

DOI: <https://doi.org/10.60797/IRJ.2024.144.179>**ENHANCING IMAGE PROCESSING WITH A WAVELET-BASED CONVOLUTIONAL NEURAL NETWORK MODEL**

Research article

Liyanage P.^{1,*}, Kruglova L.V.², Ranasinghe N.³¹ORCID : 0009-0005-5030-3266;²ORCID : 0000-0002-8824-1241;³ORCID : 0009-0008-1193-4681;^{1,2,3} Peoples' Friendship University of Russia, Moscow, Russian Federation

* Corresponding author (1032225219[at]pfur.ru)

Abstract

Image processing has emerged as a pivotal technology in the fast advancing domains of computer science and technology. It allows for the extraction of useful information, increase of visual quality, and improvement of decision-making processes. This paper investigates the novel combination of wavelet transformation and convolutional neural networks (CNNs) to tackle different image processing difficulties. The wavelet transformation is well-known for its capability to handle non-stationary signals and perform multi-resolution analysis. It is especially useful for applications that involve extracting detailed features. The objective of the wavelet-based CNN model is to improve the precision and effectiveness of image analysis, denoising, compression, and feature extraction.

This work introduces a holistic methodology that merges the advantages of wavelet processing with CNNs. We analyze the fundamental principles of both techniques and provide a comprehensive explanation of the procedure for building the wavelet-based CNN model. The model underwent training using the Common Objects in Context (COCO) dataset, with emphasis on a selection of the most discernible categories. The results indicate substantial enhancements in both the accuracy and efficiency of picture categorization when utilizing wavelet processing in comparison to conventional CNN models.

The results emphasize the capability of wavelet-based convolutional neural networks (CNNs) to enhance image processing by utilizing the combined advantages of wavelet transformation and deep learning. This work establishes a basis for future investigation and advancement in constructing more resilient and effective models for diverse image processing applications, hence facilitating the emergence of inventive solutions in computer vision.

Keywords: convolutional neural networks, deep learning, feature extraction, image processing, wavelet transformation.**УЛУЧШЕНИЕ ОБРАБОТКИ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ МОДЕЛИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ НА ОСНОВЕ ВЕЙВЛЕТОВ**

Научная статья

Лиянаге П.^{1,*}, Круглова Л.В.², Ранасингхе Н.³¹ORCID : 0009-0005-5030-3266;²ORCID : 0000-0002-8824-1241;³ORCID : 0009-0008-1193-4681;^{1,2,3} Российский университет дружбы народов, Москва, Российская Федерация

* Корреспондирующий автор (1032225219[at]pfur.ru)

Аннотация

Обработка изображений стала одной из важнейших технологий в быстро развивающихся областях компьютерной науки и техники. Она позволяет извлекать полезную информацию, повышать качество изображения и улучшать процессы принятия решений. В данной статье исследуется новое сочетание вейвлет-преобразования и сверточных нейронных сетей (CNN) для решения различных задач обработки изображений. Вейвлет-преобразование хорошо известно своей способностью обрабатывать нестационарные сигналы и выполнять анализ в нескольких разрешениях. Оно особенно полезно для приложений, связанных с извлечением детальных характеристик. Цель CNN-модели на основе вейвлет-преобразования – повысить точность и эффективность анализа изображений, их обесцвечивания, сжатия и извлечения особенностей.

В данной работе представлена целостная методология, объединяющая преимущества вейвлет-обработки и CNN. Мы анализируем фундаментальные принципы обеих методик и даем исчерпывающее объяснение процедуры построения CNN-модели на основе вейвлетов. Модель прошла обучение на наборе данных Common Objects in Context (COCO) с акцентом на выборку наиболее различных категорий. Полученные результаты свидетельствуют о значительном повышении точности и эффективности категоризации изображений при использовании вейвлет-обработки по сравнению с обычными моделями CNN.

Полученные результаты подчеркивают способность сверточных нейронных сетей (CNN) на основе вейвлетов улучшать обработку изображений за счет использования комбинированных преимуществ вейвлет-преобразования и глубокого обучения. Данная работа закладывает основу для будущих исследований и продвижения в построении более устойчивых и эффективных моделей для различных приложений обработки изображений, что способствует появлению инновационных решений в области компьютерного зрения.

Ключевые слова: сверточные нейронные сети, глубокое обучение, извлечение признаков, обработка изображений, вейвлет-преобразование.

