Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

Improving Traffic Sign Recognition by Using Wavelet Convolutional Neural Network

Hamidreza Emami[1], Ramin Shaghaghi Kandowan[*2], and Seyyed Abolfazl Hosseini[3]

1 Department of Communication, Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran.
Email: emami.hamidr@gmail.com
2 Department of Communication, Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran.
*Corresponding author: Ramin Shaghaghi Kandowan; ramin_shaghaghi@yahoo.com
3 Department of Communication, Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran.
Email: universizen@yahoo.com

Abstract:

Traffic sign recognition (TSR) considered as a challenging subject in image processing for many years. Nowadays, after achievements in processing power of processors and easily accessible datasets, many researches has been done by using convolutional neural networks (CNN) In many applications including TSR. CNN is a popular deep learning method that has a reasonable functionality in image classification and pattern recognition. Important factors in performance of a CNN can be written as follows: accuracy, efficiency and the precision. Therefore, in this paper we try to achieve better results in these important factors. In this particular usage, we must consider complexity and processing time of whole procedure to be in an acceptable range. As will be shown in the following sections, because of wavelet characteristics, we will use wavelet in CNN to improve performance of it in two parts of its structure: In first step, we use wavelet convolutional neural network (wCNN) instead of CNN. Therefore, wavelet transform function is replaced in the convolutional layers of CNN. In next step, wavelet convolutional wavelet neural network (wCwNN) is suggested, so that fully connected neural network (FCNN) of wCNN and CNN is changed by wavelet neural network (wNN). So, we can compare results of these methods (wCNN and wCwNN) with CNN. We can obviously see that, accuracy results is about 4 percent better than CNN and the mean square error and the rate of error are decreased. These results achieved by increasing the calculations of the CNN algorithm.

***Keywords**: Wavelet convolutional neural network, traffic sign recognition, wavelet neural network, smart cars, deep learning.*

1. Introduction

Traffic sign recognition (TSR) is a subset of advanced driver assistant systems (ADAS). In recent years, it has become more important for smart car systems and can help drivers for a

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

safer and simpler driving. It has engaged many researchers for years by using different methods from simple machine learning technics [1] to deep learning ones. Because of recent developments and achievements in processing power of processing unit hardware and easily accessible datasets, deep learning methods are more practically usable and riches better results in accuracy and processing time. Then, these methods are widely used in many usages like TSRs. Also, CNN is a common model of deep learning methods which is based on convolutional features of image [2]. It is highly used in detection and classification [3]. CNN can solve difficult issues which are hard for other machine learning algorithms. Especially in the classification and object recognition, CNN is doing separation of labels very powerful with very low error rate. The neuron weights [4] of CNN are modified and adjusted step by step in forward propagation and error back propagation [5]. Nowadays, the performance of CNNs becomes better and better because of new processors and computing technologies like cloud computing and distributed computing resulting that computing power has been greatly improved. In addition to pattern recognition and image processing [6], CNN are also applied in the other fields [7] such as text classifications [8], control systems [9] and object tracking [10].

Indeed, CNN in image classification can work as a non-linear transform of an image. Usually, capability of getting more input data can be good for better adaptation of CNN and offering better quality results by using more spatial data. In general, input data length can be expanded by increasing network layers or filter size or pooling algorithm. Because of computational cost of working on increasing filter size or network layers, pooling can large input data and lead to better results and efficiency by increasing spatial resolution of feature map. Also, it can result in losing some parts of data. At recent researches, dilated filtering [11] by using zero holes in convolutional filtering can balance efficiency and input data size. Also, dilated filtering has their own disadvantages and restrictions. Recently, a better CNN based algorithm proposed [12] as multi-level wavelet CNN (MWCNN) by using discrete wavelet transform (DWT) instead of common pooling. Because of benefits of DWT, image data and intermediate features are preserved by the presented method. In addition, frequency and location data of feature maps are considered by DWT [13], [14], that is useful for keeping detailed texture when using multi-frequency features. In other ways, The Multi-Path Learnable Wavelet Neural Network for Image Classification was proposed by De Silva et al [15]. This schema proposes a multi-path algorithm with multi levels of wavelet decompositions. In the part of prediction, a convolutional LSTM network using the wavelet decomposition has been introduced in 2018 [16]. It consider the wavelet decomposition as the procedure of feature extraction instead of common feature extraction methods, which has been also proposed by Kiskin et al. in 2017 [17].

