КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

UDC 519.7
DOI: https://doi.org/10.17721/1812-5409.2025/1.12

Igor AIZENBERG, PhD (Comp. Science), Prof.
ORCID ID: 0000-0002-5994-6568
e-mail: igor.aizenberg@manhattan.edu
Manhattan University, Riverdale, New York, USA

Alexander VASKO, PhD Student
ORCID ID: 0009-0006-1527-505X
e-mail: oleksandrvasko@uzhnu.edu.ua
Uzhhorod National University, Uzhhorod, Ukraine

COMPARATIVE ANALYSIS OF CNNMVN AND MLMVN AS FREQUENCY DOMAIN CNN CONVOLUTIONS

Each convolutional layer in any convolutional neural network produces a feature map containing the most important information, which a network needs to recognize respective images. To further improve these neural networks and better understand their capabilities, it is essential to discover, which features are actually extracted and how the images to be recognized are transformed by convolutions resulted from the learning process. This paper presents a comparative analysis of convolutions obtained via two complex-valued neural networks based on multi-valued neurons. The first network is a convolutional neural network based on multi-valued neurons (CNNMVN) which has a traditional convolutional neural network topology except of that it employs complex-valued convolutional kernels in its convolutional part and multi-valued neurons in its fully connected part. The second one is the multi-valued neural network based on multi-valued neurons (MLMVN) which is a fully connected multilayer neural network employed as a convolutional network in the frequency domain.

Considering that both neural networks are complex-valued and the obtained filters operate in the complex domain, the conducted research indicates that the kernels of both networks produce filters similar to existing digital image processing filters. The analysis of CNNMVN kernels revealed that they implement unsharp masking filters and edge detection filters for identifying shapes in images, while the MLMVN kernels enhance specific frequency sub-bands. The latter means that the respective filters are mostly not similar to the ones known as unsharp masking or sharpening filters. Thus, the kernels of both convolutional networks contribute to improving image recognition performance in their own ways.

K e y w o r d s : ***convolutional neural networks; CNNMVN; complex-valued neural networks; multi-valued neuron; MLMVN; image recognition; frequency domain.***

AMS 2020 classification: 68T07.

Introduction

A convolutional neural network (CNN) has a special place in the field of neural networks and deep learning since its introduction in (Krizhevsky, Sutskever, & Hinton, 2012). Originally inspired by the human visual system (Jarrett et al., 2009; LeCun et al., 2004), CNNs become a primary tool for visual information processing. In general, all real-world images may contain many details that may not be always useful for classification or recognition. The main distinction of CNN from the other neural networks is that CNN has special convolutional layers that can first extract useful information from visual data and then perform classification and recognition based only on it. This approach showed excellent results compared to regular feedforward neural networks (Beysolow Ii, 2017; Caldeira et al., 2019; Gad, 2018; Kaur, & Garg, 2020; Lin et al., 2019; Lin et al., 2020; Sarabu, & Santra, 2021; Venkatesan, & Li, 2017; Wu, Zhang, & Zhao, 2020; Xiao et al., 2018; Yar et al., 2021).

Among all known neural networks, we would like to focus our attention in this work on complex-valued networks. They use complex numbers as weights, employ complex-valued activation functions, and their inputs and outputs are also complex-valued, accordingly. Various neural networks based on complex-valued neurons with a complex hyperbolic tangent activation function were considered during the last decades (Hirose, 2011; Kobayashi, 2017; Valle, 2014). These networks demonstrated a number of advantages over their real-valued counterparts (Boonsatit et al., 2022; Nitta, 2004; Nitta, 2008; Suresh, Sundararajan, & Savitha, 2013; Zhang et al., 2022). The first complex-valued convolutional neural network (CV-CNN) was introduced in (Bruna et al., 2015) and the first quaternion-based convolutional neural network (QV-CNN) was introduced in (Zhu et al., 2018). Over the past years, significant progress has been made in the development and application of CV-CNNs and QV-CNN, leading to their successful implementation across various domains. Subsequent studies (Chatterjee et al., 2023; Fuchs et al., 2021; Guberman, 2016; Hongo et al., 2020; Popa, 2017; Rawat, Rana, & Kumar, 2021; Sunaga, Natsuaki, & Hirose, 2020; Yadav, & Jerripothula, 2023; Zhang et al., 2017) introduce both CV-CNNs and QV-CNNs as powerful tools for complex image processing tasks. Notably, these models have demonstrated their effectiveness in synthetic-aperture radar (SAR) image analysis, where they have been employed for both image recognition and noise filtering. In addition to SAR applications, CV-CNNs and QV-CNNs have proven to be valuable in broader image recognition tasks, offering advantages in handling complex-valued data and capturing richer spatial and spectral features compared to traditional real-valued CNNs.

Another class of complex-valued neural networks is based on multi-valued neurons (MVN). MVN and MVN-based networks are comprehensively described in (Aizenberg, 2011). The multilayer feedforward neural network with multi-valued neurons (MLMVN) was first introduced in (Aizenberg, & Moraga, 2007). The following features should be mentioned among advantages: derivative-free learning algorithm not suffering from the local minima problem; it often learns faster and develop better generalization capability compared to regular MLP (Aizenberg, 2011; Aizenberg, & Moraga, 2007). A convolutional neural network based on MVN (CNNMVN) was introduced in (Aizenberg, & Vasko, 2020), and its modification was presented in (Aizenberg, Herman, & Vasko, 2022). CNNMVN is a network with a standard convolutional neural network topology. It consists of the convolutional layers that can be followed by pooling layers, other convolutional layers or the dense layers.

ISSN 1812-5409 (Print), ISSN 2218-2055 (Online)

It was shown in (Aizenberg, & Vasko, 2024) that CNNMVN is the further generalization of MLMVN. The use of MLMVN as a frequency domain CNN (MLMVN-FCNN) was suggested in (Aizenberg, & Vasko, 2023). According to the Convolution Theorem, the convolution operation can be calculated in both spatial and frequency domains. Thus, it was suggested to interpret and use the first hidden layer of MLMVN as a convolutional layer in the frequency domain.

In this paper, we present a deep analysis of feature maps for both CNNMVN and MLMVN-FCNN. This should help to better understand what kind of information is extracted and which convolutional kernels are constructed during the learning process.

1. CNNMVN

CNNMVN combines the principles of convolutional neural networks and multi-valued neurons (Aizenberg, Herman, & Vasko, 2022; Aizenberg, & Vasko, 2020; 2024). It has a topology traditional for CNNs, employes complex-valued weights and continuous MVN activation function for its convolutional kernels, and has an MVN-based fully connected part.

As any convolutional neural network, CNNMVN consists of convolutional, pooling and the dense layers. For proper processing, intensities of a real-valued input image should be converted to complex-valued ones according to the following rule

$$z_j = e^{i\varphi_j}, \qquad \varphi_j = \frac{x_j - \min\{X\}}{\max\{X\} - \min\{X\}} \cdot \lambda, \tag{1}$$

where x_j and z_j stand for a real-valued datapoint and a datapoint projected onto the unit circle, respectively; φ_j is the argument (phase) of the projected datapoint z_j; $\min\{X\}$ and $\max\{X\}$ are respectively the minimum and maximum values over a real-valued dataset; parameter λ defines a pre-determined part of the unit circle to which real-valued data are projected (a standard value is $\lambda = 3\pi/2$).

