L., Zhai, X., Unterthiner, T. et. al. (2021). MLP-Mixer: An all-MLP Architecture for Vision. arXiv. doi: <https://doi.org/10.48550/arXiv.2105.01601>
37. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016.
38. J. L. Ba, J. R. Kiros, and G. E. Hinton. Layer normalization. arXiv preprint arXiv:1607.06450, 2016.
39. Krizhevsky A. The CIFAR-10 dataset.
URL: <https://www.cs.toronto.edu/~kriz/cifar.html>
40. Liu Z., Wang Y., Vaidya S., Ruehle F., Halverson J., Soljacić M., Hou T.Y., Tegmark M.. KAN: Kolmogorov–Arnold Networks. // arXiv: 2404.19656v1 [cs.LG] 30 Apr 2024.
41. G. Hinton, O. Vinyals, J. Dean, Distilling the knowledge in a neural network, In NIPS Deep Learning and Representation Learning Workshop, 2015.arXiv: 1503.02531 (2015).
42. C. Li, J. Peng, L. Yuan, G. Wang, X. Liang, L. Lin, and X. Chang. Blockwisely supervised neural architecture search with knowledge distillation, 2019.
43. Yonathan Aflalo*, Asaf Noy, Ming Lin, Itamar Friedman, Lihi Zelnik. Knapsack Pruning with Inner Distillation . arXiv: 2002.08258v3[cs.LG] 3 Jun 2020
44. Zhenyu Wang, Fu Li *, Guangming Shi, Xuemei Xie, Fangyu Wang. Network pruning using sparse learning and genetic algorithm . *Neurocomputing* 404 (2020) 247–256 journal homepage: www.elsevier.com/locate/neucom
45. E.L. Denton , W. Zaremba , J. Bruna , Y. LeCun , R. Fergus , Exploiting linear structure within convolutional networks for efficient evaluation, in: *Advances in Neural Information Processing Systems*, 2014, pp. 1269–1277 . . arXiv: 1404.0736v2[cs.CV] 9 Jun 2014

46. Yawei Li¹, Shuhang Gu¹, Christoph Mayer¹, Luc Van Gool^{1;2}, Radu Timofte¹
Group Sparsity: The Hinge Between Filter Pruning and Decomposition for Network
Compression . arXiv: 2003.08935v1[cs.CV] 19 Mar 2020
47. S. Han, H. Mao, W.J. Dally, Deep Compression: Compressing Deep Neural
Networks with Pruning, Trained Quantization and Huffman Coding, arXiv: 1510. 00149
(2015).
48. Qiu J., Wang J., Yao S., Guo K., Li B., Zhou E., Yu J., Tang T., Xu N., Song S.,
Wang Y., Yang H.. (2016) Going Deeper with Embedded FPGA Platform for
Convolutional Neural Network . FPGA'16, February 21-23, Monterey, CA, USA.
URL: <http://cadlab.cs.ucla.edu/~jaywang/papers/fpga16-cnn.pdf> ,
DOI: <http://dx.doi.org/10.1145/2847263.2847265>
49. LeCun Y., Denker J. S., Solla S. A. Optimal Brain Damage.// NIPS, 1990. URL:
<http://yann.lecun.com/exdb/publis/pdf/lecun-90b.pdf>
50. Blalock D., Gonzalez Ortiz J. J., Frankle J., Gutttag J. WHAT IS THE STATE OF
NEURAL NETWORK PRUNING? .- arXiv: 2003.03033v1 [cs.LG] 6 Mar 2020.
51. Han S., Pool J., Tran J., Dally W. J.. Learning both Weights and Connections for
Efficient Neural Networks //ArXiv: 1506.02626v3 [cs.NE] 30 Oct 2015. URL:
<https://arxiv.org/pdf/1506.02626v3.pdf>
52. Christos Louizos, Max Welling, Diederik P. Kingma. LEARNING SPARSE
NEURAL NETWORKS THROUGH L0 REGULARIZATION. ICLR 2018, arXiv:
1712.01312v2[stat.ML] 22 Jun 2018.
53. Jiashi Li, Qi Qi, Jingyu Wang, Ce Ge, Yujian Li, Zhangzhang Yue, and Haifeng
Sun. OICSR: Out-In-Channel Sparsity Regularization for Compact Deep Neural
Networks. arXiv: 1905.11664v5[cs.CV] 1 Jul 2019.
54. Aidan N. Gomez, Ivan Zhang , Siddhartha Rao Kamalakara, Divyam Madaan, Kevin
Swersky, Yarin Gal, Geoffrey E. Hinton. Learning Sparse Networks Using Targeted
Dropout. arXiv: 1905.13678v5[cs.LG] 9 Sep 2019.
55. Tabik S., Peralta D., Herrera-Poyatos A., Herrera F..(2017) A snapshot of image pre-
processing for convolutional neural networks:case study of MNIST. International
Journal of Computational Intelligence Systems, Vol. 10 , 555–568.