Introduction

Image processing has become a powerful and influential factor in several fields within the fast-advancing field of computer science and technology. This study focuses on manipulating visual inputs to extract significant data, improve the quality of perception, and enable efficient decision-making in many contexts [1]. Wavelet transformation has gained considerable interest in image processing due to its distinctive characteristics.

The distinguishing feature of wavelet transformation is its capacity to process non-stationary signals and effectively represent data that possesses both global and local characteristics. This adaptability makes it especially well-suited for applications that need thorough feature extraction together with analysis at many resolutions [2]. With the ongoing progress in image processing, the integration of wavelet transformation with neural networks presents exciting opportunities for innovation.

This paper presents a novel approach that integrates wavelet transformation with convolutional neural networks (CNNs) to improve image processing tasks. This model seeks to enhance accuracy and efficiency in image analysis, denoising, compression, and feature extraction by utilizing the advantages of both approaches.

The subsequent sections will offer a thorough examination of the theoretical underpinnings of wavelet transformation along with CNNs, delineate the approach employed to construct the wavelet-based CNN model, followed by showcasing the results of the study. This study emphasizes the possibility of combining wavelets with CNNs, demonstrating their efficacy in expanding the domain of image processing.

Literature review

2.1. Image Processing Fundamentals

Image processing is a major field of computer science that manages and analyses visual inputs to extract useful information. It aims to improve image quality, pattern detection, as well as application decision-making [3]. Image processing has advanced due to improvements in visual data collection and analysis.

Image processing includes two major objectives which are image restoration and enhancement. Image Restoration includes restoring images which have been damaged by noise, blur, and other factors [1]. Common methods include deblurring and denoising. Image enhancement improves the look and suitability of an image for various activities. Colour correction, contrast enhancement, and edge enhancement are common.

Scaling, cropping, filtering, segmentation, and transformation are used to modify visual data. The advancement of computer vision, remote sensing, and medical imaging depends on feature extraction, pattern recognition, and object identification [4].

Image processing has grown into a vibrant, multidisciplinary area throughout the years. In the mid-20th century, histogram equalization and contrast stretching were employed to improve images. These approaches enhanced image quality for interpretation and analysis [5]. In the 1980s and 1990s, complex mathematical methods and algorithms for picture segmentation and edge detection were introduced. FFT and morphological processing became popular [6].

Digital images and cameras pushed more improvements in the 2000s. Wavelet-based compression (e.g., JPEG2000) and more sophisticated image processing algorithms were developed during this time [7]. AI and ML transformed image processing in the 2010s and beyond. In object identification, image classification, and facial recognition, CNNs and deep learning achieved unparalleled accuracy and efficiency. These advances made image processing a flexible tool for large data and complicated visual information difficulties [8].

Image processing evolves due to computer power, algorithmic innovation, and big datasets. Wavelet transformation using neural network topologies is a recent breakthrough that promises improved image processing.

2.2. Fundamentals of Wavelet Transformation

Wavelet transformation divides an image into frequency components for multiscale analysis. Hierarchical decomposition divides the image into approximation and detail coefficients [9].

Approximation Coefficients: These capture the low-frequency form and structure of the image.

Detail Coefficients: These capture the high-frequency components such as edges, textures, and minute details.

The wavelet transform can be continuous or discrete. The computationally efficient Discrete Wavelet Transform (DWT) is employed more in real applications [10].

$$A_k[n] = \sum x[n] \cdot \phi_k[n - 2^k] \quad (1)$$

$$D_k[n] = \sum x[n] \cdot \psi_k[n - 2^k] \quad (2)$$

where A_k and D_k are the approximation and detail coefficients at level k , respectively, and ϕ and ψ are the scaling and wavelet functions.

The continuous wavelet transform (CWT) [11] of a signal $x(t)$ is given by:

$$CWT_s(\tau) = \int_{-\infty}^{\infty} x(t) \cdot \frac{1}{\sqrt{|s|}} \psi^* \left(\frac{t-\tau}{s} \right) dt \quad (3)$$

where s is the scaling factor, τ is the translation parameter, and ψ is the wavelet function.

2.3. Image Processing using Wavelet Transformation

Image-compression

Wavelet transformation reduces data redundancy, making it an excellent image compression method [12]. Wavelet-based approaches can quantize and encode coefficients by dividing an image into frequency bands, emphasizing perceptually relevant components and eliminating less important features.