Each convolutional layer consists of kernels whose purpose is to extract some useful features from an input image. All kernels slide over an input image (or a feature map) performing a convolution operation. It computes a weighted sum of the input values within its receptive field. This process helps to extract specific features such as edges, shapes and textures from images to be analyzed. Convolutional kernels also capture spatial relationships between pixels. It is widely known that adjacent pixels in an image are often highly correlated, and kernels can identify local patterns because of this property. Also, in a case when multiple convolutional layers are used, kernels can identify larger features by combining smaller local features that were extracted in preceding layer(s).

A layer following a convolutional one can be either another convolutional layer or a pooling layer. If the latter is the case, its main purpose is downsampling and reduction of the information to be processed. The convolutional and the pooling layers are followed by fully connected layers. This part of CNNMVN is nothing else but MLMVN.

A CNNMVN learning algorithm, that is the process of error backpropagation and error correction in CNNMVN was presented in (Aizenberg & Vasko, 2020), further develop in (Aizenberg, Herman, & Vasko, 2022) and then presented in detail in (Aizenberg, & Vasko, 2024).

2. MLMVN as a Frequency domain CNN (MLMVN-FCNN)

The idea to use MLMVN as a frequency-domain convolutional neural network was introduced in (Aizenberg, & Vasko, 2023) and then considered in detail in (Aizenberg, & Vasko, 2024). It was inspired by the fact that convolution operation can be performed in either of the spatial or frequency domains. The convolution of two signals $f(t)$ and $g(t)$ can be obtained as

$$(f * g)(t) = \mathcal{F}^{-1}(F \cdot G). \tag{2}$$

where $\mathcal{F}^{-1}$ stands for the inverse Fourier transform, F and G are the Fourier transforms of signals $f(t)$ and $g(t)$. According to (2) the Convolutions Theorem states that spatial domain convolution of two signals is equivalent to the frequency-domain product of the respective Fourier transforms of these signals.

Let us now consider the first hidden layer in MLMVN. Let vector $(x_1, \dots, x_n)$ and vector $\left(w_0^i, w_1^i, \dots w_n^i\right)$ be the i^{th} neuron's inputs and weights, respectively. Then a weighted sum z_i is calculated as

$$z_i = w_0^i + w_1^i x_1 + w_2^i x_2 + \cdots + w_n^i x_n \tag{3}$$

According to (3), if the input vector $(x_1, \dots, x_n)$ represents a vectorized 2D Fourier transform of an image, we can assume that each neuron in the first layer of MLMVN performs a convolution in the frequency domain. Consequently, the first layer can be interpreted as a convolutional layer operating in the frequency domain and the neurons' weights interpreted as convolutional kernels in the frequency domain.

This approach has several key distinctions when compared to traditional CNNs. The first one is that in the first hidden layer, the result of each neuron's convolution is not a feature map as it is in regular CNNs. Instead, it is a sum of the frequency-domain feature map coefficients. And as a result, there is no possibility to apply another convolutional layer. The second distinction is that the kernel size in the spatial domain cannot be manually set, and therefore, we cannot directly control the size of the features being extracted.

To improve performance and network's generalization capability, some modifications including frequency domain pooling were suggested. The main purpose of the frequency domain pooling is to reduce the number of calculations and remove frequencies, which are not important for recognition of objects of interest. Frequency domain pooling is reduced to extraction of Fourier coefficients corresponding only to the 2D frequencies, which are essential to recognition of certain objects – from the 1st frequency to the k-th frequency, which is a cut-off frequency in this case.

3. Comparative analysis of convolutions

As it was said before, the main purpose of any convolutional layer is to extract important features from an image to make it easier to classify/recognize it. Considering that both CNNMVN and MLMVN-FCNN are complex valued, it is interesting to visualize and compare feature maps that are generated in convolutional layers. Analyzing these feature maps, we may better understand and hopefully improve the learning process of both networks.

We used in our experiments two popular benchmarks: MNIST – the dataset of handwritten digits (LeCun et al., n. d.) and Fashion MNIST – the dataset of fashion images (Xiao, Rasul, & Vollgraf, 2017). Both datasets contain 60 000 images for training and 10 000 images for testing. All images are gray-scale and therefore their intensity range is $[0, 255]$. The size of each image is 28×28 pixels.

To visualize and analyze the feature maps, we used trained networks, which were presented in (Aizenberg, & Vasko, 2024). For CNNMVN we analyzed models with same topology for both MNIST and Fashion MNIST datasets. It was 32C5-h256-o10 topology, where 32C5 stands for the convolutional layer with 32 kernels of size 5×5, h-256 is the number of neurons in the hidden layer and

o10 is the number of output neurons. The accuracy was 97,71% and 88,06% for MNIST and Fashion MNIST datasets, respectively. To analyze and visualize feature maps of MLMVN-FCNN we used a model with the h2048-o10 topology. For this model, we got 98.11% accuracy on the MNIST dataset with 5 extracted 2D frequencies. To analyze Fashion MNIST feature maps, we used a model with the h3072-o10 topology. For this model, we got 90% accuracy with 7 extracted 2D frequencies.

