

DOI: 10.31319/2519-8106.1(50)2024.304790

UDC 004.9

**Yalova Kateryna**, Candidate of Technical Sciences, Associate Professor, Department of the Systems software

**Ялова К.М.**, кандидат технічних наук, доцент, кафедра програмного забезпечення систем

ORCID: 0000-0002-2687-5863

e-mail: yalovakateryna@gmail.com

**Ahrymeev Oleksandr**, master's degree student, Department of the Systems software

**Ахримєєв О.О.**, здобувач другого (магістерського) рівня вищої освіти, кафедри програмного забезпечення систем

email: achrin@ukr.net

Dniprovsky State Technical University, Kamianske

Дніпровський державний технічний університет, м. Кам'янське

## HANDWRITTEN DIGITS RECOGNITION WITH A CONVOLUTIONAL NEURAL NETWORK

## РОЗПІЗНАВАННЯ РУКОПИСНИХ ЦИФР ЗАСОБАМИ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

*The patterns recognition problem's formulation using the handwritten digit recognition task as an example is presented in the paper. To solve the problem, the feasibility of applying a convolutional neural network is justified. The architecture of the developed convolutional network is presented and its training and testing parameters are described. The database of handwritten digits MNIST is used for training the neural network.*

*The results of software development within an integrated development environment Visual Studio.net using C# programming language are presented. The output of the developed application is a response to the user regarding the results of handwritten digit recognition submitted to the application in the form of a graphic image. The causes of recognition errors have been analyzed, an assumption regarding the optimal setting of the developed neural network architecture has been made. The approach of convolutional neural network topology construction, which is effective for handwritten digit recognition, is proposed. The validity of the proposed solutions effectiveness lies in the results of comparing the obtained data with the data of other authors. Recognition accuracy for the developed convolutional network is 96.83 %.*

**Keywords:** *patterns recognition, handwritten digit recognition, convolutional neural network.*

*У статті здійснено постановку завдання розпізнавання образів на прикладі розпізнавання рукописних цифр. Для розв'язання поставленого завдання обґрунтовано доцільність застосування згорткової нейронної мережі. Представлено архітектуру застосованої згорткової нейронної мережі і описано параметри її навчання і тестування. Для проведення навчання нейронної мережі застосовано корпус рукописних цифр MNIST.*

*Наведено результати створення програмного забезпечення, що було розроблено в рамках інтегрованого середовища розробки програмного забезпечення Visual Studio.net засобами мови програмування С#. Результатом роботи розробленого застосунку є відповідь користувачу щодо результатів розпізнавання рукописних цифр, поданих до застосунку у вигляді графічних зображень. Проаналізовано причини помилок розпізнавання, зроблено припущення щодо оптимізації архітектури розробленої нейронної мережі. Запропоновано підхід до побудови топології згорткової нейронної мережі, яка є ефективною для розпізнавання рукописних цифр, здійснено налаштування нейронної мережі для підвищення якості розпізнавання. Обґрунтова-*

ність ефективності запропонованих рішень полягає в результатах порівнянні отриманих даних із даними інших авторів. Точність розпізнавання для розробленої згорткової мережі 96,83 %.

**Ключові слова:** розпізнавання образів, розпізнавання рукописних цифр, згорткова нейронна мережа.

### Problem's Formulation

Handwritten digit recognition is subject of study in the pattern recognition and classification problem. As a direction of science development, the theory of pattern recognition was created in the 50s of the twentieth century. The task of pattern recognition is to find elements that belong to a certain class, within a set of objects that can or belong to different classes, where class is a merged set that is formed using specified rules according to the general properties inherent in each object of that class. The pattern is a representation of the input element as an instance of the corresponding pattern class [1].

The basis of the pattern recognition theory is the scientific task, namely — to develop a mechanism or software machine that would be used to classify events, situations, images, various objects. Setting a scientific problem made it possible to form a number of scientific tasks in the theory of pattern recognition, for example [2]:

- development of models to describe the process of perception and classification of the inputted data;
- algorithms, methods and models construction for pattern recognition within specific data domains;
- find ways to automate the recognition process.

In general, the pattern recognition task consists of two parts:

- training;
- recognition.

Training process is carried out by demonstrating individual objects with their belonging to one or another pattern. As a result of training, the recognition system must acquire the ability to respond equally to all objects of the same class and in different ways to all objects of different classes. It is very important that the learning process should end only by showing the finite number of objects without any other hints.