Denoising

Wavelet-based Image noise reduction is vital, and wavelet transformation excels at it [13]. Wavelet-based denoising algorithms may locate and reduce noise in the high-frequency detail coefficients while keeping significant image properties by splitting an image into approximation and detail coefficients.

Extraction of Features

Wavelets extract features at different resolutions, capturing minute details and larger patterns [14]. Object identification, texture analysis, and image retrieval need multi-scale feature recognition, making this capability beneficial.

Detecting and Enhancing Edges

Wavelets can examine images at different scales for edge identification [15]. Wavelets may detect abrupt intensity fluctuations due to high-frequency detail coefficients highlighting edges and transitions. Wavelet coefficients can enhance features or suppress artifacts, improving image quality.

Wavelet transformation combines information from several images, such as multi-focus or multi-sensor images, into a single image. This method decomposes each image into wavelet coefficients, merges them according to criteria, and reconstructs the final image.

CNN integration

Image processing utilizing wavelet transformation integrated with CNNs is a cutting-edge concept which is being developed [16]. CNN model can use wavelet transformations to combine the strengths of both methods.

Wavelet transformations can pre-process images to give CNNs multi-resolution representations of detailed and contextual information [17]. CNNs can benefit from wavelet-based denoising to lower noise sensitivity and improve model performance. Wavelets minimize input data dimensionality while keeping critical properties, making CNNs more efficient and effective. This integration produces powerful models that can accurately classify, segment, and improve images, demonstrating wavelet-based CNNs' promise to advance image processing.

Methodology

The methodology section of the article provides a comprehensive explanation of the methods employed in this study.

Selection of Dataset

The first step in building the solution was carefully selecting an appropriate dataset to use as the foundation for the research effort. After a thorough assessment of several options, the Common Objects in Context (COCO) dataset [17] created by Microsoft was chosen as the best suitable choice for this research. The COCO dataset is renowned for its extensive collection of high-quality photos that have been meticulously annotated with data for object recognition, segmentation, and captioning. The COCO dataset encompasses a wide range of item categories and scene types, rendering it a highly valuable asset for computer vision applications.

The COCO dataset is distinguished by its extensive size, comprising more than 300,000 pictures that encompass 80 categories. Once the COCO dataset was chosen, the subsequent step involved preparing the data for both model training as well as evaluation. The dataset was thereafter partitioned into many groups based on the defined item categories using the provided JSON files. This categorization approach facilitated efficient data management and organization, ensuring that images belonging to the same category were grouped together for subsequent processing.

Creating a database

The division of these photos resulted in the creation of a database including 80 categories based on the annotation. Due to the arduous nature of training a model with 80 classes, it was necessary to decrease the number of classes. After careful consideration, the top 10 image categories were chosen based on their unique characteristics, with each category including between 1000 to 2000 images. The purpose of this selection strategy was to create a balanced combination of preserving diversity within the dataset and ensuring an adequate representation of each class throughout the model training.

In this essential stage, the task is to traverse the directory structure of the generated image database and gather all available image files, together with their corresponding class labels. The COCO dataset is renowned for its comprehensive collection of images that depict a diverse array of objects in different situations. The script systematically traverses the directory structure, recognizing each image file and extracting its matching class label based on the directory it is found in. This technique ensures that each image is correctly matched with its corresponding class, laying the groundwork for subsequent training and validation steps. More precisely, the script iterates over every class directory in the database and documents the file paths of each image, along with their corresponding class labels. This meticulous extraction technique ensures that no image is overlooked, and that every category in the database is accounted for. Moreover, the extensive size and diverse content of the COCO dataset present both benefits and challenges. The extensive coverage of the dataset enables comprehensive training of deep learning models, but its substantial size necessitates meticulous control to avoid resource constraints and processing delays. Therefore, the script employs efficient methods to expedite the data loading process, ensuring that all images are successfully processed and incorporated into the training and validation sets.

Image preprocessing

Subsequently, the validation images underwent resizing. The intended picture size dictates the ratios by which the input photos will be reduced before being fed into the model. The picture dimensions for this investigation were configured as 224×224 . This ensures uniformity in input dimensions, allowing this model to process pictures of consistent sizes. Moreover, the selection of a batch size allows for the regulation of the number of data processed in each iteration, hence enhancing memory usage and computational efficiency.

The system incorporated the wavelet transformation as an innovative concept. Wavelet transformation is a very efficient method employed in signal processing to decompose data into different frequency components, allowing for the analysis of

both spatial as well as frequency information. The utilization of wavelet transformation in image processing offers significant benefits, as it allows for the accurate representation of complex features and textures at various scales.