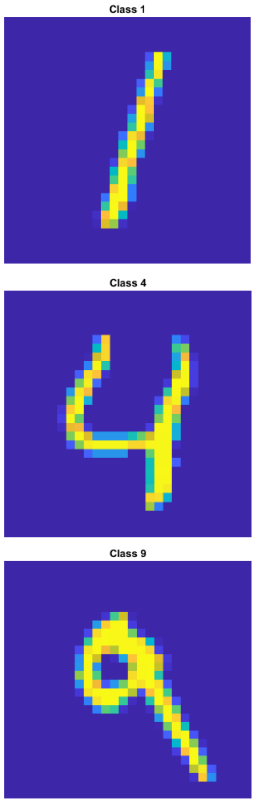


Fig. 1 (a). Original images of 1, 4 and 9 classes

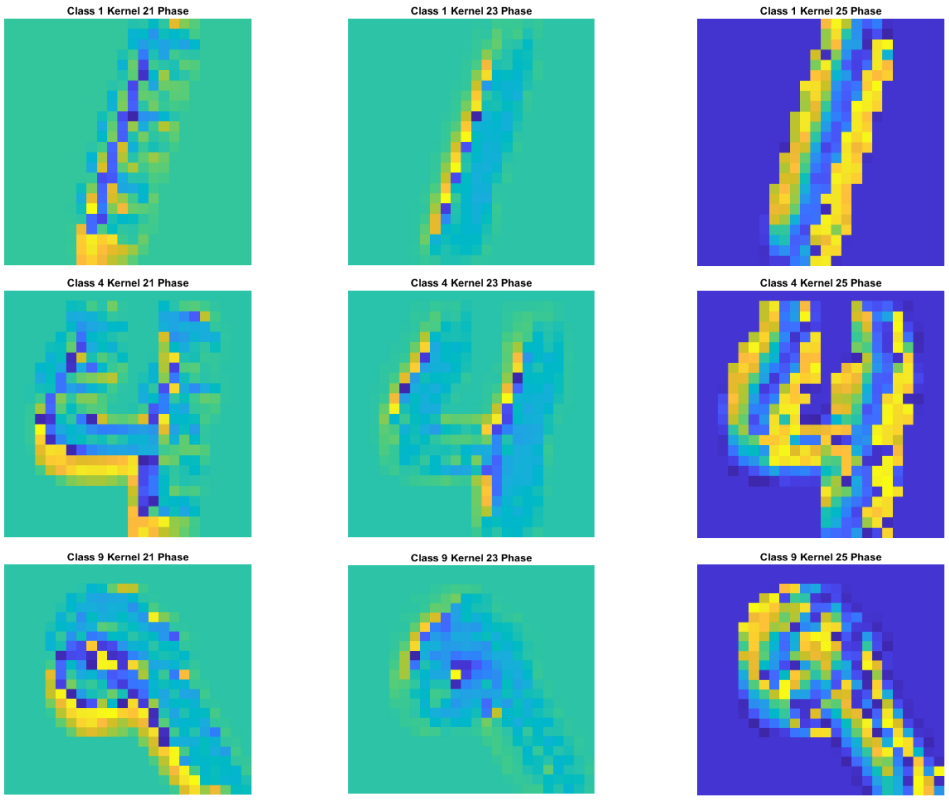


Fig. 1 (b). Phases of the feature maps for 21, 23 and 25 kernels of CNNMVN in MNIST dataset

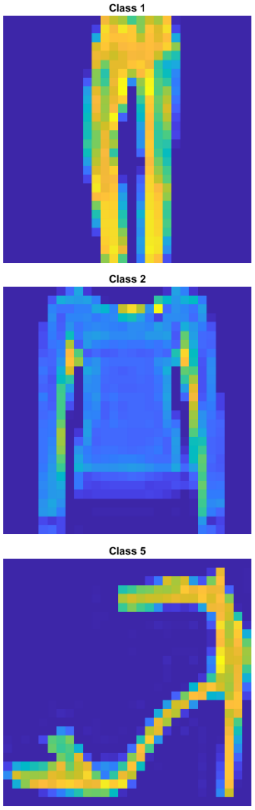


Fig. 1 (c). Original images of 1, 2 and 5 classes

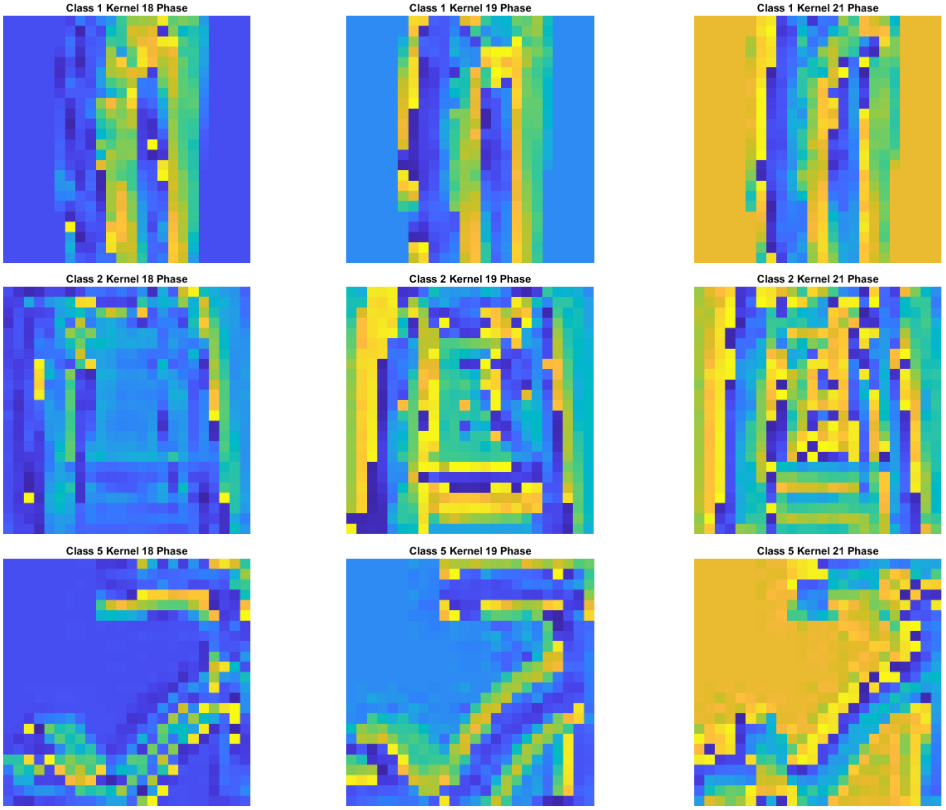


Fig. 1 (d). Phases of the feature maps for 18, 19 and 21 kernels of CNNMVN in Fashion MNIST dataset

The visualization of the feature maps for MNIST and Fashion MNIST datasets of CNNMVN are presented in Fig. 1. In Fig. 1 (a) we observe the phases of the original images of MNIST dataset for the classes 1,4 and 9, transformed according to (1). Figure 1 (b) presents the phases of the feature maps generated by the kernels 21, 23 and 25 of CNNMVN after the learning process was completed. One can observe that kernel 21 is performing horizontal and diagonal lines detection while kernel 23 better detects vertical lines. The kernel 25 performs high frequency emphasizing, that is, it performs filtering, which can be compared to unsharp masking. Similar results are observed for the Fashion MNIST dataset. Fig. 1 (c) illustrates the phases of the original images for the classes 1, 2 and 6, transformed by (1). Figure 2 (d) depicts phases of the respective feature maps for kernels 18, 19, and 21 after the CNNMVN training. We can observe that kernels 18 and 19 are performing high-pass filtering, which leads to shape highlighting of the image. Kernel 21 performs some kind of high frequency emphasizing which can be compared to unsharp masking. Summarizing the above, we can conclude that CNNMVN kernels produce feature maps similar to those of traditional CNNs. Like their real-valued counterparts, CNNMVN kernels aims to detect shapes and edges in an input image through the development of various types of filters.

While CNNMVN performs convolution in the spatial domain, MLMVN-FCNN performs it in the frequency domain. As mentioned earlier, this approach has several limitations. We cannot manually set the kernel size in the spatial domain and therefore we cannot have control over the size of details extracted by a respective convolution. Additionally, as we are dealing in this case with the frequency domain kernels, the learning process can be resulted in kernels that often are difficult to compare to their spatial domain counterparts.