Objects of training can be either pictures, or other visual images (letters, digits), or various phenomena of the outside world. It is important that during the training process only the objects themselves and their belonging to the pattern are specified. The training is followed by the process of recognizing new objects, which characterizes the actions of an already trained system. Automation of these procedures poses a problem of learning to recognize pattern. Images of any object can be used as an input for the pattern recognition task [3]. In the presented paper as input data the handwritten digit images are applied for the recognition. As the image of the digit can be treated as a pattern class instance, so it's possible to claim that the task of handwritten digit recognition is one form the numerous tasks of the pattern recognition problem. Also the handwritten digit recognition can be specified as a supervised machine learning task.

For banks, industrial and trading enterprises, the problem of paper record-keeping inefficiency is becoming more urgent. To move to automated electronic document flow using unified and standardized processes based on paperless technologies, it is necessary to solve the problem of handwritten text information recognition for translating printed and handwritten documents into electronic form. Hand-writing recognition systems are used in a variety data domain, such as reading and archiving old documents, bank checks, and letters. In addition to these examples, handwriting recognition is widely used in the field of education, and on some electronic devices, in particular tablets there is a built-in software to recognize handwriting text, turning handwriting text into printed.

### Analysis of recent research and publications

The rapid development of computer and information technologies and improving the performance of computing algorithms have become the basis for the creation of new and optimization of developed methods and algorithms for pattern recognition. In recent years, the development of image processing and image recognition methods, methods of computer vision and artificial intelligence for

individual application tasks has made it possible to create optical character recognition systems (OCR systems). The development of software for solving practical problems in this data domain is still an up-to-date scientific and practical problem that requires the further development of theoretical knowledge and the search for solutions to increase speed, accuracy, completeness of the recognition and efficiency of the recognition process, and reduction of memory used by the recognition process.

The first work on handwritten digit recognition appeared in 1980—90, in which Shildhar and Badreldin proposed an algorithm that could be used to recognize handwritten digits, and in 1989 Yann LeCun and others proposed the most basic LeNet convolutional network for pattern recognition, and already in 1990 he proposed to use backpropagation method for neural network models [4]. Every year, the structure of neural networks becomes more and more effective, if the first convolutional neural network that was proposed by Jan LeCun in 1998, gave an error of 1.1 %, then in 2012 Dan Krizan described in his article the Ciresan network, which managed to show the result of the error in 0,23 %, which indicates that the problem of pattern recognition is relevant and requires further development.

Approaches by various scientists have been proposed to extract character recognition features. The main way to reduce the error is to change the topology of the developed neural networks, search for variations with the number and type of layers, or activation function, even without preprocessing of the inputted images. In search of optimal solutions, scientists conduct comparative analyses of various qualification, template and neural network algorithms on a variety of datasets and databases of handwritten digits, in particular Performance and Modified National Institute of Standards and Technology (MNIST). Multivariate studies substantively prove the feasibility of using convolutional neural networks during the decision of handwritten digit recognition.

Handwritten digits can be recognized by a variety of methods, and their precision, recall, productivity capacity differ from each other. Tabl. 1 shows the comparison characteristics of handwritten digit recognition's results obtained by different scientist using different recognition methods.

Table 1. Comparison characteristics of handwritten digit recognition's results

Scientists	Algorithm or neural network type	Dataset	Precision, %
Yang Gong, Pan Zhang [5]	CNN	MNIST	98-99,9
Rabia Karakaya, Serap Kazan [6]	Support vector machine (SVM)	MNIST	84-90
	Decision tree		85
	Random forest		97
	K-nearest neighbor		96
Ritik Dixit, and other [7]	SVM	MNIST	94
	MLP		98
	CNN		99
Syed Sohail Ahmed, and other [8]	BiFPN	MNIST	97-99
Nidhal Azawi [9]	CNN	MNIST	97-99
Ming Chen, and other [10]	FOBP	MNIST	93-96
Zufar Kaymov, and other [11]	Combined CNN	MNIST	97-99

Analyzing the data of numerical experiments obtained by other scientists, the conclusions can be as following:

1. The most justified architecture of the neural network used for handwritten digit recognition is convolutional neural networks.
2. It is advisable to use the MNIST dataset for training and testing the developed neural network.
3. The average recognition precision depends significantly on the input image and ranges from 90—99 %.