56. Cireşan D., Meier U., Gambardella L. M., Schmidhuber J. (2010) Deep, big, simple neural nets for handwritten digit recognition . *Neural Computation*, vol. 22, 3207–3220,
57. Simard P. Y., Steinkraus D., Platt, J.C. (2003) Best practices for convolutional neural networks applied to visual document analysis . *Intl. Conf. Document Analysis and Recognition*. San Mateo, CA: IEEE Computer Society Press. 958–962.
58. Galchonkov O., Nevrev A., Glava M., Babych M. EXPLORING THE EFFICIENCY OF THE COMBINED APPLICATION OF CONNECTION PRUNING AND SOURCE DATA PRE-PROCESSING WHEN TRAINING A MULTILAYER PERCEPTRON // *Eastern – European Journal of Enterprise Technologies . Information and controlling system– 2020. - VOL 2, NO 9 (104) -P. 6-13*
59. Galchonkov, O., Nevrev, A., Shevchuk, B., Baranov, N. (2022). Definition of the influence of the choice of the pruning procedure parameters on the quality of training of a multilayer perceptron. *Eastern-European Journal of Enterprise Technologies*, 1 (9 (115)), 75–83. doi: <https://doi.org/10.15587/1729-4061.2022.253103>
60. Braun J., Griebel M. (2009) On a Constructive Proof of Kolmogorov’s Superposition Theorem. *Constructive Approximation*, 30:653–675, 2009. DOI 10.1007/s00365-009-9054-2
61. Haihong Guo^{1,2,3}, Fengxin Li^{1,3}, Jiao Li^{2,*}, Hongyan Liu. KAN v.s. MLP for Offline Reinforcement Learning. // arXiv: 2409.09653v1 [cs.LG] 15 Sep 2024.
62. Runpeng Yu, Weihao Yu, and Xinchao Wang. KAN or MLP: A Fairer Comparison. // arXiv: 2407.16674v1 [cs.LG] 23 Jul 2024.
63. Yunhong He* Yifeng Xie* Zhengqing Yuan² Lichao Sun. MLP-KAN: UNIFYING DEEP REPRESENTATION AND FUNCTION LEARNING. // arXiv: 2410.03027v1 [cs.LG] 3 Oct 2024.
64. Kunpeng Xu, Lifei Chen, Shengrui Wang. KOLMOGOROV-ARNOLD NETWORKS FOR TIME SERIES: BRIDGING PREDICTIVE POWER AND INTERPRETABILITY. // arXiv: 2406.02496v1 [cs.LG] 4 Jun 2024.
65. Yuan Gao^a, Zehuan Hu^b, Wei-An Chen^c, Mingzhe Li^d, Yingjun Ruane. A revolutionary neural network architecture with interpretability and flexibility based on