The visualization of the MLMVN feature maps for both datasets is presented in Fig. 2. In Fig. 2 (a) samples from the MNIST dataset after frequency domain pooling are presented with focusing on classes 1, 4, and 9, and with the frequencies (1–5) extracted. Fig. 2 (c) depicts the Fashion MNIST classes after the frequency domain pooling, specifically classes 1, 2, and 6, with the frequencies (1–7) were extracted. Figures 2 (b) and 2 (d) visualize the feature maps resulted from the learning process of MLMVN. Since the convolution process in MLMVN is performed in the frequency domain, we applied the inverse Fourier transform to the feature maps, and then their magnitudes were extracted to visualize them. It should also be mentioned that according to the frequency domain pooling algorithm, the 2D frequencies are extracted in a "diamond" shape (Fig. 3), have been flattened and then processed by the first hidden layer neurons in MLMVN. Thus, a network learns employing only Fourier coefficients extracted from the kernel according to the "diamond" rule. It should also be taken into consideration that kernels and feature maps have ripple effects because of this when visualized. It is significant when analyzing kernels and feature maps in the frequency domain.

It is clearly seen in Fig. 2 (b) that all filters resulted from the learning process in the frequency domain distinguish certain features in the MNIST images. These filters emphasize medium and high frequencies and can be compared to unsharp masking filters. At the same time, while unsharp masking filter usually emphasize all high frequencies and possibly some medium frequencies, MLMVN neurons as a result of learning emphasize very specific narrow frequency sub-bands. Thus, they develop filters distinguishing specific details. For example, as it is seen from Fig. 2 (b), neuron 351 distinguishes details of certain specific sizes by emphasizing respective medium frequencies. It can also be seen from Fig. 2 (b) that neuron 461 distinguishes vertically oriented details in images by emphasizing vertical frequencies, while neuron 466 distinguishes diagonally oriented details (from bottom-left to top-right) by emphasizing respective frequencies.

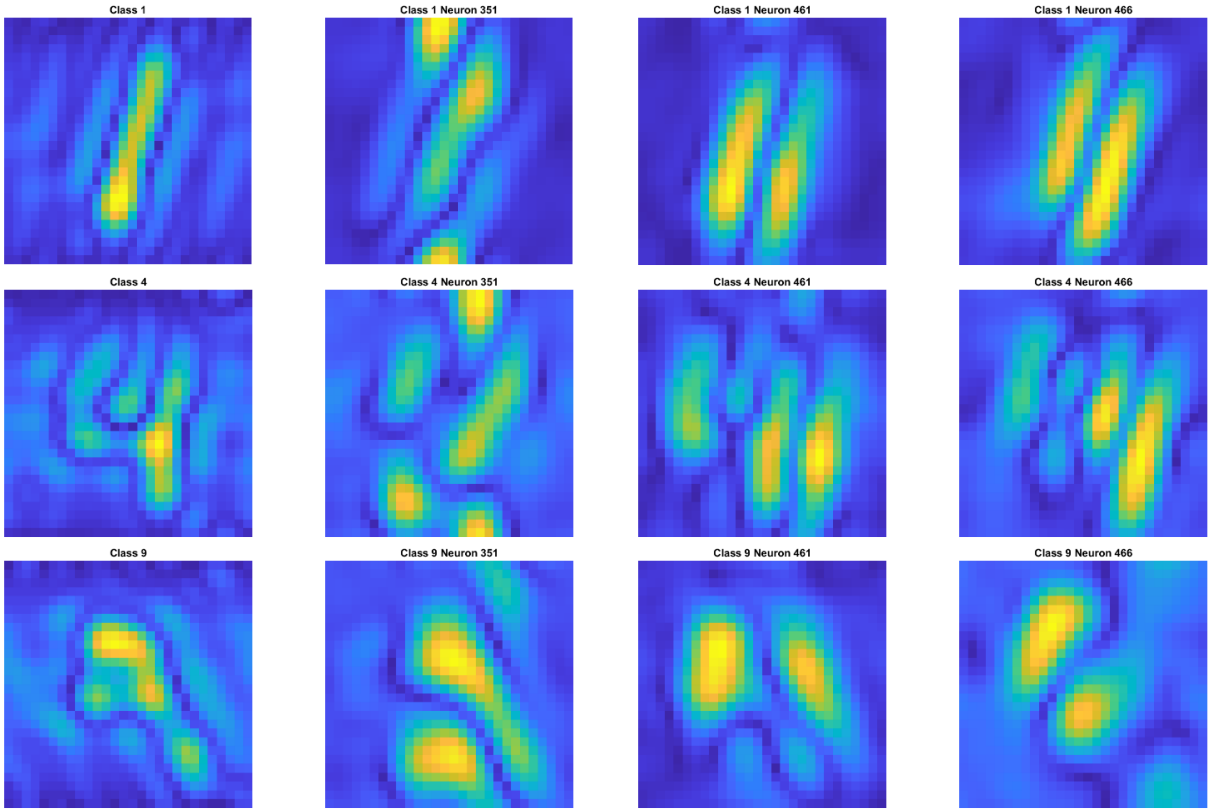


Fig. 2 (a). Classes 1,4, and 9 after frequency domain pooling

Fig. 2 (b). Magnitudes of the inverse Fourier transform for neurons 351, 461, and 466 neurons of MLMVN-FCNN in MNIST dataset