#### Formulation of the study purpose

The problem of handwritten digit recognition has been widely studied in recent years and a large number of pre-processing methods and classification algorithms have been developed for this purpose. However, it is still an actual scientific and practical task. The main difficulty in recognizing

is a serious variance in size, translation, stroke thickness, rotation and deformation of the digit images, since digits can be written by different people, and their writing style is unique or can differ significantly from each other.

The aim of the article is to present the results of development, validation and testing of a convolutional neural network for solving the problem of handwritten digit recognition. The objectives of the research are: construction of the neural network topology, neural network training on a training dataset, testing the neural network and determining optimal parameters of its using, software development to provide a dialogue with users to get the results of recognition.

### Presenting main material

Intelligent systems based on artificial neural networks can successfully solve problems of pattern recognition, prediction execution, optimization, associative memory and control. The use of a neural network approach to solve the problem of pattern recognition is a popular and highly effective method.

To achieve the study purpose, an information system was developed, consisting of:

- neural network;
- software for setting and updating the main parameters of the neural network to analyze the effectiveness of its use.
- software providing interaction with users for submission an incoming image and providing recognition results.

To solve the handwritten digit recognition task, the convolutional neural network was constructed. A convolutional neural network is a feed-forward artificial neural network with a specific topology. Convolutional network neurons are located not in a row, but in a matrix. Matrices for such neurons are of different sizes, the commonly used sizes are  $3 \times 3$ ,  $5 \times 5$  or  $7 \times 7$ . Such a neuron matrix passes through the entire image in increments of 1 and forms the so-called map of features. Due to the fact that the matrix gives just one output the network is called convolutional. After such a matrix has passed through the entire image, each output passes through the activation function. In Fig. 2, the common architecture of a convolution neural network is presented. Convolutional networks have convolutional, pooling, fully-connected and softmax layers [12].

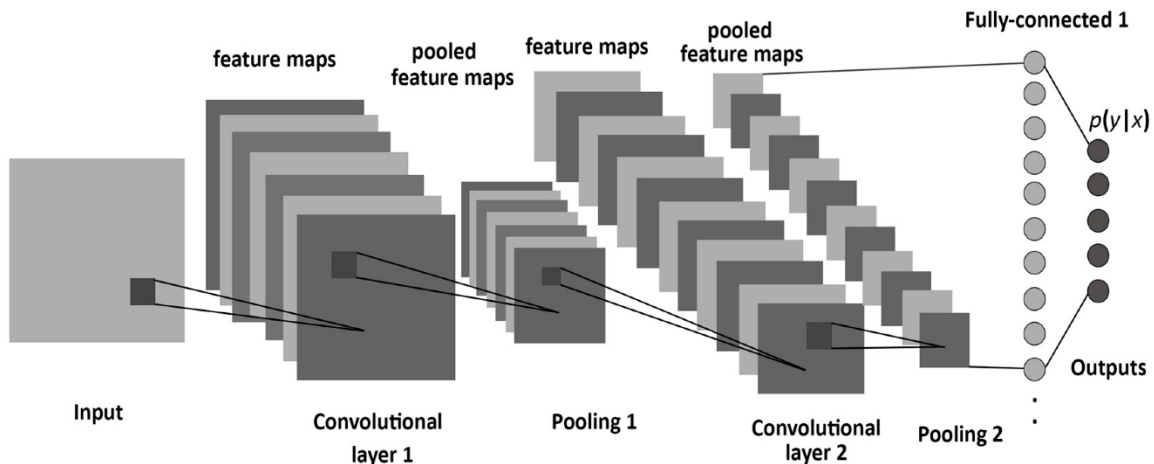


Fig. 1. The common architecture of the convolutional neural network

The convolutional layer commonly used to get the feature information from the inputted image. This mechanism can be described with the formula:

$$Y(x, y) = \sum_{a=1} \sum_{b=1} K(a, b) \cdot X(x+1-a, y+1-b), \quad (1)$$

where  $X$  — is the input matrix,  $K$  — is the kernel matrix,  $Y$  — is an output matrix.

For the inputted image with height  $m_1$  and width  $m_2$  the  $ij_{th}$  entry of the feature map will be as following:

$$f[i, j] = \sum_x^{m_1} \sum_y^{m_2} \sum_z^{m_c} K_{[x,y,z]} I_{[i+x-1, j+y-1, z]}, \quad (2)$$

where  $I, K$  — are the chosen parts of the image,  $m_c$  — number of channels.

