# Incorporating Kernels into Convolutional Neural Networks for Enhanced Feature Extraction

**Sackineh Shamil Jasim** [*1]

[1] Department of Statistics - College of Administration and Economics -University of Karbala-Iraq
Emails: sackineh.sh@uokerbala.edu.iq

**Abstract**

The simulation was used to evaluate the method of kernel K , Neural network NN, Convolution Neural network CNN bys using (MINIST) data set. The accuracy of the method was tested and compared with the convolutional neural network as well as with the kernel function for the same input data (training and testing). The results of simulation showed that there is a high accuracy of the method, and at the same time there is a decreasing loss over the epochs, which indicates the. We note high smooth by method for recognize among features.

**Keywords:** Convolution; MINIST; kernel; neural network; convolution; simulation

## 1. Introduction

Convolutional neural networks are considered a type of deep machine learning that mimics the human brain, while traditional neural networks rely only on input matrices and do not use convolution on data. Convolutional neural networks have been trained on millions of images, but in the case of very large data and complex data, we need a faster time to complete the training. This network takes its idea from the communication between artificial neurons and natural cells of humans or animals, which achieves greater speed and lower error rates than traditional neural networks, as it is very effective in classification. This type of uncertainty has been applied in drug discovery and in identifying potential treatments by predicting interactions between biological molecules and proteins. This type of network was developed and published for the first time through the discovery of handwritten numbers. Since then, this type of network has been used to read postal codes and security numbers.

## 2. Convolutional neural network CNNs

It is termed a filtration neural network, and it is a feed forward neural network, and its function is analogous to the action of biological cells in the visual lobe of the human brain. It's a particular kind of deep learning technique that's mostly used for tasks involving object identification, such segmentation, detection, and classification of images. CNNs provide solutions for numerous computer vision issues in artificial intelligence, including image and video processing, and are used in a wide range of real-world applications, including security camera systems and self-driving cars. The French computer scientist Yann André LeCun (1998) introduced one of the most well-known models with a real-world use [1].



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024

### A. CNNs Structure

The fold or convolution mathematical procedure, which is applied to two functions, f and g, yields the function o, which is a modified copy of the two functions and indicates the overlapping area between them, is the source of the network's name. Image processing is significantly impacted by this process. These layers are the outcome of applying various filters or kernels together with the mathematical convolution or convolution process on the input elements (pixels of the input picture or neurons of the preceding layer). To generate single values, the above convolution procedure is repeated for every set of input elements with a filter size of one. We then create a new feature map for each filter by repeating all of the preceding procedures from the feature maps, and so on. These filters' values are regarded as the network's weights. The following are seen in these layers [2]:

### B. Shared Weights

Applying the filter convolution process again to all of the input parts will result in a lot more connections, but these connections will have shared weights, which will significantly boost efficiency [3]. This is one of the most crucial characteristics of this kind of network as it improves learning efficacy and qualifies the network to provide superior generalization outcomes. As an example, every resultant neuron in the feature map will be linked to the number of filter elements; however, the weights of these connections will match the weights of the neurons in the same feature map that are next to it. Since it joins neighboring points and enables the representation of regions, this approach will also guarantee the identification of image characteristics based on where they are located within the fundamental picture [4].

### C. Duplicate filters

Different sets of features in the image can be identified by using multiple convolutional filters with the same input because each will produce a unique feature map. However, the number of filters should be chosen so as to take into account the complexity of the calculations involved, as this depends on the quantity of training examples that are available. While a single filter's size, or its dimensions, is mostly determined by the data in the collection of training examples, the task, contrast, and dimensions of the input picture all play a role [5].

### D. Subsampling layers

- In the network architecture, these levels are optional, which means that their presence is not necessary. In the event that they exist, they will follow each convolutional layer and work to decrease the number of samples or neurons by reducing each set of input neurons of a certain size to a single neuron. The network architecture determines its size, and its value is The best size is $2 \times 2$, since increasing it might result in information loss. Reduction can be accomplished in a number of methods, such as [6]:
- Maximum pooling value: takes the highest value among them.
- Average pooling: takes the average of all values.
- Fully connected layers:
- After numerous layers of the previous two kinds, these layers come to link all the neurons of the preceding layer (regardless of their type) and make them input to every neuron in it, like in ordinary neural networks. They don't have to be of a certain number, but being the last layers in the network, they often have two consecutive layers since they can't appear before a convolutional layer [7].



