UDC 004.8:004.932.2 **doi: 10.26906/SUNZ.2026.2.180**

Oleksandr Sokolov, Anton Poroshenko, Roman Yaroshevych, Vladislav Kholiev

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

APPLICATION AND ANALYSIS OF MACHINE LEARNING METHODS FOR IMAGE CLASSIFICATION

Abstract. Relevance. Image classification is a key task in computer vision, which has wide applications in medicine, transportation, industry, and security. The use of optimized CNN architectures allows high accuracy to be achieved with limited resources, which is relevant for mobile and embedded systems. **Research subject:** machine learning methods and neural network architectures for image classification. **The purpose of the article** is to develop and evaluate a modified convolutional neural network that provides a balance between classification accuracy and computational efficiency, as well as to compare its results with classical and modern models. **Research results.** The proposed CNN achieved 93.8% accuracy on the Fashion MNIST dataset, exceeding the performance of LeNet-5 (91.2%) and classical algorithms (KNN – 85.3%, Decision Tree – 82.7%, XGBoost – 90.4%). On the more complex CIFAR-10 dataset, the model showed an accuracy of 80.2%, exceeding LeNet-5 but falling short of ResNet and EfficientNet. This confirms the effectiveness of the model for tasks of medium complexity and systems with limited resources. **Conclusions.** Modified CNN is a compromise between simple classical methods and complex modern architectures. It provides an optimal balance between accuracy and learning speed, making it suitable for practical application in mobile and embedded systems. Further research may focus on the use of more complex datasets, automatic hyperparameter selection, and the integration of self-attention mechanisms. Scope of application of the results obtained: medium-complexity computer vision systems, mobile and embedded devices with limited resources, applied image classification tasks.

Keywords: image classification, CNN, Vision Transformer, Fashion MNIST, neural networks, machine learning.

Introduction

Problem statement. Image classification is one of the key tasks in computer vision, which finds application in many areas: from medicine and autonomous vehicles to the fashion industry and face recognition. Recent advances in machine learning have made it possible to create models that can analyze and classify images with high accuracy. The development of deep learning methods, such as convolutional neural networks (CNNs), as well as architectures based on the self-attention mechanism (e.g., Vision Transformer), have opened up new horizons for solving this problem. This article examines the effectiveness of various approaches to image classification based on the Fashion MNIST dataset, which contains images of clothing, footwear, and accessories. Both classical machine learning methods (K-Nearest Neighbors, Decision Tree, XGBoost) and modern deep learning models were selected for the study: LeNet-5, VGG16, ResNet, EfficientNet, Vision Transformer [1, 2]. Particular attention is paid to the presentation and analysis of my implementation, which demonstrates significant advantages over other architectures.

The aim of this work is to study existing machine learning methods and modify the classical convolutional neural network. The proposed model is a modification of the classical convolutional neural network, built on the basis of well-known architectural solutions widely used in image classification tasks. The main focus of the work is on the practical analysis of the influence of architectural parameters (number of convolutional layers, use of MaxPooling, choice of activation functions and optimization algorithm) on the classification accuracy and computational efficiency of the model. Thus, the research is applied in nature and aims to evaluate the feasibility of using simplified CNN architectures in tasks with limited computational resources. A comparative analysis of the proposed model with other approaches allows us to assess how effective my implementation is in visual data analysis tasks

Main material

Analysis of methods. To achieve the set goal, a number of classification methods were applied, covering both classical machine learning algorithms and modern neural networks. Classic approaches include KNN, Decision Tree, and XGBoost. These methods are based on geometric and statistical classification principles, but their application to image processing tasks is limited due to the local nature of feature computation and insufficient generalization ability on complex datasets.

Neural networks, in particular ResNet, EfficientNet, and Vision Transformer, demonstrate higher efficiency due to their ability to consider both local and global features of images. Within the scope of this study, both the modern architectures and the implemented convolutional neural network were used to classify images from the Fashion MNIST dataset.