The max-pooling layer most often places in the network immediately behind the convolution layer. It processes the feature maps from the convolutional layer. In convolutional neural network where the convolutional level is  $Conv(I, K)=C$  the max-pooling layer can be describe as:

$$P = f_p(C), \quad (3)$$

where  $f_p$  — is a pooling function.

The max-pooling layer does not process data in depth, each neuron of such a layer is connected to a single feature map. Such a neuron usually does not have an activation function, it passes through the feature maps with kernel matrix for example,  $2 \times 2$  with step 2, or  $4 \times 4$  with step 4, and simply selects the maximum value from this kernel matrix, all other values are ignored. On one hand this highlights the strongest features on the network, on the other, it reduces the image. When passing through the  $2 \times 2$  kernel matrix in increments of 2 from the image  $24 \times 24$ , the feature map  $12 \times 12$  will remain.

The fully-connected layer deploys the resulting data into a single vector. We can say that after all layers of convolution and subassembly, a conventional neural network is connected to convolution neural network at the end. Mathematically the fully-connected layer can be described as below:

$$X = \sum_i w_i P + b', \quad (4)$$

where  $P$  — is a max-pooling layer.

The softmax layer is the same as in a conventional neural network. It has a number of neurons that corresponds to the number of outputs. It is connected to a fully-connected layer and is the output layer of the network. The output of such a layer is a vector of values from 0 to 1, it is the response of the network.

The proposed convolutional neural network topology consists of:

- an input layer that represents images of  $28 \times 28$  pixels;
- convolution layer with kernel matrix  $5 \times 5$ , in this layer 8 cores;
- ReLU activation function;
- maxpool layer  $2 \times 2$  with step 2;
- convolution layer with kernel matrix  $5 \times 5$ , in this layer 16 units;
- ReLU activation function;
- maxpool layer  $3 \times 3$  with step 3;
- full-connected layer with neurons;
- softmax layer.

ReLU stands for Rectified Linear Unit and is described as:

$$f(x) = \max(0, x). \quad (5)$$

Despite the fact that the mathematical description gives the impression of a linear function, ReLU has a derived function and allows backpropagation while making it computationally efficient [13]. A training algorithm for a convolutional neural network is a backpropagation method. The image passes through the network, an output is obtained, the output is compared with the labels and the back search of errors begins through all layers and for each neuron.

The input for the application will be images of handwritten digits from the MNIST database [14]. This database is a standard that is widely used for training, testing, calibration, topology changes, optimization of neural networks. The database contains 60,000 training images and 10,000 testing images. All images are  $28 \times 28$  pixels in size. Also, a label with the image name is attached to each image, that is, the label will tell the network what the digit is in the image [15].

Software for implementing the constructed neural network model was developed as a user application in the C # programming language. The main application's main window (Fig. 2). There are 6 main elements displayed on the application's main form:

1. The set of functional buttons to operate with the neural network. Using the corresponding button, it is possible to start training the neural network, to provide the testing of the model or to pause the process.

2. The set of controls to get and set data about the image under the recognition. The Text Box control named “Current” is used to output the information about the current image’s index, it is not editable. The Text Box control named “Next” allows to input the image’s index that will be used in the system next. The Drop-Down List control named “Zoom” is used to enlarge the images that are displayed on the application’s form.

3. The Picture Box control that is used to display the current digit’s image in the initial form.

4. The List Box control outputs the detailed information about the process of training and testing the neural network. User get the information about the Loss function, training and test accuracy, time of a recognition.

5. Chart control is used to visualize the information about dependence of the loss function and the number of epoch.

6. Track Bar control allows to output the information about the progress of the recognition process in the graphical form.