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024

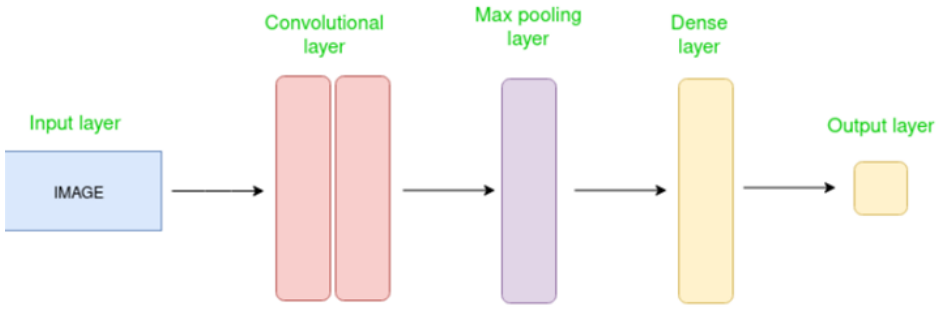


Figure 1: Structural of convolution neural net work

Now we take a little portion of this picture and represent it vertically using a small neural network called a filter or kernel, for example, with K outputs. Upon running the neural network over the whole picture, we will get an alternate image that varies in width, height, and depth. Currently, we possess additional channels that are more slender and taller in comparison to the original R, G, and B channels. This process is often known as convolution. A neural network is considered normal when the patch size and image size are the same [8].

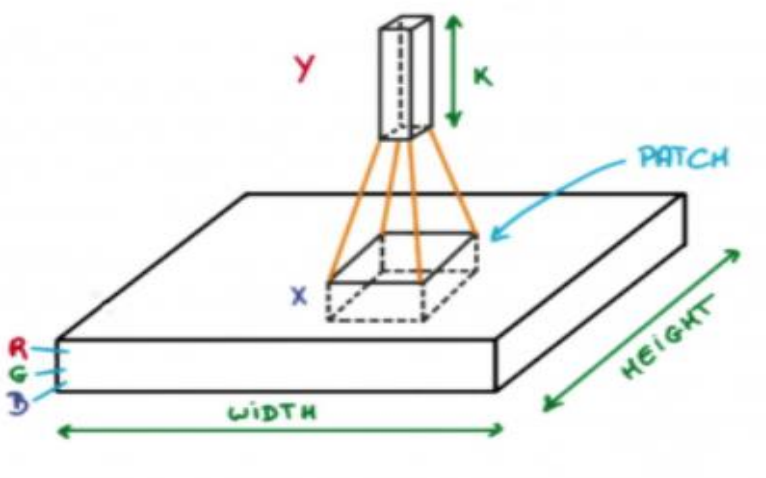


Figure 2: A perpendicular function on an CNN

### 3. Convolutional neural networks (CNNs) [9]

This type of network was used for the purpose of distinguishing patterns and features in images and finding the characteristics of the image with high accuracy, which depends on convolutional hidden layers through which accurate discrimination is carried out according to the following steps [10]:

### A. Method of convolution

It is the method that is used to show the basic features in the data by convolution and activating the role of filters or kernel functions on the input data, according to the following mathematical formula [11]:

$$K(\alpha: \beta).(n+m, j+Z(i*\sum_{\alpha}.\sum_{\beta} = K(i,j)*Z) \qquad \ldots(1)$$

Input matrix (image), K is kernel function, (i,j) the point in the input matrix, α, β are kernel inputs.

### B. Activation Layer [12]:

Then we apply, the activation function to the model to add nonlinearity. The Gaussian, function is the most often used.



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024

$$\boldsymbol{f(x) = max(0, x)} \quad \ldots (2)$$

This function keeps positive values as they are and zeros out negative values.