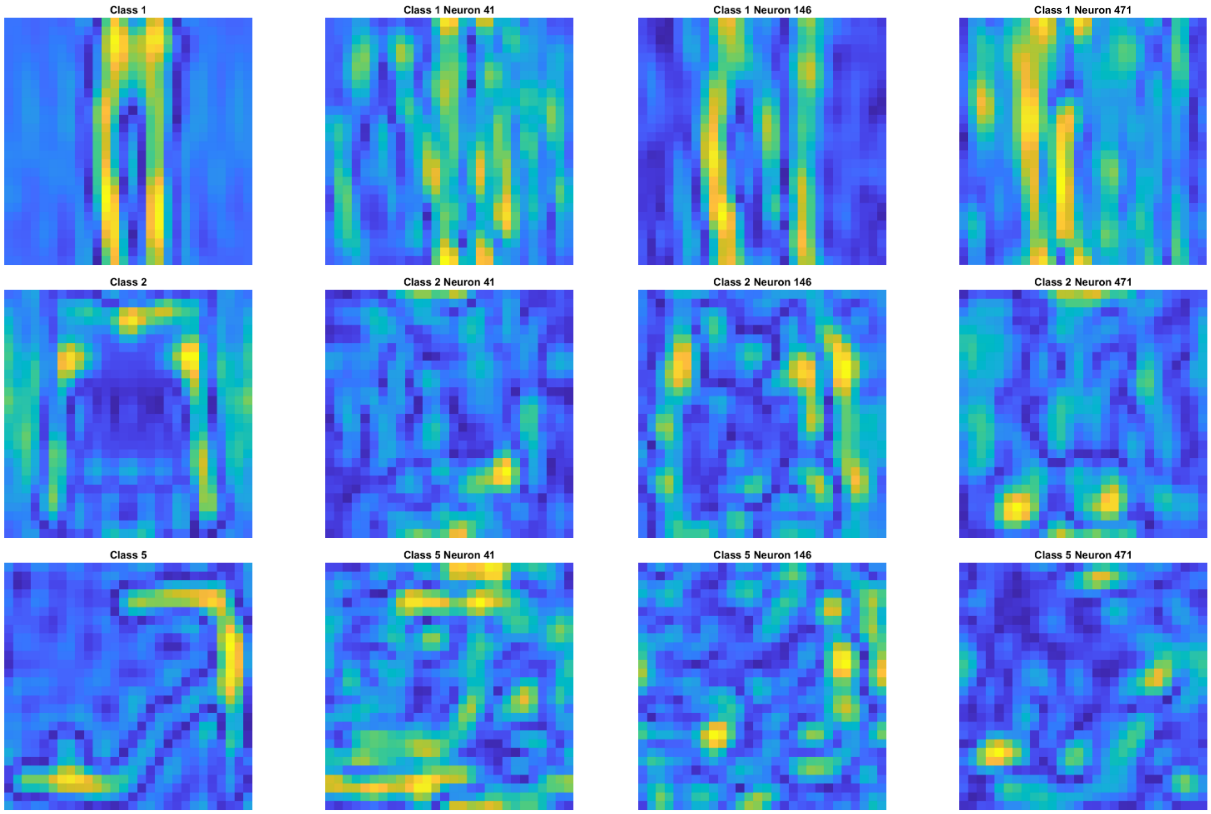
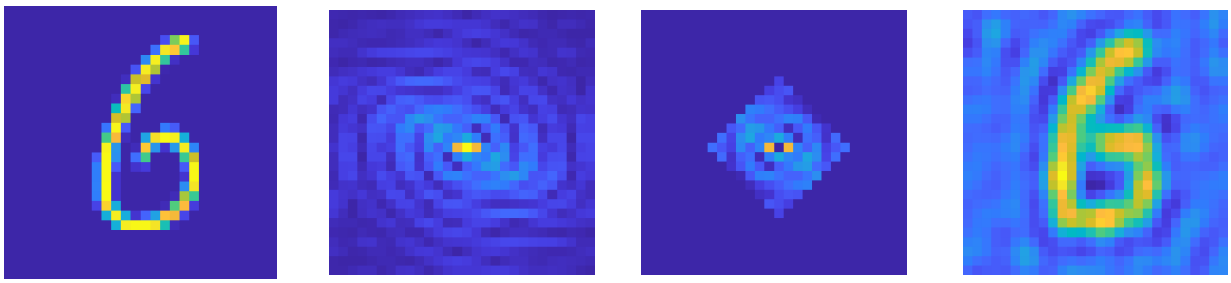


Fig. 2 (c). Classes 1,2 and 6 after frequency domain pooling

Fig. 2 (d). Magnitudes of the inverse Fourier transform for neurons 41, 146, and 471 of MLMVN-FCNN in Fashion MNIST dataset



(a) (b) (c) (d)

Fig. 3. Frequency pooling process: a) original image; b) Fourier frequencies of the image with circular shift; c) extracted Fourier coefficients corresponding to the frequencies 1-5 around the DC frequency; d) inverse Fourier transform of the extracted Fourier coefficients corresponding to the frequencies 1–5

A similar result can be observed in Fig. 2 (d) for the Fashion MNIST images, but with some distinctions. Since these images are more complicated and lost some details due to the frequency domain pooling (the frequencies (1–7) were extracted), they retain the general shape. For example, neuron 41 detects horizontally oriented details in the images from classes 2 and 6. At the same time, while there are no horizontally oriented details in the image belonging to class 1, it highlights some isolated elements belonging to vertically oriented details in this image. Neuron 146 highlights vertically oriented details for all classes and the kernel 471 highlights various shapes for different classes by emphasizing medium frequencies.

It is worth looking at the illustrations showing exactly which frequencies are amplified by the filters resulting from training both neural networks. To discover this, we need to look at Fourier power spectra of the respective filter kernels. To obtain respective power spectra of CNNMVN kernels, we took Fourier transform of the spatial domain kernels resulting from the learning process. Figure 4 (a) represents respective power spectra of kernels 21, 23, 25 for the MNIST dataset, and Fig. 4 (b) represents respective powers spectra of kernels 18, 19, 21 for the Fashion MNIST dataset. Thus, power spectra in Fig. 4 (a) represent the filters used to create feature maps shown in Fig. 1 (b), while power spectra in Fig. 4 (b) represent the filters used to create feature maps shown in Fig. 1 (d).

Power spectra of MLMVN-FCNN kernels can easily be obtained by taking absolute values of the respective weights and reshaping them into 2D structures. Figure 5 (a) represents respective power spectra of kernels 351, 451, 466 for the MNIST dataset. We may observe from Fig. 5 (a) that medium frequencies are emphasized by neuron 351, and this leads to the extraction of details of certain specific sizes as it can be seen in Fig. 2 (b). This example also shows that vertical frequencies are emphasized by neuron 461, and this is resulted in distinguishing vertical details as can be seen in Fig. 2 (b). As we see from Fig. 5 (a), neuron 466 emphases frequences corresponding to the diagonally oriented image elements, and this can be seen in Fig. 2 (b). Figure 5 (b) represents respective powers spectra of kernels 41, 146, 471 for the Fashion MNIST dataset, that is they represent the filters used to create feature maps shown in Fig. 2 (d).

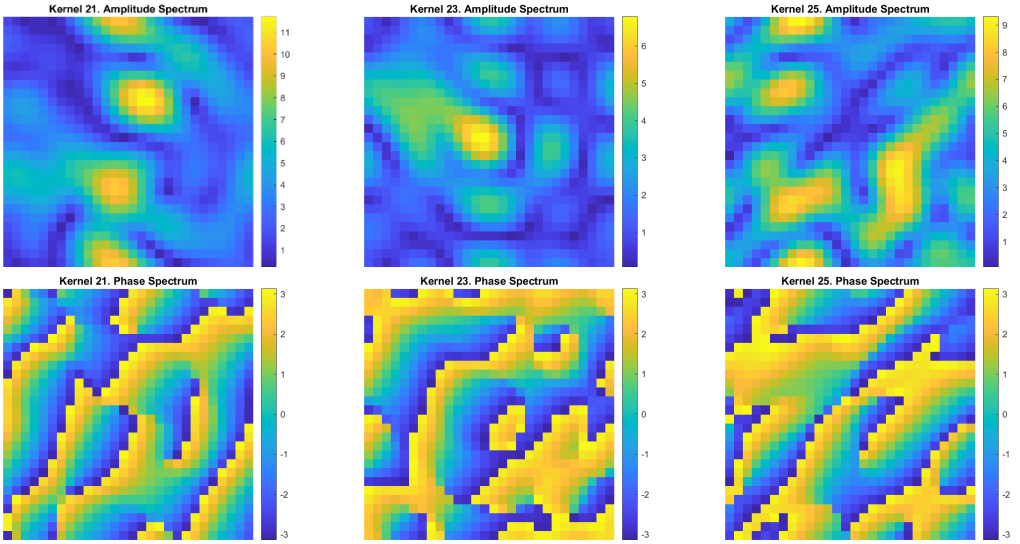


Fig 4 (a). Amplitude and Phase spectrum of kernels 21, 23, 25 of CNNMVN in MNIST dataset