The model employs a wavelet transformation function that performs a two-dimensional discrete wavelet transform (DWT) on each input image, using the Haar wavelet [18]. The Haar wavelet, renowned for its simplicity, is very proficient at detecting and portraying abrupt changes or discontinuities in signals or images.

In the context of a two-dimensional discrete wavelet transform (DWT) for images, the wavelet function ψ_k is utilized in both the horizontal and vertical orientations. This program divides down the image into approximation (LL) and detail (LH, HL, HH) coefficients. The coefficients correspond to the multi-resolution attributes of the image, encompassing its overall structure (LL) as well as fine details in various orientations (LH, HL, HH).

Building and training the model

The CNN is a powerful structure specifically designed to handle organized matrix data, such as images [20]. In the field of image classification, the CNN undergoes a sequence of modifications, each tailored to extract and get increasingly intricate and abstract representations of the input data.

A customised CNN was built using appropriate optimization and initialising step which has three convolutional together with three pooling layers, a flatten layer along with four fully connected layers.

The three convolutional layers were created using 64, 32, and 16 filters respectively and were constructed with kernel initialization to simplify the training process by assisting in the initialization of the weights.

The rectified linear unit (ReLU) activation function [21] was then employed on all hidden layers. A pooling layer serves as a means to minimize the dimensionality of the data. In this case, three pooling layers were used to lower the output size from each of the convolutional layers by half. A flatten layer was employed to convert the matrix from the final pooling layer into a one-dimensional (1-D) column vector for the purpose of classification. Four completely linked layers were implemented in this research. To mitigate overfitting and regularize the output, three activation functions and two dropout layers were used. The last fully connected layer was constructed with the 'Softmax' activation function to convert the data into probabilities.

Furthermore, the model was trained using a learning rate of 0.001, a batch size of 32, with 50 epochs and L2 regularization.

Results and discussion

This section provides the findings of this study for training a custom wavelet based CNN.

Table 1 - Performance Metrics Value

DOI: <https://doi.org/10.60797/IRJ.2024.144.179.1>

Accuracy, %	97.8
Validation Accuracy, %	91.5
Loss, %	8.7
Validation Loss, %	10.7

The ability of the model to accurately categorize the training images is demonstrated by the attained accuracy of 97.8% on the training data. The high accuracy of the model showcases its proficiency in comprehending intricate patterns and discerning crucial aspects within the training dataset. However, when assessing the model's performance on unseen data, as evidenced by the validation accuracy of 91.5%, a little decrease relative to the training accuracy is seen as illustrated in figure 1. The discrepancy in accuracies between the training and validation sets and the spikes seen in figure 1, indicates that the model is experiencing overfitting, wherein it has effectively remembered the training data rather than acquiring the ability to generalize its knowledge to novel instances. Despite the relatively low validation accuracy, the model has demonstrated strong performance overall, in line with expectations.



Figure 1 - Accuracy and Validation Accuracy functions against the number of epochs

DOI: <https://doi.org/10.60797/IRJ.2024.144.179.2>

Analysing the loss numbers provides a deeper comprehension of the model's performance during the training procedure. The training loss of 8.7% and validation loss of 10.7% indicate the extent of disparity or divergence between the model's predictions and the actual labels. A lower loss number indicates a more robust correlation between the predicted and real labels, implying that the model effectively minimized the errors it generated during the training phase, as shown in figure 2. However, the somewhat higher validation loss compared to the training loss suggests that the model struggled to apply its expertise to new data, maybe due to overfitting or insufficient regularization.

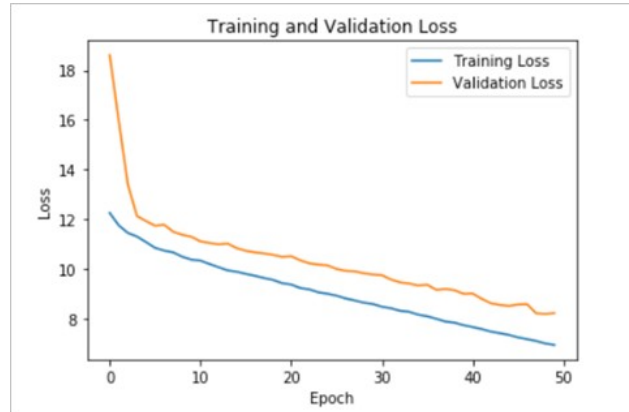


Figure 2 - Loss and Validation Loss functions against the number of epochs
 DOI: <https://doi.org/10.60797/IRJ.2024.144.179.3>