**C. Pooling Layer** [12]**:**

The pooling layer reduces the dimensions of the matrix generated by the convolution layer, which reduces computational complexity and helps avoid overfitting. There are two basic types of aggregation:

- **Max Pooling**

$$(\boldsymbol{\alpha + \beta j + \underset{R \in (\alpha, \beta)}{Max} Z\left(i = P(i, j)\right)} \quad \ldots (3)$$

- **Average Pooling**

$$(\boldsymbol{\alpha} + \beta \boldsymbol{j} + \boldsymbol{Z}\left(\underset{R \in (\alpha, \beta)}{i} \sum \frac{\mathbf{1}}{|\mathbf{R}|} = \boldsymbol{P(i, j)} \quad \ldots (4)$$

**3.4. Fully Connected Layers:**

Finally used to produce the final forecast. Each node in this layer is connected to every node in the previous layer. The mathematical equation [13]:

$$\boldsymbol{y = W \cdot x + by} \quad \ldots (5)$$

Y output, W is a weight matrix, X is the input matrix, B is the biased vector.

**4. General application of the mathematical model** [14]**:**

A convolutional neural network consists of several successive layers of convolution, activation, and pooling, followed by fully connected layers. The final mathematical model can be summarized by the following formula:

$$\boldsymbol{F(x) = fL\left(WL \cdot \left(\ldots f2\left(W2 \cdot \left(f1(W1 \cdot x + b1)\right) + b2\right) \ldots\right) + bL\right)} \ldots (6)$$

x Initial input, W and b are the weights and biases of each layer , f is the activation function , L is the number of layers in the network.

**5. Kernels in CNN**

Filters, also known as small matrices, are utilized in the convolution process. As the kernels move through the input data, they perform multiplication with the corresponding pixels, which produces a feature map that highlights specific patterns or traits of the input. Features in Convolutional Neural Networks (CNNs) refer to the important patterns or unique characteristics obtained from the input data through the convolutional layers. The mathematical modeling for the argument the Kernel in CNN as following:

$$(Z * K)(x, y) = \sum_{i=1}^{\beta-1} \sum_{j=1}^{\alpha-1} Z(x + i, y + j) K(ji, j) \quad (1)$$

Z input image, k kernel, and x, y position of kernel on image, m, and n kernel dimensions.



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024

## 6. Simulation study

The famous data (MINIST), which means (Modified National Institute of Standards and Technology database), was used to train and test the K (Kernel), NN (Neural Network), KCNN((Convolution Neural Network)) method. It is one of the most widely used data in research and practical applications, which consists of an input image that includes (60,000) training images, (10,000) test images, and all image consisting of a number written by hand from (0) to (9). The image size is 28x28. It also consists of image labels, which express the number written in the image from (0) to (9). A kernel function with a size of (3x3) and (8) feature maps were used for the purpose of teaching the network on it. The accuracy of the method was tested and compared with the convolutional neural network as well as with the kernel function for the same input data (training and testing), as shown in Table (1). The algorithm iteration was (600).

Table 1: Accuracy of the methods

| Method | Kernel | NN | CNN |
|---|---|---|---|
| Accuracy | 0.89 | 0.93 | 0.98 |

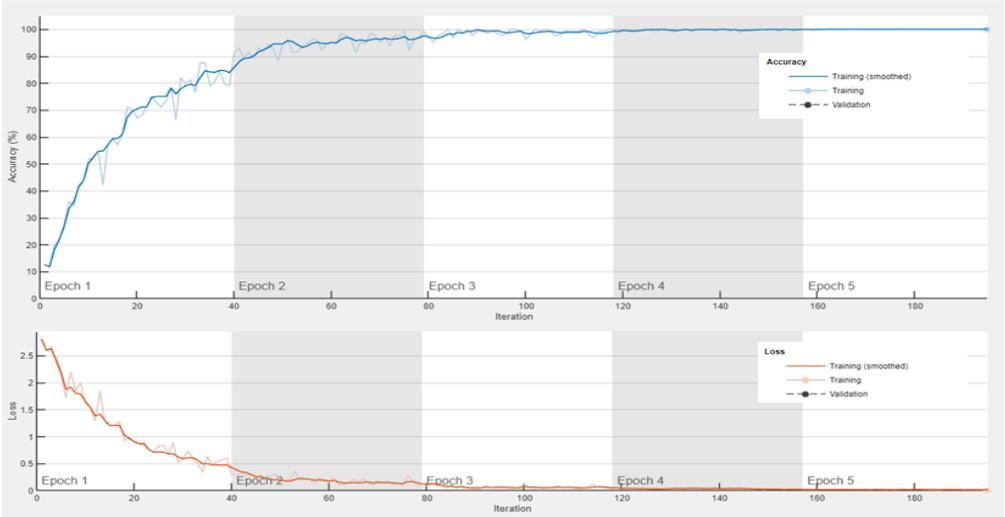


Figure 3: Simulated training method and accuracy, Loss in method.