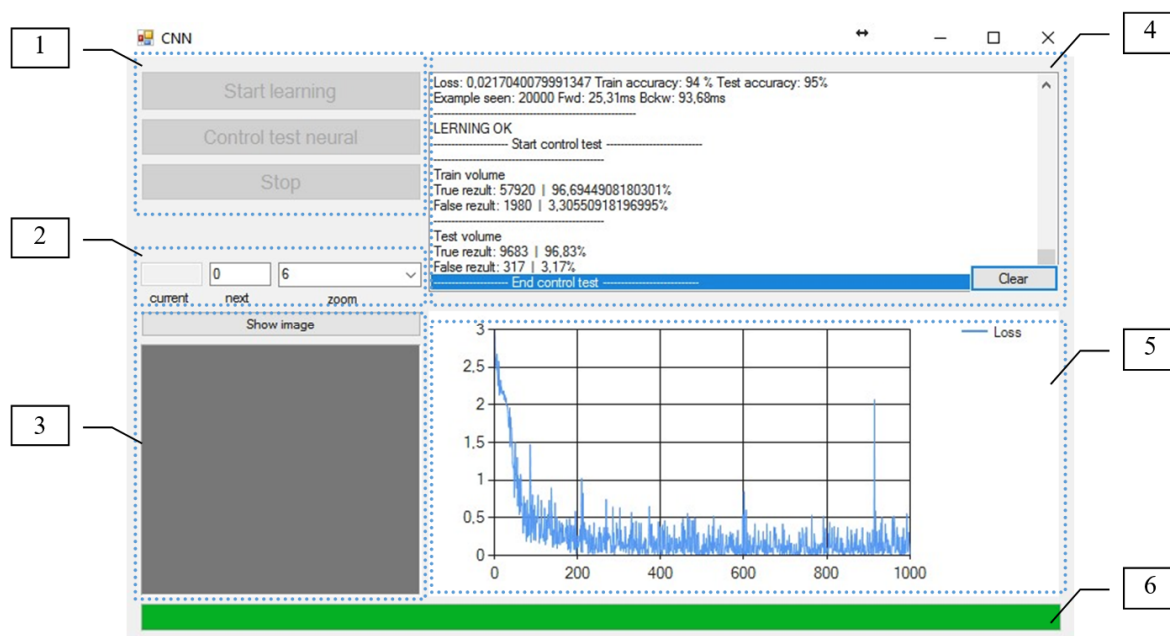


Fig. 2. The example of the application’s main form

The dialog with user consists of the following steps:

- download the image from the database;
- start training, clicking the corresponding button;
- after the neural network is trained, do the control test, clicking the corresponding button.

The following settings are available for user configuration:

- speed of training;
- sample size of images to be submitted for a single entry;
- moment;
- number of training epochs.

In an empirical way the following recommendations for the neural network settings were determined and can be listed:

1. As for the speed of training, this parameter should be set small, because there may be cases when the network will simply rewind the result we need, which leads to a longer training time.

2. The size of the subsets is usually given by more than the number of outputs. The larger the size of the subsets, the longer one training epoch will take. The most optimal subset size was 20. This is sufficient to achieve good results.

3. The moment parameter is selected close to one, for example here 0.9 was used. This parameter is responsible for ensuring that the neural network during training tries to find not a local minimum, but a global one.

4. The number of training epochs as a parameter was introduced in this work in order to experimentally obtain training results with a different number of epochs and compare them. Tests were carried out with the number of epochs 200, 400, 600, 800 and 1000.

The result of the information system using is an answer for the user about the recognized digit. Recognition accuracy was estimated for various combinations of neural network parameters. Fig. 3 illustrates how the average accuracy depends on the number of epochs.