- Kolmogorov–Arnoldforsolarradiation andtemperature forecasting. *Applied Energy* 378 (2025) 124844. <https://doi.org/10.1016/j.apenergy.2024.124844>
66. Ali Jamali, Swalpa Kumar Roy, Danfeng Hong, Bing Lu, Pedram Ghamis. How to Learn More? Exploring Kolmogorov–Arnold Networks for Hyperspectral Image Classification. *Remote Sensing*, 2024, 16, 4015. <https://doi.org/10.3390/rs16214015>.
67. Swalpa Kumar Roy; Gopal Krishna; Shiv Ram Dubey; Bidyut B. Chaudhuri. HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters* (Volume: 17, Issue: 2, February 2020), Page(s): 277 – 281. DOI: 10.1109/LGRS.2019.2918719
68. Wei Li, Lu Li, Man Peng and Ran Tao. KANDiff: Kolmogorov–Arnold Network and Diffusion ModelBased Network for Hyperspectral and Multispectral Image Fusion/Remote Sensing, 2025, 17, 145, p.1-25. <https://doi.org/10.3390/rs17010145>
69. Guo L., Wang Y., Guo M., Zhou X. YOLO-IRS: Infrared Ship Detection Algorithm Based on Self-Attention Mechanism and KAN in Complex Marine Background. *Remote Sens.* 2025, 17, 20.p.1-19. <https://doi.org/10.3390/rs17010020>
70. Mohamed Abd Elaziz; Ibrahim Ahmed Fares; Ahmad O. Aseeri. CKAN: Convolutional Kolmogorov–Arnold Networks Model for Intrusion Detection in IoT Environment. *IEEE Access* (Volume: 12), 2024, p. 134837 – 134851. DOI: 10.1109/ACCESS.2024.3462297
71. Bodner A.D., Spolski J.N., Tepsich A.S., Pourteau S.. Convolutional Kolmogorov–Arnold Networks. // arXiv: 2406.13155v1 [cs.CV] 19 June 2024.
72. Fashion MNIST. <https://github.com/zalandoresearch/fashion-mnist>
73. Drokin I. Kolmogorov-Arnold Convolutions: Design Principles and Empirical Studies. // arXiv: 2407.01092v1 [cs.CV] 1 Jul 2024.
74. Wang, C.; Zhang, X.; Liu, L. FloodKAN: Integrating Kolmogorov–Arnold Networks for Efficient Flood Extent Extraction. *Remote Sensing*. 2025, 17, 564, p.1-13. <https://doi.org/10.3390/rs17040564>
75. Seydi S.T. Exploring the Potential of Polynomial Basis Functions in Kolmogorov–Arnold Nenworks: A Comparative Study of Different Groups of Polynomials. // arXiv: 2406.02583v1[cs.LG] 30 May 2024.

76. Ismayilova A., Ismayilov M.. On the universal approximation property of radial basis function neural networks. // arXiv: 2304.02220v2 [cs.LG] 8 Apr 2023.
77. Zan-Rong He, Yan-Ting Lin, Shie-Jue Lee, Chih-Hung Wu. A RBF Network Approach for Function Approximation. // Proceeding of the IEEE International Conference on Information and Automation, Wuyi Mountain, China, August 2018, p.105-109.
78. Panda S., Panda G. On the Development and Performance Evaluation of Improved Radial Basis Function Neural Networks. // IEEE Trans. ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, June 2022, v.52, Issue: 6, p. 3873 – 3884, DOI: 10.1109/TSMC.2021.3076747.
79. Abd-Krim Seghouane, Navid Shokouhi. Adaptive Learning for Robust Radial Basis Function Networks. // IEEE Trans. ON CYBERNETICS, May 2021, v. 51, Issue: 5, p. 2847 – 2856, DOI: 10.1109/TCYB.2019.2951811
80. Chen Wu, Xian-ren Kong, Zhen-guo Yang. (2018) An online self-adaption learning algorithm for hyper basis function neural network. // 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC 2018), p.215-220.
81. Brownlee J. Better Deep Learning. Copyright 2019 Jason Brownlee, 575p. URL: <https://machinelearningmastery.com/better-deep-learning/>
82. Keras. Code examples. Computer vision.
URL: <https://keras.io/examples/vision/>
doi: <https://doi.org/10.15587/1729-4061.2022.258187>
83. Brownlee J. Weight Initialization for Deep Learning Neural Networks.// URL: <https://machinelearningmastery.com/weight-initialization-for-deep-learning-neural-networks/>.
84. Colab //URL: <https://colab.research.google.com/notebooks/welcome.ipynb>
85. Brownlee J. Deep Learning for Computer Vision. Image Classification, Object Detection and Face Recognition in Python.// Copyright 2019.- Jason Brownlee.- 563 p. <https://machinelearningmastery.com/deep-learning-for-computer-vision/>
86. Minjong Cheon. DEMONSTRATING THE EFFICACY OF KOLMOGOROV-