Comparing with base CNN

The performance matrices obtained during testing of this CNN model against a comparable CNN model that does not include wavelet processing are as follows:

Table 2 - Model Accuracy achieved

DOI: <https://doi.org/10.60797/IRJ.2024.144.179.4>

CNN model with wavelet transformation, %	91.5
Base CNN model, %	75.8

The comparison between the CNN model using wavelet transformation and the CNN model without wavelets underscores the significant impact of integrating wavelet transformation on model performance. The CNN model using wavelet modification demonstrated a validation accuracy of 91.5%, representing a substantial improvement above the 75.8% validation accuracy attained by the CNN model without wavelets. As shown in figure 3, although the shapes of the accuracy and loss functions from the base CNN appear to be the same as the wavelet integrated CNN model, the values at which the graphs plateau are widely different between the two models.

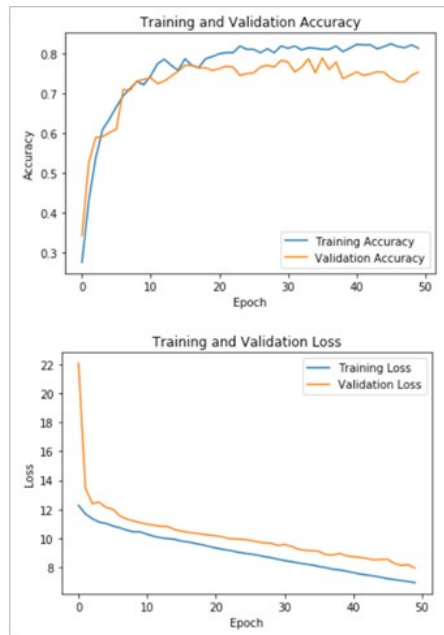


Figure 3 - The results obtained from the base CNN
DOI: <https://doi.org/10.60797/IRJ.2024.144.179.5>

The substantial improvement in validation accuracy achieved with the integration of wavelet transformation highlights the efficacy of this approach in addressing key challenges associated with image classification tasks. The utilization of wavelet transformation boosts the representation of features and promotes the understanding of visual structures, hence enabling the CNN model to provide predictions that are more accurate and dependable.

Conclusion

The finished work primarily concentrated on the development of a deep learning model that employs wavelet transformation for the purpose of image processing. Image processing is a significant field in computer science and technology. It is having a transformative impact on various sectors. Wavelet transformation is a highly effective tool in image processing due to its distinctive capability to provide a multi-resolution analysis. It can simultaneously capture both high and low-frequency components.

Thus, it can be concluded that the integration of wavelet transformation into neural networks has been a triumph. The efficacy of wavelet transformation in improving model performance underscores the necessity of exploring innovative preprocessing tactics and employing advanced signal processing techniques in deep learning applications. This integration has resulted in more accurate and dependable forecasts, thereby addressing the primary challenges associated with picture categorization.

The model was just employed to execute a classification function. This wavelet-based CNN has the capability to be improved in order to conduct a segmentation procedure, as well as do real-time object identification tasks. In order to include these characteristics, it will be necessary to make adjustments to both the structure of the model and the input data.

In addition, future study might explore additional preprocessing methods and assess their effectiveness when combined with wavelet transformation to enhance the model's performance. Moreover, the information obtained from this research may be utilized to improve the development of highly efficient and reliable deep learning models for a wide range of image processing applications. This will ultimately advance the field of computer vision and enable the creation of innovative solutions in several domains.

Конфликт интересов

Не указан.

Рецензия

Сообщество рецензентов Международного научно-исследовательского журнала
DOI: <https://doi.org/10.60797/IRJ.2024.144.179.6>

Conflict of Interest

None declared.

Review

International Research Journal Reviewers Community
DOI: <https://doi.org/10.60797/IRJ.2024.144.179.6>

Список литературы на английском языке / References in English

- Chitradevi B. An Overview on Image Processing Techniques / B. Chitradevi, P. Srimathi, A. Professor // International Journal of Innovative Research in Computer and Communication Engineering. — 2007. — 11.
- Kingsbury N. Wavelet Transforms in Image Processing / N. Kingsbury, J. Magarey. — 1998. — p. 27-46. DOI: [10.1007/978-1-4612-1768-8_2](https://doi.org/10.1007/978-1-4612-1768-8_2).