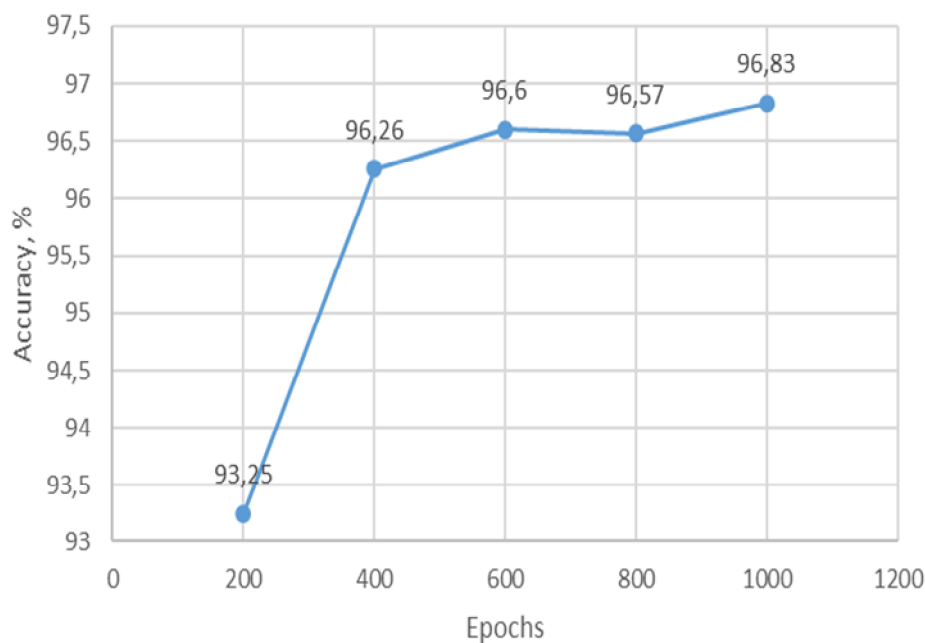


Fig. 3. Dependence of the recognition accuracy from the number of epochs

The graph (Fig. 3) shows that in the best case out of 10 thousand images, the network did not recognize only 317 or 3.17 %. To improve results, it is possible to:

- complicate the neural network topology;
- increase the number of epochs;
- try to set training speed, subsets, or moment settings differently.

Complication of the neural network topology will lead to an increase in its training time. On average, training in 200 eras on a given convolutional network topology takes about 20 minutes, and the full test after network training takes about 40 minutes.

In order to analyze the results of recognition, a transaction log was created. After each iteration of the input image recognition, the report record was added to the log consisting of fields: the index of the input image, the recognition response, the mark of the operation correctness. A sample of data on the most often unrecognized digits (the recognition error is close to 100 %) showed, that even for human being it would be difficult to recognize digit because of the specific handwriting.

### Conclusions

The object of the presented paper was the process of handwritten digits' recognition using a neural network approach. The paper presents the results of the design and software implementation of the information system, which uses the proposed model of convolutional neural network and provides interaction with users. Input data for information system is a graphical image of handwritten digits,

and the result of the neural network is a response with the results of recognition. As a system's training and testing dataset a database MNIST, containing images of handwritten digits written in various handwriting was used. The recognition precision for the developed neural network is 96.83 %. It can be assumed that the proposed convolutional network topology can be improved by debugging it and optimizing the learning algorithm. Also, the way of the recognition results improvement can be the pre-processing of the input images.

### References

- [1] Serey, J., Alfaro, M., Fuertes, G., Vargas, M., Duran, C., Ternero, R., Rivera, R., Sabattin, J. (2023). Pattern recognition and deep learning technologies, enables of industry 4.0 and their role in engineering research. *Symmetry*, 15(2), 535—539. doi: 10.3390/sym15020535.
- [2] Boesch, G. (2024). What is a pattern recognition? A gentle introduction. Retrieved from <https://viso.ai/deep-learning/pattern-recognition/>.
- [3] Bhamare, D., Suryawanshi, P. (2018). Review on Reliable Pattern Recognition with Machine Learning Techniques. *Fuzzy Information and Engineering*, 10(3), 362—377. doi: 10.1080/16168658.2019.1611030
- [4] Sushmitha, V., Hussain, S. (2022). A survey on handwritten digit recognition. *International Journal of Innovative Research in Computer and Communication Engineering*, 10, 1—5.
- [5] Gong, Y., Zhang, P. (2021). Research on MNIST Handwritten Numbers Recognition based on CNN. *Journal of Physics*, 2138, 1—7. doi:10.1088/1742-6596/2138/1/012002.
- [6] Karakaya, R., Kazan, S. (2021). Handwritten digit recognition using machine learning. *Sakarya University Journal of science*, 25(1), 65—71. doi: 10.16984/saufenbilder.801684.
- [7] Dixit, R., Kushwah, R., Pashine, S. (2020). Handwritten digit recognition using machine and deep learning algorithms. *International Journal of Computer Applications*, 176(42), 27—35.
- [8] Ahmed, S.S., Mehmood, Z., Awan, I.A., Yousaf, R.M. (2023). A Novel Technique for Handwritten Digit Recognition Using Deep Learning. *Journal of sensors*, 2023, 1—15. doi: 10.1155/2023/2753941.
- [9] Azawi, N. (2023). Handwritten digits recognition using transfer learning. *Computers and electrical engineering*, 106, 1—10.
- [10] Chen, M.R., Chen, B.P., Zeng, G.Q., Lu, K.D., Chu, P. (2019). An adaptive fractional-order BP neural network based on extremal optimization for handwritten digits recognition. *Neurocomputing*, 391, 260—272. doi: 10.1016/j.neucom.2018.10.090.
- [11] Kaymov, Z., Tumakov, D., Mosin, S. (2020). Hierarchical convolutional neural network for handwritten digits recognition. *Procedia Computer Science*, 171, 1927—1934.
- [12] Swetha, K., Hithaishi, Y., Tejaswini, N.L., Pathasaradhi, P., Rao, P.V. (2021). Handwritten digit recognition using OpenCv and CNN. *International journal of creative research thoughts*, 9, 211—220.
- [13] Hamin, S.M., Miah, M.B., Sarker, A., Rana, M., Jobair, A. (2018). Handwritten Digit Recognition using Machine Learning Algorithms. *Global Journal of Computer Science and Technology: Neural & Artificial Intelligence*, 18, 17—23.
- [14] Patil, P., Kaur, B. (2020). Handwritten digit recognition using various machine learning algorithms and models. *International Journal of Innovative Research in Computer Science & Technology*, 8, 337—340. doi: 10.21276/ijirest.2020.8.4.16.
- [15] Ahlawat, S., Choudhary, A., Nayyar, A., Singh, S., Yoon, B. (2020). Improved handwritten digit recognition using convolutional neural networks. *Sensors*, 20, 1—18. doi:10.3390/s20123344.