ARNOLD NETWORKS IN VISION TASKS. // arXiv: 2406.14916v1 [cs.CV] 21 Jun 2024.

87. Fan Zhang, Xin Zhang. GraphKAN: Enhancing Feature Extraction with Graph Kolmogorov Arnold Networks. // arXiv: 2406.13597v1 [cs.LG] 19 June 2024.

88. Jinfeng Xu, Zheyu Chen, Jinze Li, Shuo Yang, Wei Wang, Xiping Hu, Edith C.-H. Ngai. FourierKAN-GCF: Fourier Kolmogorov-Arnold Network – An Effective and Efficient Feature Transformation for Graph Collaborative Filtering. // arXiv: 2406.01034v2 [cs.IR] 4 June 2024.

89. Vaca-Rubio C.J., Blanco L., Pereira R., Caus M. Kolmogorov-Arnold Networks (KANs) for Time Series Analysis. // arXiv: 2405.08790v1 [eess.SP] 14 May 2024.

90. Inzirillo H., Genet R. SigKAN: Signature-Weighted Kolmogorov-Arnold Networks for Time Series. // arXiv: 2406.17890v1 [cs.LG] 25 June 2024.

91. Li Ziyao. Kolmogorov-Arnold Networks are Radial Basis Function Networks. // arXiv: 2405.06721v1 [cs.LG] 10 May 2024.

92. Seydi S.T. Unveiling the Power of Wavelets: A Wavelet-based Kolmogorov-Arnold Network for Hyperspectral Image Classification. // arXiv: 2406.07869v1[cs.CV] 12 Jun 2024.

93. Qi Qiu, Tao Zhu, Helin Gong, Liming Chen, Huansheng Ning. ReLU-KAN: New Kolmogorov-Arnold Networks that Only Need Matrix Addition, Dot Multiplication, and ReLU. // arXiv: 2406.02075v2[cs.LG] 12 Aug 2024.

94. Hao Hao, Xiaoqun Zhang, Bingdong Li, Aimin Zhou. A First Look at Kolmogorov-Arnold Networks in Surrogate-assisted Evolutionary Algorithms. // arXiv: 2405.16494v1 [cs.NE] 26 May 2024.

95. Corke, P. Robotics, Vision and Control: Fundamental Algorithms in MATLAB / P. Corke. — 2nd ed. — Cham: Springer, 2017. — 650 p.

96. Springer Handbook of Robotics / B. Siciliano, O. Khatib, L. Villani, J. Santos-Victor. — 2nd ed. — Berlin: Springer-Verlag, 2016. — 2240 p.

97. Szeliski, R. Computer Vision: Algorithms and Applications / R. Szeliski. — 2nd ed. — Cham: Springer Nature, 2022. — 950 p.

98. LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. — 2015. — Vol. 521, № 7553. — P. 436–444.
99. Redmon, J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — P. 779–788.
100. Shi, W. Edge Computing: Vision and Challenges / W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu // IEEE Internet of Things Journal. — 2016. — Vol. 3, № 5. — P. 637–646.
101. Wang, Y. Generalizing from a Few Examples: A Survey on Few-Shot Learning / Y. Wang, Q. Yao, J. T. Kwok, L. M. Ni // ACM Computing Surveys. — 2020. — Vol. 53, № 3. — P. 63:1–63:34.
102. Li, C. Perceptual Generative Adversarial Networks for Small Object Detection / C. Li, X. Xu, Z. Deng, L. Meng, Q. Zhao // IEEE International Conference on Computer Vision (ICCV). — 2017. — P. 1222–1230.
103. Shaikh, S. A. Real-time object detection in constrained environments: A review / S. A. Shaikh, A. M. Hameed, R. K. Zafar // Journal of Real-Time Image Processing. — 2020. — Vol. 17, № 4. — P. 1109–1125.
104. Maalouf, M. A. Robotic Humanitarian Demining: A Review of Current Methods and Future Challenges / M. A. Maalouf, B. A. Al-Haddad, C. K. M. Lee // IEEE Robotics & Automation Magazine. — 2022. — Vol. 29, № 2. — P. 34–48.
105. Ech-Cherif, A. Unmanned Aerial Vehicles for Landmine Detection: Trends, Challenges, and Opportunities / A. Ech-Cherif, M. B. Boukhriss, R. A. Mansour // Sensors. — 2023. — Vol. 23, № 9. — P. 4412–4435.
106. Kaur, S. Deep learning for detection of buried objects: A review / S. Kaur, K. P. Singh, J. Singh // Remote Sensing of Environment. — 2021. — Vol. 263. — P. 112548.
107. Landmines Detection Dataset (by Northumbria)/ Landmine_new_v1_yolov11 <https://universe.roboflow.com/northumbria/landmines-detection-dataset>
108. <https://ru.freepik.com/free-photos-vectors/>
109. Landmine_new_64_3 <https://github.com/oleksii-m-baranov/landmines-recognition>

110. Li Yuan, Yunpeng Chen, Tao Wang, Weihao Yu, Yujun Shi, Francis EH Tay, Jiashi Feng, and Shuicheng Yan. Tokensto-token vit: Training vision transformers from scratch on imagenet.// arXiv preprint arXiv:2101.11986, 2021.
111. d’Ascoli S., Touvron H., Leavitt M., Morcos A., Biroli G., Sagun L. Convit: Improving vision transformers with soft convolutional inductive biases.// arXiv preprint arXiv:2103.10697, 2021.
112. Kun Yuan, Shaopeng Guo, Ziwei Liu, Aojun Zhou, Fengwei Yu, Wei Wu. Incorporating convolution designs into visual transformers.// arXiv preprint arXiv:2103.11816, 2021.
113. Haiping Wu, Bin Xiao, Noel Codella, Mengchen Liu, Xiyang Dai, Lu Yuan, and Lei Zhang. Cvt: Introducing convolutions to vision transformers.// arXiv preprint arXiv:2103.15808, 2021.
114. <https://www.st.com/en/microcontrollers-microprocessors/stm32mp1-series.html>