3. Kumar K. A Review on Image Processing / K. Kumar, S. Anand, D. Soni, N. Gaur. — 2013. — №7. — URL: <https://www.iiste.org/> (accessed: 08.04.2024)
4. Dastres R. Advanced Image Processing Systems. / R. Dastres, M. Soori. — 2021 — URL: <https://www.ceserp.com/cp-jour> (accessed: 08.04.2024)
5. Billingsley F.C. Digital image processing for information extraction / F.C. Billingsley // Int J Man Mach Stud. — 1973. — 5. — p. 203-214. DOI: 10.1016/S0020-7373(73)80032-X.
6. Schreiber W.F. Image Processing for Quality Improvement / W.F. Schreiber // Proceedings of the IEEE. — 1978. — 66. — p. 1640-1651. DOI: 10.1109/PROC.1978.11172.
7. Chen Y. Introduction to Wavelet Analysis with Applications to Image and JPEG 2000 / Y. Chen // ACM International Conference Proceeding Series. — 2022. — p. 49-57. DOI: 10.1145/3524086.3524094.
8. Kuruvilla J. A review on image processing and image segmentation / J. Kuruvilla, D. Sukumaran, A. Sankar, S.P. Joy // 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE). — 2016. — p. 198-203. DOI: 10.1109/SAPIENCE.2016.7684170.
9. Saha M. Advanced Wavelet Transform for Image Processing – A Survey / M. Saha, M.K. Naskar, B.N. Chatterji // Lecture Notes in Networks and Systems. — 2020. — 79. — p. 185-194. DOI: 10.1007/978-981-32-9453-0_19/COVER.
10. Zhao W. Discrete-time continuous-dilation wavelet transforms / W. Zhao, R.M. Rao // Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis. — 1998. — p. 233-236. DOI: 10.1109/TFSA.1998.721404.
11. Gao R.X. Continuous Wavelet Transform / R.X. Gao, R. Yan. — 2011 — URL: https://link.springer.com/chapter/10.1007/978-1-4419-1545-0_3 (accessed: 10.04.2024) DOI: 10.1007/978-1-4419-1545-0_3.
12. Song M.S. Wavelet Image Compression / M.S. Song. — 2006 — URL: https://www.researchgate.net/publication/266018963_Wavelet_image_compression (accessed: 10.04.2024)
13. Agarwal S. Analysis and comparison of wavelet transforms for denoising MRI image / S. Agarwal, O.P. Singh, D. Nagaria // Biomedical and Pharmacology Journal. — 2017. — 10. — p. 831-836. DOI: 10.13005/BPJ/1174.
14. Nafoarniță C. A second order statistical analysis of the 2D discrete wavelet transform / C. Nafoarniță, I. Firoju, A. Isar, D. Isar, J.M. Boucher // 2010 8th International Conference on Communications, COMM 2010. — 2010. — p. 145-148. DOI: 10.1109/ICCOMM.2010.5509080.
15. Ruo-fei C. Wavelet transformation and its application in image processing / C. Ruo-fei // Journal of Hebei Institute of Technology. — 2003.
16. Liyanage P. An in-depth look into combining image processing and wavelet transformation / P. Liyanage, L.V. Kruglova // International Research Journal. — 2024. — 3(141). DOI: 10.23670/IRJ.2024.141.59.
17. Kaszta T. Wavelet-Based Prototype Networks / T. Kaszta, H.G. Lieb, L. Csato // SISY 2023 - IEEE 21st International Symposium on Intelligent Systems and Informatics, Proceedings. — 2023. — p. 277-282. DOI: 10.1109/SISY60376.2023.10417903.
18. COCO - Common Objects in Context. — 2014 — URL: <https://cocodataset.org/#home> (accessed: 11.04.2024)
19. Kanagaraj H. Image compression using HAAR discrete wavelet transform / H. Kanagaraj, V. Muneeswaran // IEEE International Conference on Distributed Computing Systems. — 2020. — p. 271-274. DOI: 10.1109/ICDCS48716.2020.243596.
20. Saha S. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way / S. Saha. — 2018 — URL: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (accessed: 12.04.2024)
21. Baheti P. Activation Functions in Neural Networks [12 Types & Use Cases] / P. Baheti. — 2021 — URL: <https://www.v7labs.com/blog/neural-networks-activation-functions> (accessed: 12.04.2024)