The Fashion MNIST dataset consists of 28×28 pixel grayscale images and contains 10 classes corresponding to different categories of clothing, footwear, and accessories [3]. The data is divided into a training sample of 60,000 images and a test sample of 10,000 images. In addition, the study used the CIFAR-10 dataset, which contains 32×32 pixel color images belonging to 10 classes, including vehicles, animals, and everyday objects. The use of CIFAR-10 made it possible to evaluate the ability of models to work with more complex and diverse visual data.

Classic machine learning methods have demonstrated limited effectiveness in image classification tasks. The KNN algorithm is simple to implement [4], but its performance decreases on large samples due to significant computational complexity. Decision Tree provides convenient interpretation of results but is prone to overfitting. XGBoost shows better results thanks to its ensemble approach but requires significant computational resources. In contrast, neural networks are capable of automatically forming complex features from input data, making them more suitable for image analysis.

Modified CNN. The implemented convolutional neural network was optimized (Fig. 1) to work with relatively simple datasets. Its architecture includes several convolutional layers, MaxPooling operations to reduce feature dimensions, as well as modern ReLU activation functions and the Adam optimization algorithm. This combination allows for improved classification accuracy without a significant increase in computational costs. The proposed model is built in accordance with generally accepted principles of CNN development and does not contain fundamentally new architectural components, which ensures the correctness of comparing its results with classical and modern models such as LeNet-5, ResNet, EfficientNet, and Vision Transformer [6].

The general structure of the proposed convolutional neural network is shown in Fig. 1.

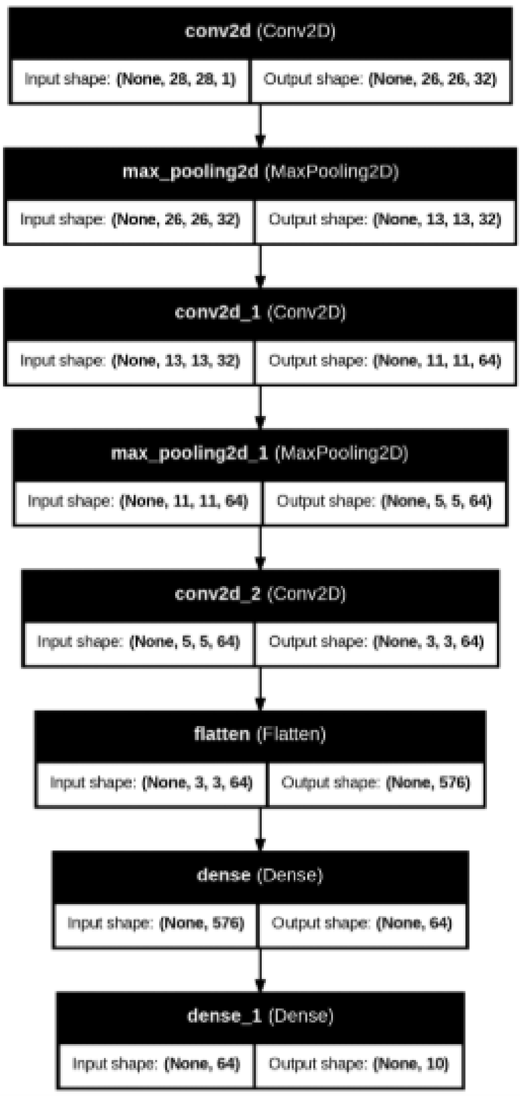

Fig. 1. Model architecture

The architecture was designed with an emphasis on simplicity and computational efficiency while preserving sufficient representational capacity for image classification tasks. Particular attention was paid to the selection of the number of layers and their configuration in order to ensure stable training and reliable performance on benchmark datasets.