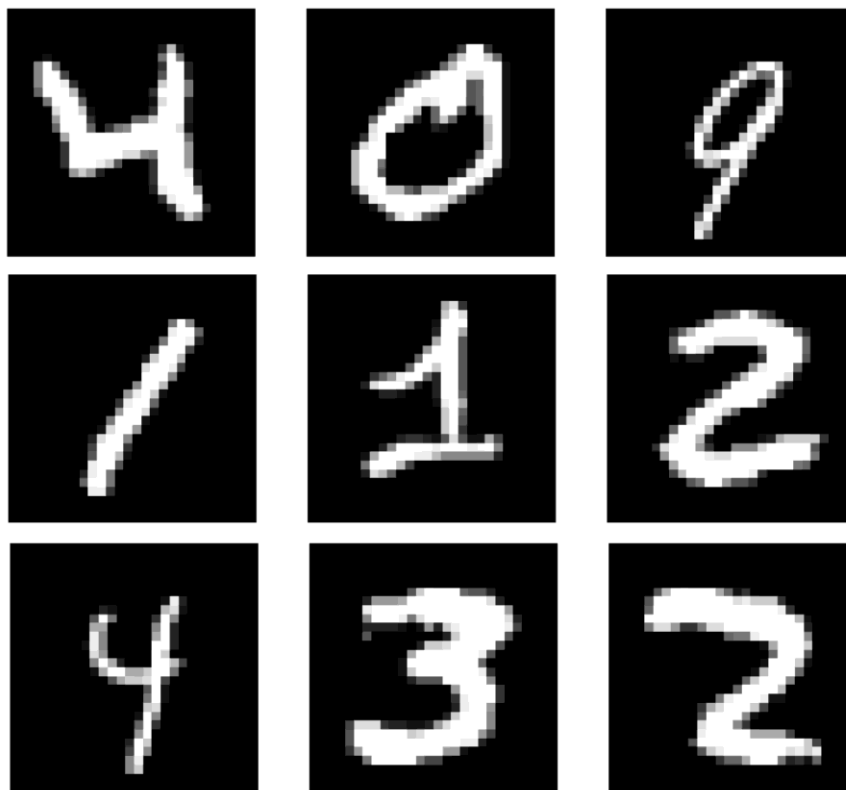
ДОДАТОК А**Приклади зображень з наборів даних MNIST, CIFAR10**

Рисунок А.1 – Приклади зображень цифр із набору MNIST [27]



Рисунок А.2 – Приклади зображень із набору CIFAR-10 [23]

ДОДАТОК Б

Список публікацій здобувача

1. Galchonkov O., Baranov O., Babych M., Kuvaieva V., Babych Y. Improving the quality of object classification in images by ensemble classifiers with stacking. Eastern-European Journal of Enterprise Technologies, 2023, 3 (9 (123)), 70–77.

DOI: <https://doi.org/10.15587/1729-4061.2023.279372> (Indexed in SCOPUS)

2. Galchonkov O., Baranov O., Chervonenko P., Babilunga O. Reducing the volume of computations when building analogue of neural networks for the first stage of an ensemble classifier with stacking. Eastern-European Journal of Enterprise Technologies, 2024, 2 (9 (128)), 27–35.

DOI: <https://doi.org/10.15587/1729-4061.2024.299734> (Indexed in SCOPUS)

3. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babych M. Development of a neural network with a layer of trainable activation functions for the second stage of the ensemble classifier with stacking. Eastern-European Journal of Enterprise Technologies, 2024, 5 (9 (131)), 6–13.

DOI: <https://doi.org/10.15587/1729-4061.2024.311778> (Indexed in SCOPUS)

4. Galchonkov O., Baranov O., Maslov O., Babych M., Baskov I. MLP-KAN: implementation of the Kolmogorov-Arnold layer in a multilayer perceptron. Eastern-European Journal of Enterprise Technologies, 2025, 3 (4 (135)), 34–41.

DOI: <https://doi.org/10.15587/1729-4061.2025.328928> (Indexed in SCOPUS)

5. Galchonkov O., Baranov O., Baskov I. Reducing the amount of computations at the second stage of an ensemble classifier with stacking on mine-detection drones. Applied Aspects of Information Technology. 2024; Vol.7 No.4: 370–377.

DOI: <https://doi.org/10.15276/aait.07.2024.26> (Indexed in Copernicus)

6. Galchonkov O., Baranov O., Baskov I. Improving the quality of mine detection by a mine-detecting drone by selecting an architecture for classifying objects in visible-light camera images. Applied Aspects of Information Technology, 2026; Vol.9 No.1: 9–24.