Power System Technology
ISSN:1000-3673
Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

Using wavelet analysis has some benefits for us and has been extensively used in signal and image processing and analysis. Wavelet analysis solution preserved as a strong tool for looking on details of sound, image, etc. Also, wavelet transformation has some complications [18]. In addition, the handy detail extraction capability of wavelet transformation is effective and significant to solve some issues of CNN [19]. In these paper, the purpose of investigations is to answer the CNN's shortcomings by using the benefits of the WT. Therefore, the significance of the investigations is that the enhancements of CNN neurons are considered. Other than the capability of network with deeper layers, it is expected that the enhancement in operation of each neuron of CNN can lead to better features extraction and feature identification that lead to increase learning power of the CNN. In this paper we implement wavelet analysis to improve the CNN network results for traffic sign recognition.

Main parts of this paper are in two step: (1) An existing reasonable CNN method for TSR is described (2) The wavelet Convolutional Neural Network (wCNN) used, where the wavelet transformation is used as the activation function in Convolutional layers of CNN. (3) In addition of wCNN, the wavelet Convolutional wavelet Neural Network (wCwNN) is used, where the FCNN part of wCNN is changed by wavelet Neural Network (wNN). (4) Comparing examinations among CNN, wCNN and wCwNN that executed on GTSRB [20] dataset.

2. Objectives

Most of usual image processing and machine learning methods are very sensitive toward weather and light conditions of input image and even angle of the camera, causing difficulties in detection process and limit any specific pattern at real outdoor conditions [21]. Moreover, most of the latest methods rely of intrinsic defaults about the traffic signs like their color, shape or texture [22].

In [1] researchers have assumed that all important traffic signs are in a red color and used enhanced color filtering to detect and locate any possible region of interest (ROI) in the image but clearly this won't work on all signs and not for textures in signs.

Researchers at [23] have recommended a more general method and proposed to define some templates for the spatial pattern of traffic signs, moving it over the image to find the parts with highest similarities. This methodology is really slow and requires some very complex hardware to run. The same idea of color segmentation with the LAB and HIS color space is practiced in [1] and [24] with similar results and disadvantages like in [21]. Many other algorithms and methods have been used like K-Nearest Neighbor (KNN) [25], Support Vector Machine (SVM) [1] and neural network [26] for traffic signs recognition.

At present, Convolutional networks are gently replaced traditional and older computer vision methods for various usages such as object detection and classification, image processing and pattern recognition. It is implemented for processing, recognizing and the learning of depth

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

description of the traffic signs. This method solves the problem of descriptors extraction which has very dependencies to different aspects. This network receipts 2D image and apply convolution operations to it. It has the capability to learn a representative schema of image.

2.1. Convolutional neural network

The schematic of a sample CNN for TSR is showed in Fig. 1. It contains two major sections: the first section is CPNN, and the next section is FCNN. In CPNN, we start with an input layer, and the next layers are some convolutional and pooling layers and dropout layers. In next part where we have FCNN, we have an input flatten layer, and the next layer of FCNN is an output dense layer. The relations and details of CPNN and FCNN parts are as follows:

(1) The first layer of CPNN is the first layer of CNN.

(2) The final layer of CPNN is the first layer of FCNN.

(3) The output layer of FCNN is the output layer of CNN.

(4) The activation of the convolutional layer in CPNN is ReLU and the output layer in FCNN is using softmax.

(5) There are not any activation functions in the pooling layer of CPNN and the input layer of FCNN

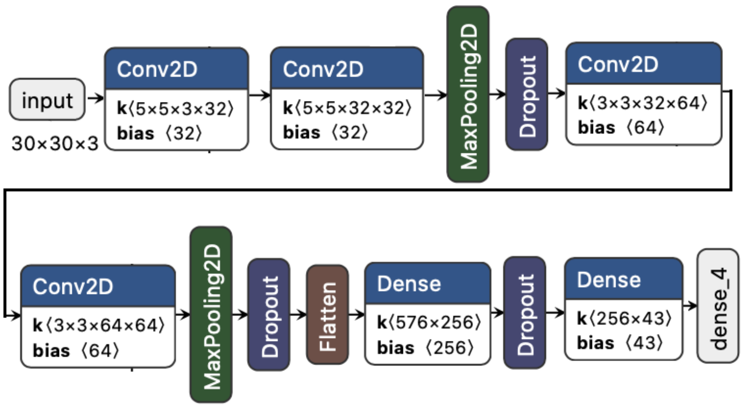

Fig. 1. Schematic of a sample CNN for TSR.

2.1.1. Structure of CNN

Main parts of CNN can be explained as:

(1) Adjusting weights and bias in neurons of layers.

(2) Using forward propagation method.

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

(3) According to the loss function, determining the mean square error (MSE).