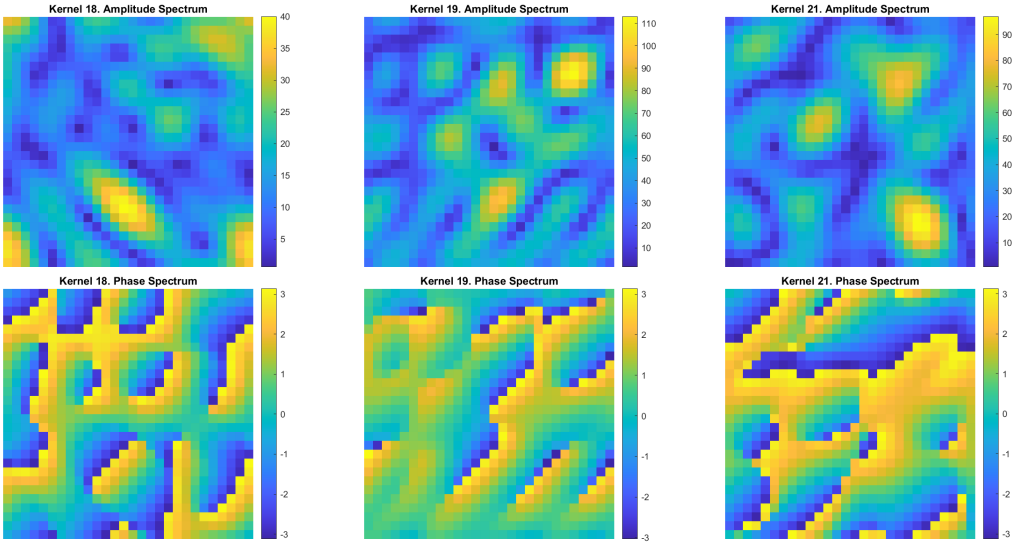


Fig. 4 (b). Amplitude and Phase spectrum of kernels 18, 19, 21 of CNNMVN in Fashion MNIST dataset

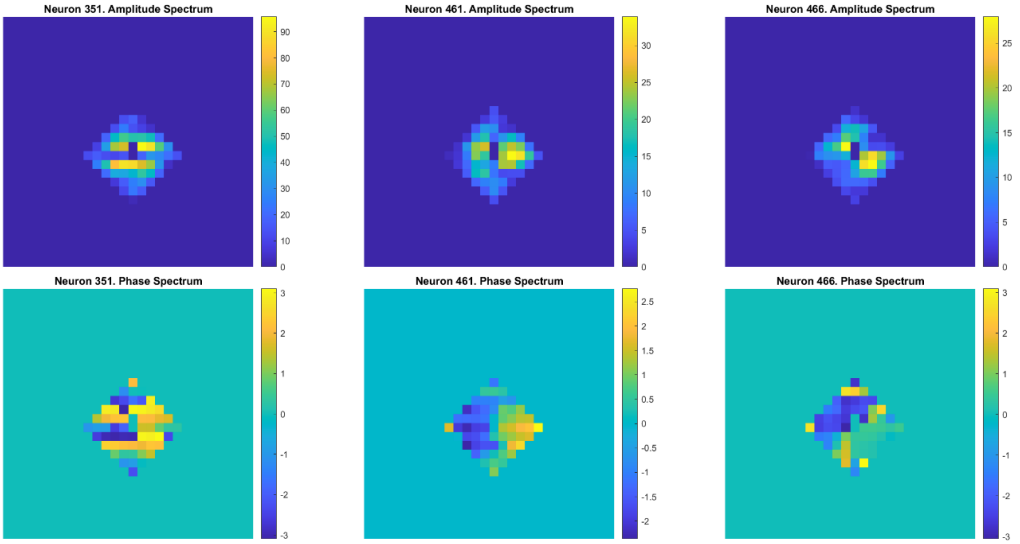


Fig. 5 (a). Amplitude and Phase spectrum of kernels 351, 451, 466 of MLMVN-FCNN in MNIST dataset

ISSN 1728-3817

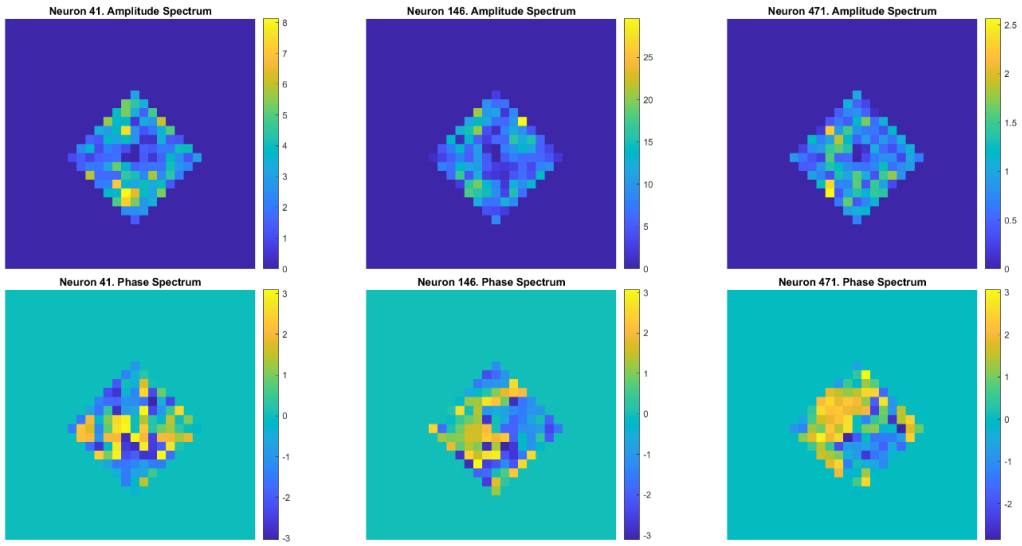


Fig. 5 (b). Amplitude and Phase spectrum of kernels 41, 146, 471 of MLMVN-FCNN in Fashion MNIST dataset

Considering that both networks reached high accuracy rate we can conclude that the extracted features, which are visualized in Fig. 2 (b) and 2 (d) are crucial for image classification. Phase spectra of the respective kernels are also shown in Fig. 4 (a), 4 (b), 5 (a) and 5 (b), but it can be seen that it does not carry some visually significant information. However, it is worth noting that visualization of the magnitude of complex-valued feature maps may not provide a complete understanding of their nature.