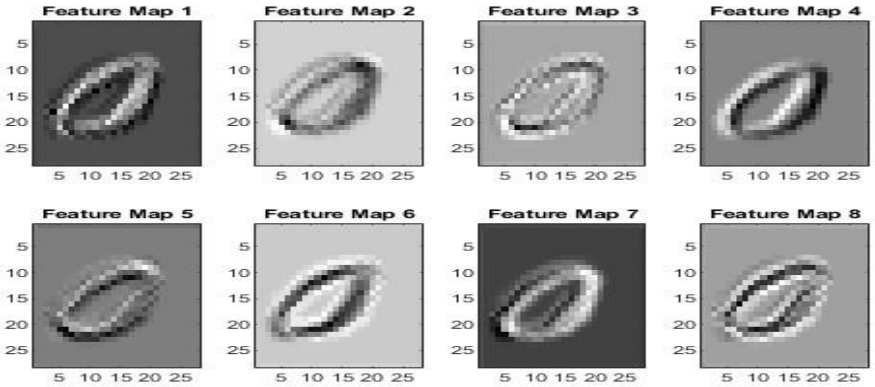


Figure 4: Output of the convolutional operation on the input images



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024

## 7. Result Discussion

We can see from table (1) there are high accuracy for CNN is (0.98) compared with K & NN, Figure (3) that there is a high accuracy of the method after the training data set is completely passed through the neural network at once over the epochs (Epoch), and at the same time there is a decreasing loss over the epochs, which indicates the accuracy of the proposed method. From figure (4) we note high smooth by method for recognize among features.

## References

[1] Wu J. Introduction to convolutional neural networks. Natl Key Lab Nov Softw Technol Nanjing Univ China 2017;5:495.

[2] O'shea K, Nash R. An introduction to convolutional neural networks. ArXiv Prepr ArXiv151108458 2015.

[3] Anuj L, Gopalakrishna MT. Resnet50-yolov2-convolutional neural network based hybrid deep structural learning for moving vehicle tracking under occlusion. Solid State Technol 2020;63:3237–58.

[4] Al-Haija QA, Smadi M, Al-Bataineh OM. Identifying phasic dopamine releases using darknet-19 convolutional neural network. 2021 IEEE Int. IoT, Electron. Mechatronics Conf., IEEE; 2021, p. 1–5.

[5] Bhatt D, Patel C, Talsania H, Patel J, Vaghela R, Pandya S, et al. CNN variants for computer vision: History, architecture, application, challenges and future scope. Electronics 2021;10:2470.

[6] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. ArXiv Prepr ArXiv170404861 2017.

[7] Zhao X, Wang L, Zhang Y, Han X, Deveci M, Parmar M. A review of convolutional neural networks in computer vision. Artif Intell Rev 2024;57:99.

[8] Ji X, Yan Q, Huang D, Wu B, Xu X, Zhang A, et al. Filtered selective search and evenly distributed convolutional neural networks for casting defects recognition. J Mater Process Technol 2021;292:117064.

[9] Kuo C-CJ. Understanding convolutional neural networks with a mathematical model. J Vis Commun Image Represent 2016;41:406–13.

[10] Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Trans Neural Networks Learn Syst 2021;33:6999–7019.

[11] Sainath TN, Kingsbury B, Mohamed A, Dahl GE, Saon G, Soltau H, et al. Improvements to deep convolutional neural networks for LVCSR. 2013 IEEE Work. Autom. speech Recognit. Underst., IEEE; 2013, p. 315–20.

[12] Shetty S. Application of convolutional neural network for image classification on Pascal VOC challenge 2012 dataset. ArXiv Prepr ArXiv160703785 2016.

[13] Shi X, Chen Z, Wang H, Yeung D-Y, Wong W-K, Woo W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Adv Neural Inf Process Syst 2015;28.

[14] Wiatowski T, Bölcskei H. A mathematical theory of deep convolutional neural networks for feature extraction. IEEE Trans Inf Theory 2017;64:1845–66.



Doi: https://doi.org/10.54216/PMTCS.040102
Received: December 13, 2023 Revised: March 17, 2024 Accepted: May 24, 2024