The design of the modified convolutional neural network was guided by the need to achieve a balance between model simplicity and classification performance. Unlike deep modern architectures that rely on many layers and complex structural elements, the proposed CNN focuses on a limited number of carefully selected components that are sufficient for effective feature extraction. This approach allows the model to remain computationally efficient while avoiding excessive overfitting, which is especially important when working with relatively small and low-resolution datasets such as Fashion MNIST.

Results. The results of the study showed that this model demonstrates significant improvements in accuracy compared to classical machine learning methods and the classical LeNet-5. On the Fashion MNIST dataset, the developed CNN achieved an accuracy of 93.8%, while the classic version of LeNet-5 achieved only 91.2% (Table 1). This improvement was achieved using additional convolutional layers, modern activation functions (ReLU), and the Adam optimization algorithm.

After the convolutional feature extraction stage, the network includes fully connected layers that perform the final classification. The Adam optimization algorithm was selected due to its ability to adaptively adjust the learning rate, which leads to faster convergence and more stable training. Overall, the proposed architecture follows well-established CNN design principles and serves as a representative baseline for comparison with both classical and modern deep learning models.

Classic machine learning methods such as KNN, Decision Tree, and XGBoost demonstrated significantly lower performance. KNN achieved 85.3% accuracy, which can be explained by its high computational complexity and low ability to work with high-dimensional data. Decision Tree demonstrated 82.7% accuracy but suffered from overfitting. XGBoost achieved 90.4% accuracy thanks to its powerful ensemble approach but required significant computational resources and did not outperform neural networks.

Modern neural networks such as ResNet, EfficientNet, and Vision Transformer showed the highest results. ResNet achieved 95.5% accuracy by using residual connections that ensure stable training of deep models [8]. EfficientNet achieved 96.1% accuracy by optimizing the width, depth, and resolution of the model. Vision Transformer showed the best accuracy 96.3% using a self-attention mechanism that allows it to analyze the global context in images. However, these models require significantly more computational resources, which limits their application on less powerful hardware systems.

Discussion. It should be noted that the results of this study were obtained from a limited number of datasets, which may affect the generalizability of the conclusions.

Table 1 – **Test results**

	KNN	Decision tree	XG-Boost	Dence NN	LeNet-5	**My CNN**	VGG16	Res-Net	Efficient Net-80	Vision Transformer
Accuracy FASHION	85	75	88	90	92	**94**	94	96	96	97
Loss FASHION	-	-	0.6	0.4	0.35	**0.3**	0.3	0.25	0.2	0.18
Accuracy CIFAR-10	40	30	50	65	75	**80**	88	93	94	94
Loss CIFAR-10	-	-	1.2	1.1	0.8	**0.44**	0.4	0.4	0.35	0.3

In particular, the Fashion MNIST dataset contains relatively simple images with low resolution, which makes the classification task easier compared to real-world application scenarios. In addition, the selection of model hyperparameters was done empirically and did not cover the entire space of possible configurations. In future studies, it would be advisable to expand the set of experiments using more complex datasets and methods for automatic hyperparameter selection.

Although CNN is inferior in accuracy to ResNet, EfficientNet, and Vision Transformer, it has several key advantages. It provides an optimal balance between accuracy and computational efficiency, making it ideal for tasks of medium complexity and systems with limited resources. For example, the training time for this network is significantly shorter than for EfficientNet and Vision Transformer, allowing the model to be adapted to new tasks more quickly.

On the CIFAR-10 dataset, the implementation showed an accuracy of 80.2%, which surpasses the classic LeNet-5 (78.1%) but is inferior to modern models such as ResNet (91.3%) and EfficientNet (92.5%). This indicates that it is effective for tasks of low and medium complexity, but for high complexity, it is recommended to use deeper models with global data processing.

It should be noted that the results of this study were obtained from a limited number of datasets, which may affect the generalizability of the conclusions. In particular, the Fashion MNIST dataset contains relatively simple images with low resolution, which makes the classification task easier compared to real-world application scenarios.