DOI: <https://doi.org/10.15276/aait.09.2026.01> (Indexed in Copernicus)

7. Galchonkov O., Baranov O., Antoshchuk S., Maslov O., Babilunga O., Shunin O. Using Stacked Ensemble Classifiers to Speed Up Object Recognition in Images for Mine Detection Drones. Proceedings of 2024 IEEE 7th International Conference on ACTUAL PROBLEMS OF UNMANNED AERIAL VEHICLES DEVELOPMENTS (APUAVD). October 22-24, 2024 Kyiv, Ukraine. P.262-265.

DOI: <https://ieeexplore.ieee.org/xpl/conhome/10765843/proceeding> (Indexed in SCOPUS)

8. Galchonkov O., Baranov O., Bascov I. Building a Stacked Ensemble Object Classifier on Images Using the Base CCT Classifier and Its Clones. Proceedings of 2025 IEEE 8th International Conference on Methods and Systems of Navigation and Motion Control (MSNMC). October 21-24, 2025 Kyiv, Ukraine. P.280-284.

DOI: 10.1109/MSNMC61130.2025.11399700 (Indexed in SCOPUS)

ДОДАТОК В

Акти впровадження



Довідка № 1009/64-07 від 11.06.2026

про використання результатів дисертаційної роботи

Баранова Олексія Миколайовича

**«МОДЕЛІ АНСАМБЛЕВОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ
СТЕКІНГУ НЕЙРОННИХ МЕРЕЖ»**

у науково-дослідницької діяльності

Національного університету «Одеська Політехніка»

Довідка видана в тому, що у науково-дослідницькій діяльності Національного університету «Одеська політехніка» використані наступні наукові результати, отримані у дисертаційній роботі Баранова Олексія Миколайовича:

- методи побудови ансамблевих класифікаторів зображень на основі стекінгу моделей нейронних мереж для автономних робототехнічних комплексів;
- методи побудови нейронних мереж MLP-KAN.

Дисертацію виконано згідно тематичних планів НДР Одеської політехніки за період 2016 - 2026 рр.

Баранов Олексій Миколайович був виконавцем був виконавцем у держбюджетній темі № 723-43 «Розробка інтелектуальних сенсорів дистанційного комплексного моніторингу деградації ґрунтів повоєнних територій» (номер держ. реєстрації 0126U001468).

Баранов Олексій Миколайович був виконавцем у гранті № 131/0161 Національного фонду досліджень України «Використання штучного інтелекту для підвищення якості визначення вибухових пристроїв мультисенсорним дроном-міношукачем» (номер держ. реєстрації 0124U003664).

Проректор

Дмитро ДМИТРИШИН

Керівник

НДР №723-43 та №131/0161

Світлана АНТОЩУК



«ЗАТВЕРДЖУЮ»

Перший проректор, проректор з науково-педагогічної та виховної роботи Національного університету «Одеська політехніка»
професор Сергій НЕСТЕРЕНКО



ДОВІДКА

про впровадження результатів дисертації доктора філософії
БАРАНОВА Олексія Миколайовича
«МОДЕЛІ АНСАМБЛЕВОЇ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ НА ОСНОВІ
СТЕКІНГУ НЕЙРОННИХ МЕРЕЖ»
у навчальному процесі Національного університету «Одеська політехніка»

Чинна довідка видана в тому, що в програмі, навчально-методичних матеріалах та курсу лекцій з дисципліни «Проектування інформаційних систем», який викладається бакалаврам за спеціальністю 122 «Комп'ютерні науки» на кафедрі Інформаційних систем Інституту комп'ютерних систем Національного університету «Одеська політехніка», використовуються наукові результати отримані в дисертаційній роботі Баранова О. М., а саме:

- базові алгоритми ансамблевої класифікації зображень для систем технічного зору – тема «Системотехніка інформаційних систем».
- алгоритм нейронної мережі MLP-KAN для обробки сигналів – тема «Проектування систем реального часу».

Директор
Інституту Комп'ютерних систем,
д.т.н., професор

Світлана АНТОЩУК

Голова
методичної ради ІКС
к.т.н., доцент

Анатолій НІКОЛЕНКО