### Список використаної літератури

1. Serey J., Alfaro M., Fuertes G., Vargas M., Duran C., Ternero R., Rivera R., Sabattin J. Pattern recognition and deep learning technologies, enables of industry 4.0 and their role in engineering research. *Symmetry*. 2023. 15(2). 535—539. doi: 10.3390/sym15020535.



2. Boesch G. What is a pattern recognition? A gentle introduction. URL: <https://viso.ai/deep-learning/pattern-recognition/> (дата звернення 01.02.2024).
3. Bhamare D., Suryawanshi P. Review on Reliable Pattern Recognition with Machine Learning Techniques. *Fuzzy Information and Engineering*. 2018. 10(3). 362—377. doi: 10.1080/16168658.2019.1611030
4. Sushmitha V., Hussain S. A survey on handwritten digit recognition. *International Journal of Innovative Research in Computer and Communication Engineering*. 10. 2022. 1—5.
5. Gong Y., Zhang P. Research on Mnist Handwritten Numbers Recognition based on CNN. *Journal of Physics*. 2021. 2138. 1—7. doi:10.1088/1742-6596/2138/1/012002
6. Karakaya R., Kazan S. Handwritten digit recognition using machine learning. *Sakarya University Journal of science*. 2021. 25(1). 65—71. doi: 10.16984/saufenbilder.801684.
7. Dixit R., Kushwah R., Pashine S. Handwritten digit recognition using machine and deep learning algorithms. *International Journal of Computer Applications*. 2020. 176(42). 2020. 27—35.
8. Ahmed S.S., Mehmood Z., Awan I.A., Yousaf R.M. A Novel Technique for Handwritten Digit Recognition Using Deep Learning. *Journal of sensors*. 2023. 1-15. doi:10.1155/2023/2753941.
9. Azawi N. Handwritten digits recognition using transfer learning. *Computers and electrical engineering*. 2023. 106. 1—10.
10. Chen M.R., Chen B.P., Zeng G.Q., Lu K.D., Chu P. An adaptive fractional-order BP neural network based on extremal optimization for handwritten digits recognition. *Neurocomputing*. 2019. 391. 260—272. doi: 10.1016/j.neucom.2018.10.090.
11. Kaymov Z., Tumakov D., Mosin S. Hierarchical convolutional neural network for handwritten digits recognition. *Procedia Computer Science*. 2020. 171. 1927—1934.
12. Swetha K., Hithaishi Y., Tejaswini N.L., Pathasaradhi P., Rao P.V. Handwritten digit recognition using OpenCv and CNN. *International journal of creative research thoughts*. 2021. 9. 211—220.
13. Hamin S.M., Miah M.B., Sarker A., Rana M., Jobair A. Handwritten Digit Recognition using Machine Learning Algorithms. *Global Journal of Computer Science and Technology: Neural & Artificial Intelligence*. 2018. 18. 17—23.
14. Patil P., Kaur B. Handwritten digit recognition using various machine learning algorithms and models. *International Journal of Innovative Research in Computer Science & Technology*. 2020. 8. 337—340. doi:10.21276/ijirest.2020.8.4.16.
15. Ahlawat S., Choudhary A., Nayyar A., Singh S., Yoon B. Improved handwritten digit recognition using convolutional neural networks. *Sensors*. 2020. 20. 1—18. doi:10.3390/s20123344.

Надійшла до редколегії 03.04.2024