In addition, the selection of model hyperparameters was performed empirically and did not cover the entire space of possible configurations. In further research, it is advisable to expand the set of experiments using more complex datasets and methods of automatic hyperparameter selection.

Conclusions

Thus, the proposed model demonstrates practical effectiveness as a compromise solution between simple classical machine learning methods and complex modern deep models. This confirms the feasibility of using similar CNN architectures in applied image classification tasks of medium complexity.

In this article, the following results were obtained:

1. A comparative analysis of classical machine learning methods and modern deep learning models for image classification was conducted using the Fashion MNIST and CIFAR-10 datasets, demonstrating the limitations of traditional approaches when applied to visual data.

2. A modified convolutional neural network based on standard CNN architectural principles was implemented and evaluated. The proposed model showed improved classification accuracy compared to the classical LeNet-5 architecture while maintaining lower computational complexity than deeper modern models.

3. Experimental results confirmed that the proposed CNN provides a balanced trade-off between accuracy and computational efficiency, making it suitable for image classification tasks of low to medium complexity, especially in environments with limited computational resources.

Future research may focus on extending the proposed approach by exploring more complex datasets, applying automated hyperparameter optimization techniques, and integrating advanced architectural components such as attention mechanisms or lightweight residual connections to further improve model performance and generalization ability.

From a practical perspective, the modified CNN can be effectively applied in mobile applications, embedded vision systems, and industrial monitoring tasks, where a compromise between accuracy, inference speed, and memory consumption is required. This confirms the relevance of lightweight convolutional models in modern computer vision pipelines.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

1. Lebedev G. S., et al. (2020). Deep machine learning (artificial intelligence) in ultrasound diagnostics. Journal of Telemedicine and E-Health. Vol. 2020, No. 2, pp. 22–29. URL: https://doi.org/10.29188/2542-2413-2020-6-2-22-29
2. Polischuk M., Kostyuchko S., Khrystynets M. (2019). Porivnyannya metodiv optimizaciyi neyronnyh merezh na prykladi zadachi klasifikaciyi zobrazen'. [Comparison of Neural Network Optimization Methods on the Example of an Image Classification Task]. Komp'yuterno-integrovani texnologyi: osvita, nauka, vyrobnytstvo. No. 37, pp. 43–52. URL:https://doi.org/10.36910/6775-2524-0560-2019-37-7

3. Image classification using multiple convolutional neural networks on the fashion-mnist dataset / O. Nocentini та ін. *Sensors*. 2022. Т. 22, № 23. С. 9544. URL: https://doi.org/10.3390/s22239544
4. Liu Q. The development of image classification algorithms based on cnns. *Highlights in science, engineering and technology*. 2023. Т. 34. С. 275–280. URL: https://doi.org/10.54097/hset.v34i.5484
5. Mukhamediev R. I. State-of-the-Art Results with the Fashion-MNIST Dataset. *Mathematics*. 2024. Т. 12, № 20. С. 3174. URL: https://doi.org/10.3390/math12203174
6. Improved residual networks for image and video recognition / I. C. Duta та ін. *2020 25th international conference on pattern recognition (ICPR)*, м. Milan, Italy, 10–15 січ. 2021 р. 2021. URL: https://doi.org/10.1109/icpr48806.2021.9412193
7. An intelligent fashion object classification using CNN / D. Swain та ін. *EAI endorsed transactions on industrial networks and intelligent systems*. 2023. Т. 10, № 4. С. e2. URL: https://doi.org/10.4108/eetinis.v10i4.4315
8. Understanding robustness of transformers for image classification / S. Bhojanapalli та ін. *2021 IEEE/CVF international conference on computer vision (ICCV)*, м. Montreal, QC, Canada, 10–17 жовт. 2021 р. 2021. URL:https://doi.org/10.1109/iccv48922.2021.01007
9. A comprehensive survey of convolutions in deep learning: applications, challenges, and future trends / A. Younesi та ін. *IEEE access*. 2024. С. 1. URL: https://doi.org/10.1109/access.2024.3376441
10. Shermin T. Enhancing deep transfer learning for image classification : thesis. 2021. URL: http://researchonline.federation.edu.au/vital/access/HandleResolver/1959.17/179551