(4) Determining the errors of back propagating for each layer, which are the results of derivation by the chain rule.

(5) Adjust the weights and bias according to the back-propagated errors by applying gradient.

(6) Doing the step (2) to step (5) until the MSE become smaller than defined value.

(7) Evaluating the precision, accuracy and efficiency.

2.2. Wavelet transform

Wavelet transform (WT) known as a powerful method to obtain features of signals and images. WT presents a Time-Frequency sight which can offer higher and lower resolutions of details of signals and images. The imperfection of Fourier Transform (FT) [27] is that the window size cannot be changed when the frequency is changed. This problem can be solved by WT. The $\psi(x, y)$ is called wavelet generating function, which can be expressed as $\psi(x, y) = \frac{1}{\sqrt{x}} \int_{-\infty}^{\infty} f(t)*\varphi\left(\frac{t-y}{x}\right)$, where y and x are the scale parameters which control the translation and extension of function.

$\Psi(x, y)$ is designed according to the following conditions:

Just in a very small part, the value of function is not 0, and in the other parts are 0. It can be said that, considering the signal in timeline is equal to adding a window on the original signal.

The integral value of function in the horizontal axis must be 0.

The procedure must be reversible.

In Fig. 2 from [28] we can see a sample of wavelet transform and how it working. The error is the amount of difference between the primary signal and the signal of wavelet inverse transformation. The parameter of wavelet function that controls the implementation of wavelet function is named scale. With different scale parameters we can change the wavelet transform's ability of data extraction to primary signal.

In brief, by changing the value of scale and translation, wavelet transform can learn the different features. So, stronger features can be used by implementing the wavelet transformation in CNN.

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

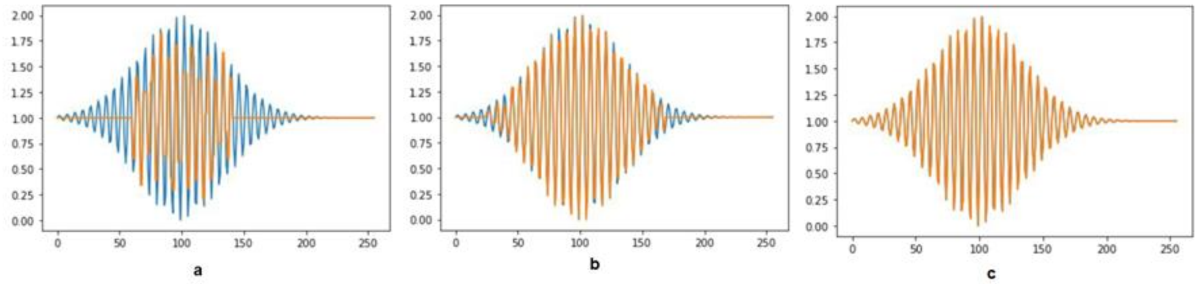

Fig. 2. Various effects of WT: a. scale= 20, error=2.0 b. scale=100, error=0.04 c. scale=200, error=0.0005

2.3. Wavelet convolutional neural network

In this novel method, the activation of the convolutional layer in CNN was F that is changed by the Ψ. The F in CNN was ReLU function, and the Ψ () of wCNN is wavelet scale transformation function.

In the first section of suggested wCNN method we offer Wavelet Convolutional Pooling Neural Network (wCPNN), and in the last section we implement FCNN. The schematic of wCNN is presented in Fig. 3.

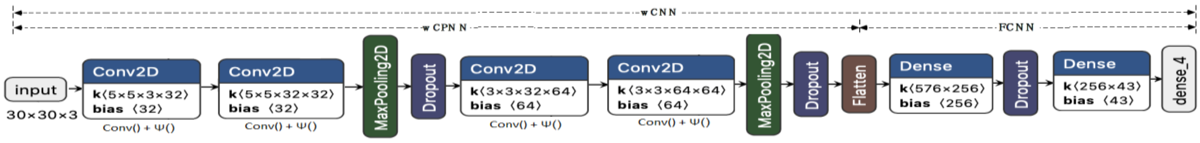

Fig. 3. Schematic of a sample wCNN for TSR.

3. Methods

Considering the limitations and vulnerabilities of the previous trials, here at the current work, according to wCNN, in the last section of it, the activation function of FCNN is changed by a wNN. The wCwNN consists of wavelet Convolutional Pooling Neural Network (wCPNN) and wNN. We test all condition for activation functions and compare results in every situation. After that, and with various number of epochs in learning procedure, the benefits of using wavelet functions were clearly seen. Then, we propose that all activation functions of convolutional layers of wCPNN and the hidden layer of the wNN, must be Ψ to achieve better accuracy and lower loss. The schematic of wCwNN is showed as Fig. 4.