Discussion and conclusions

Summarizing the above, we can conclude that both CNNMVN and MLMVN-FCNN networks successfully extract meaningful features from input images despite using different approaches. Both models demonstrate their ability to generate meaningful feature maps that emphasize important image details such as edges, shapes, and specific frequency bands. CNNMVN operates in the spatial domain and produces feature maps similar to those in traditional CNNs. Its kernels are effective at detecting specific patterns such as horizontal, vertical, and diagonal lines, as well as performing high-pass filtering to highlight image details. In contrast, MLMVN-FCNN operates in the frequency domain, emphasizing specific frequency sub-bands that are difficult to directly compare to spatial domain kernels.

Future work will focus on investigating which regions, shapes and features of an image most frequently activate the kernels, allowing for a deeper understanding of their formation and functionality.

Authors' contribution: Igor Aizenberg – conceptualization; methodology, analysis of sources, preparation of literature review or theoretical foundations of research; Alexander Vasko – conceptualization; methodology, software, empirical data collection and validation, empirical research, analysis of sources, preparation of literature review or theoretical foundations of research.

References

Aizenberg, I. (2011). *Complex-Valued Neural Networks with Multi-Valued Neurons* (Vol. 353). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20353-4

Aizenberg, I., Herman, J., & Vasko, A. (2022). A Convolutional Neural Network with Multi-Valued Neurons: A Modified Learning Algorithm and Analysis of Performance. *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference,* 0585–0591. https://doi.org/10.1109/UEMCON54665.2022.9965659

Aizenberg, I., & Moraga, C. (2007). Multilayer Feedforward Neural Network Based on Multi-valued Neurons (MLMVN) and a Backpropagation Learning Algorithm. *Soft Computing*, 11(2), 169–183. https://doi.org/10.1007/s00500-006-0075-5

Aizenberg, I., & Vasko, A. (2020). Convolutional Neural Network with Multi-Valued Neurons. *2020 IEEE Third International Conference on Data Stream Mining & Processing,* 72–77. https://doi.org/10.1109/DSMP47368.2020.9204076

Aizenberg, I., & Vasko, A. (2023). MLMVN as a Frequency Domain Convolutional Neural Network. *2023 International Conference on Computational Science and Computational Intelligence,* 341–347. https://doi.org/10.1109/CSCI62032.2023.00061

Aizenberg, I., & Vasko, A. (2024). Frequency-Domain and Spatial-Domain MLMVN-Based Convolutional Neural Networks. *Algorithms*, 17(8), Article 8. https://doi.org/10.3390/a17080361

Beysolow Ii, T. (2017). Convolutional Neural Networks (CNNs). In T. Beysolow Ii, *Introduction to Deep Learning Using R* (pp. 101–112). Apress. https://doi.org/10.1007/978-1-4842-2734-3_5

Boonsatit, N., Rajendran, S., Lim, C. P., Jirawattanapanit, A., & Mohandas, P. (2022). New Adaptive Finite-Time Cluster Synchronization of Neutral-Type Complex-Valued Coupled Neural Networks with Mixed Time Delays. *Fractal and Fractional*, 6(9), 515. https://doi.org/10.3390/fractalfract6090515

Bruna, J., Chintala, S., LeCun, Y., Piantino, S., Szlam, A., & Tygert, M. (2015). *A mathematical motivation for complex-valued convolutional networks*. https://doi.org/10.48550/ARXIV.1503.03438

Caldeira, M., Martins, P., Cecílio, J., & Furtado, P. (2019). Comparison Study on Convolution Neural Networks (CNNs) vs. Human Visual System (HVS). In S. Kozielski, D. Mrozek, P. Kasprowski, B. Małysiak-Mrozek, & D. Kostrzewa (Eds.), *Beyond Databases, Architectures and Structures. Paving the Road to Smart Data Processing and Analysis* (Vol. 1018, pp. 111–125). Springer International Publishing. https://doi.org/10.1007/978-3-030-19093-4_9

Chatterjee, S., Tummala, P., Speck, O., & Nürnberger, A. (2023). *Complex Network for Complex Problems: A comparative study of CNN and Complex-valued CNN*. https://doi.org/10.48550/ARXIV.2302.04584

Fuchs, A., Rock, J., Toth, M., Meissner, P., & Pernkopf, F. (2021). Complex-valued Convolutional Neural Networks for Enhanced Radar Signal Denoising and Interference Mitigation. *2021 IEEE Radar Conference* (pp. 1–6). https://doi.org/10.1109/RadarConf2147009.2021.9455296

Gad, A. F. (2018). Convolutional Neural Networks. In A. F. Gad, *Practical Computer Vision Applications Using Deep Learning with CNNs* (pp. 183–227). Apress. https://doi.org/10.1007/978-1-4842-4167-7_5

Guberman, N. (2016). *On Complex Valued Convolutional Neural Networks* (arXiv:1602.09046). arXiv. http://arxiv.org/abs/1602.09046

Hirose, A. (2011). Complex-valued Neural Networks. *IEEJ Transactions on Electronics, Information and Systems*, 131(1), 2–8. https://doi.org/10.1541/ieejeiss.131.2

Hongo, S., Isokawa, T., Matsui, N., Nishimura, H., & Kamiura, N. (2020). Constructing Convolutional Neural Networks Based on Quaternion. *2020 International Joint Conference on Neural Networks*, 1–6. https://doi.org/10.1109/IJCNN48605.2020.9207325

Jarrett, K., Kavukcuoglu, K., Ranzato, M. A., & LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? *2009 IEEE 12th International Conference on Computer Vision*, 2146–2153. https://doi.org/10.1109/ICCV.2009.5459469

Kaur, P., & Garg, R. (2020). Towards Convolution Neural Networks (CNNs): A Brief Overview of AI and Deep Learning. In G. Ranganathan, J. Chen, & Á. Rocha (Eds.), *Inventive Communication and Computational Technologies* (Vol. 89, pp. 399–407). Springer Singapore. https://doi.org/10.1007/978-981-15-0146-3_38