Received (Надійшла) 11.01.2026
Accepted for publication (Прийнята до друку) 15.04.2026
Publication date (Дата публікації) 22.05.2026

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Соколов Олександр Олександрович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Oleksandr Sokolov – Student, Department of Electronic Computing Machines, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine.
e-mail: oleksandr.sokolov@nure.ua; ORCID Author ID: https://orcid.org/0009-0005-5648-9663.

Порошенко Антон Ігорович – доктор філософії, старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Anton Poroshenko – PhD, Senior Lecturer, Department of Electronic Computing Machines, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: anton.poroshenko@nure.ua; ORCID Author ID: https://orcid.org/0000-0001-7266-4269;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57250025600.

Ярошевич Роман Олександрович – доктор філософії, старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Roman Yaroshevych – PhD, Senior Lecturer, Department of Electronic Computing Machines, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: roman.yaroshevych@nure.ua; ORCID Author ID: https://orcid.org/0000-0002-7949-1513;
Scopus Author ID https://www.scopus.com/authid/detail.uri?authorId=58624172500.

Холєв Владислав Олександрович – доктор філософії, асистент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Vladyslav Kholiev – PhD, Assistant Professor, Department of Electronic Computing Machines, National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: vladyslav.kholiev@nure.ua; ORCID Author ID: https://orcid.org/0000-0002-9148-1561;
Scopus Author ID https://www.scopus.com/authid/detail.uri?authorId=57224189723.

Застосування та аналіз методів машинного навчання для класифікації зображень

О. О. Соколов, А. І. Порошенко, Р. О. Ярошевич, В. О. Холєв

Анотація. Актуальність. Класифікація зображень є ключовим завданням у комп'ютерному зорі, яке має широке застосування в медицині, транспорті, промисловості та безпеці. Використання оптимізованих архітектур CNN дозволяє досягти високої точності з обмеженими ресурсами, що є актуальним для мобільних та вбудованих систем. **Тема дослідження:** методи машинного навчання та архітектури нейронних мереж для класифікації зображень. **Мета статті**. розробити та оцінити модифіковану конволюційну нейронну мережу, яка забезпечує баланс між точністю класифікації та обчислювальною ефективністю, а також порівняти її результати з класичними та сучасними моделями. **Результати дослідження.** Запропонована CNN досягла точності 93,8% на наборі даних Fashion MNIST, перевищивши продуктивність LeNet-5 (91,2%) і класичних алгоритмів (KNN – 85,3%, Decision Tree – 82,7%, XGBoost – 90,4%). На більш складному наборі даних CIFAR-10 модель показала точність 80,2%, перевищивши LeNet-5, але поступившись ResNet і EfficientNet. Це підтверджує ефективність моделі для завдань середньої складності та систем з обмеженими ресурсами. **Висновки.** Модифікована CNN є компромісом між простими класичними методами та складними сучасними архітектурами. Вона забезпечує оптимальний баланс між точністю та швидкістю навчання, що робить її придатною для практичного застосування в мобільних та вбудованих системах. Подальші дослідження можуть зосередитися на використанні більш складних наборів даних, автоматичному виборі гіперпараметрів та інтеграції механізмів самоуваги. Сфера застосування отриманих результатів: системи комп'ютерного зору середньої складності, мобільні та вбудовані пристрої з обмеженими ресурсами, прикладні завдання класифікації зображень.

Ключові слова: класифікація зображень, CNN, Vision Transformer, Fashion MNIST, нейронні мережі, машинне навчання.