Kobayashi, M. (2017). Symmetric Complex-Valued Hopfield Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 28(4), 1011–1015. https://doi.org/10.1109/TNNLS.2016.2518672
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25. https://doi.org/10.1145/3065386
LeCun, Y., Cortes, C., & Burges, C. J. C. (n. d.). *The MNIST Database of handwritten digits.* [Dataset]. Retrieved August 9, 2024, from https://www.kaggle.com/datasets/zalando-research/fashionmnist
LeCun, Y., Fu Jie, Huang, & Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004*, *2*, 97–104. https://doi.org/10.1109/CVPR.2004.1315150
Lin, L., Liang, L., Jin, L., & Chen, W. (2019). Attribute-Aware Convolutional Neural Networks for Facial Beauty Prediction. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 847–853. https://doi.org/10.24963/ijcai.2019/119
Lin, W., Ding, Y., Wei, H.-L., Pan, X., & Zhang, Y. (2020). LdsConv: Learned Depthwise Separable Convolutions by Group Pruning. *Sensors*, 20(15), 4349. https://doi.org/10.3390/s20154349
Nitta, T. (2004). Orthogonality of Decision Boundaries in Complex-Valued Neural Networks. *Neural Computation*, 16(1), 73–97. https://doi.org/10.1162/089976604603400l
Nitta, T. (2008). The uniqueness theorem for complex-valued neural networks with threshold parameters and the redundancy of the parameters. *International Journal of Neural Systems*, 18(02), 123–134. https://doi.org/10.1142/S0129065708001439
Popa, C.-A. (2017). Complex-valued convolutional neural networks for real-valued image classification. *2017 International Joint Conference on Neural Networks* (pp. 816–822). https://doi.org/10.1109/IJCNN.2017.7965936
Rawat, S., Rana, K. P. S., & Kumar, V. (2021). A novel complex-valued convolutional neural network for medical image denoising. *Biomedical Signal Processing and Control*, 69, 102859. https://doi.org/10.1016/j.bspc.2021.102859
Sarabu, A., & Santra, A. K. (2021). Human Action Recognition in Videos using Convolution Long Short-Term Memory Network with Spatio-Temporal Networks. *Emerging Science Journal*, 5(1), 25–33. https://doi.org/10.28991/esj-2021-01254
Sunaga, Y., Natsuaki, R., & Hirose, A. (2020). Similar land-form discovery: Complex absolute-value max pooling in complex-valued convolutional neural networks in interferometric synthetic aperture radar. *2020 International Joint Conference on Neural Networks* (pp. 1–7). https://doi.org/10.1109/IJCNN48605.2020.9207122
Suresh, S., Sundararajan, N., & Savitha, R. (2013). *Supervised Learning with Complex-valued Neural Networks* (Vol. 421). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29491-4
Valle, M. E. (2014). Complex-Valued Recurrent Correlation Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 25(9), 1600–1612. https://doi.org/10.1109/TNNLS.2014.2341013
Venkatesan, R., & Li, B. (2017). Modern and Novel Usages of CNNs. In R. Venkatesan & B. Li, *Convolutional Neural Networks in Visual Computing* (1st ed., pp. 117–146). CRC Press. https://doi.org/10.4324/9781315154282-5
Wu, D., Zhang, J., & Zhao, Q. (2020). A Text Emotion Analysis Method Using the Dual-Channel Convolution Neural Network in Social Networks. *Mathematical Problems in Engineering*, *2020*(1), 6182876. https://doi.org/10.1155/2020/6182876
Xiao, H., Rasul, K., & Vollgraf, R. (2017). *Fashion-MNIST: A Novel Image Dataset for Benchmarking Machine Learning Algorithms* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1708.07747
Xiao, L., Zhang, H., Chen, W., Wang, Y., & Jin, Y. (2018). Transformable Convolutional Neural Network for Text Classification. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 4496–4502. https://doi.org/10.24963/ijcai.2018/625
Yadav, S., & Jerripothula, K. R. (2023). FCCNs: Fully Complex-valued Convolutional Networks using Complex-valued Color Model and Loss Function. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 10655–10664. https://doi.org/10.1109/ICCV51070.2023.00981
Yar, H., Abbas, N., Sadad, T., & Iqbal, S. (2021). Lung Nodule Detection and Classification using 2D and 3D Convolution Neural Networks (CNNs). In L. M. Goyal, T. Saba, A. Rehman, & S. Larabi-Marie-Sainte, *Artificial Intelligence and Internet of Things* (1st ed., pp. 365–386). CRC Press. https://doi.org/10.1201/9781003097204-17
Zhang, Z., Wang, H., Xu, F., & Jin, Y.-Q. (2017). Complex-Valued Convolutional Neural Network and Its Application in Polarimetric SAR Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 7177–7188. https://doi.org/10.1109/TGRS.2017.2743222
Zhang, Z., Wang, Z., Chen, J., & Lin, C. (2022). *Complex-Valued Neural Networks Systems with Time Delay: Stability Analysis and (Anti-)Synchronization Control* (Vol. 4). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-5450-4
Zhu, X., Xu, Y., Xu, H., & Chen, C. (2018). *Quaternion Convolutional Neural Networks*, 631–647. https://doi.org/10.48550/arXiv.1903.00658

Отримано редакцією журналу / Received: 21.01.25
Прорецензовано / Revised: 24.02.25
Схвалено до друку / Accepted: 06.03.25

Ігор АЙЗЕНБЕРГ, д-р філософії, проф.
ORCID ID: 0000-0002-5994-6568
e-mail: igor.aizenberg@manhattan.edu
Манхеттенський університет, Рівердейл, Нью-Йорк, США

Олександр ВАСЬКО, асп.
ORCID ID: 0009-0006-1527-505X
e-mail: oleksandrvasko@uzhnu.edu.ua
ДВНЗ "Ужгородський національний університет", Ужгород, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЗГОРТОК, ОТРИМАНИХ МЕРЕЖАМИ CNNMVN ТА MLMVN, ЯК ЗГОРТКОВОЇ МЕРЕЖІ В ЧАСТОТНІЙ ОБЛАСТІ

Кожен згортковий шар у будь-якій згортковій нейронній мережі створює карту ознак, що містить найважливішу інформацію, необхідну мережі для розпізнавання відповідних зображень. Для подальшого вдосконалення цих нейронних мереж і кращого розуміння їхніх можливостей важливо виявити, які саме ознаки витягуються і як змінюються зображення, що підлягають розпізнаванню, унаслідок згорток, отриманих у процесі навчання.

У пропонованій статті представлено порівняльний аналіз згорток, отриманих за допомогою двох комплекснозначних нейронних мереж, які базуються на багатозначних нейронах. Перша мережа – це згорткова нейронна мережа на основі багатозначних нейронів (CNNMVN), що має традиційну топологію згорткової нейронної мережі, за винятком того, що вона використовує комплекснозначні згорткові ядра у своїй згортковій частині, та багатозначні нейрони у повнозв'язній частині. Друга мережа – це багатозначна нейронна мережа на основі багатозначних нейронів (MLMVN), що є повнозв'язною багатошаровою нейронною мережею, яку використовують як згорткову мережу в частотній області.

Ураховуючи, що обидві нейронні мережі є комплекснозначними, а отримані фільтри працюють у комплексній площині, проведене дослідження показує, що ядра обох мереж формують фільтри, подібні до наявних цифрових фільтрів для оброблення зображень. Аналіз ядер CNNMVN виявив, що вони реалізують фільтри нерізкого маскування та фільтри виявлення країв для ідентифікації форм на зображеннях, тоді як ядра MLMVN підсилюють окремі частотні піддіапазони, що ускладнює їхнє пряме порівняння з просторовими ядрами. Отже, ядра обох згорткових мереж сприяють поліпшенню точності розпізнавання зображень. Візуалізація цих ядер демонструє, що комплекснозначні згорткові нейронні мережі функціонують аналогічно до їхніх дійснозначних аналогів.

Ключові слова: ***згорткові нейромережі, комплекснозначні мережі, багатозначні нейрони, CNNMVN, MLMVN, розпізнавання зображень, частотна область.***

Автори заявляють про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.
The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.