Power System Technology
ISSN:1000-3673
Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

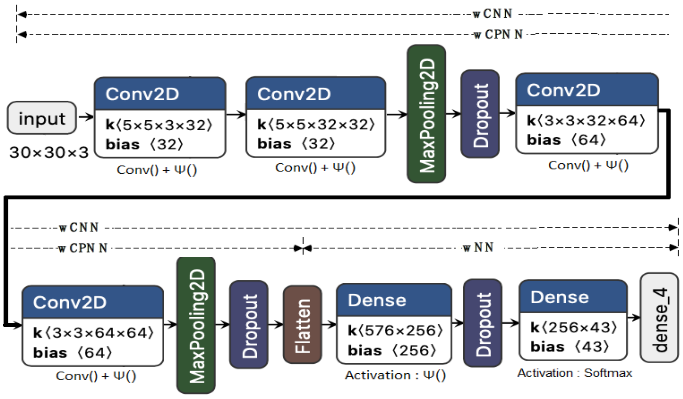

Fig. 4. Schematic of a sample wCwNN for TSR

The initial section of wCwNN is wCPNN that is similar to wCNN. The next section of wCwNN is wNN, which is different from wCNN that was FCNN. In wNN, we can use sigmoid function instead of ReLU in activation function of FCNN layer or we can use H hidden layer wavelet function. Therefore, implementation of wNN, in the following details are showed:

(1) The output layer of wNN is defined as $l = -1$. Then, number of neurons in this output layer=size − 1.

(2) Hidden layers of wNN is defined as $l = -2$. Then, number of neurons in hidden layers=size-2.

(3) The third to last layer (input layer) of wNN is defined as $l = -3$. Then, neurons in this input layer=size−3.

(4) Last layer of wCPNN (the previous layer of the input layer of wNN) is expressed as $l = -4$. Then, neurons in the output layer of wCPNN= (size−4) × (size−4).

4. Results

Dataset

Numerous freely accessible traffic sign datasets have been collected all around the word such as Belgium (Timofte et al., 2011), China (Zhu et al., 2016), Germany (Stallkamp et al., 2011) and Sweden (Larsson & Felsberg, 2011).In this paper we choose German traffic sign recognition benchmark (GTSRB) dataset. In fig.5, some sample images of GTSRB is displayed. There are many causes for selecting this dataset, such as it is extremely used in papers and is implemented for comparison to other method of traffic sign recognition results. Currently, a GTSRB website is continued and registering results are accepted, handled and shown. Such ranking aids to realize the better methods are employed for the traffic sign classification. GTSRB dataset contains different resolution traffic sign examples that produced from one second video sample. These images belongs to one of the 43 classes of these dataset.

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

Its ground truth information is trustworthy because of its semi-automatic annotation, the training set has 39,209 images, and the validation set contains 12,630 images, which are used to determine the performance of the methods. These images are RGB with pixel sizes from 15 × 15 up to 250 × 250. Through a simple initialization procedure, all the input images pixel sizes become 30 × 30 pixels.

Fig. 5. Some sample images of GTSRB datasets

4.2. Training and testing

We use Keras and Tensorflow on Python 3.8 by working on GTSRB dataset in a core i5 CPU with 8GB of ram and Geforcemx110 GPU. Source codes are available at [29]. Used a simple CNN model and after 21 epochs we have 95.81% accuracy as showed in fig. 6 with relevant training and validation losses. Then, we test sigmoid activation function instead of ReLU in different possible positions separately and together. Simulations shows that for state one that we change first convolution layer activation function to sigmoid we can reach to accuracy 95.88% in 15 epochs. In state two, when we change activation functions of first two convolution layer by sigmoid we reach to accuracy of 96.96% after 18 epoch. Next, we implement sigmoid function in activation functions of three first convolution layer sets and reach to accuracy of 96.8% after 29 epochs. Then, we use it for four layer sets and we reach to accuracy 97.6% after 34 epochs. Result of accuracy showed in table 1. In second part of system, when we use sigmoid activation function we reach to 98.67% accuracy after 32 epochs. Final results showed in fig. 6.

TABLE 1. Loss and accuracy of simulations

Model	**Epochs**	**Accuracy**
CNN	21	94.82%
wCNN (1+4)	15	95.88%

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

wCNN (2+3)	18	96.96%
wCNN (3+2)	29	96.80%
wCNN (4+1)	34	97.60%
wCwNN	32	98.67%

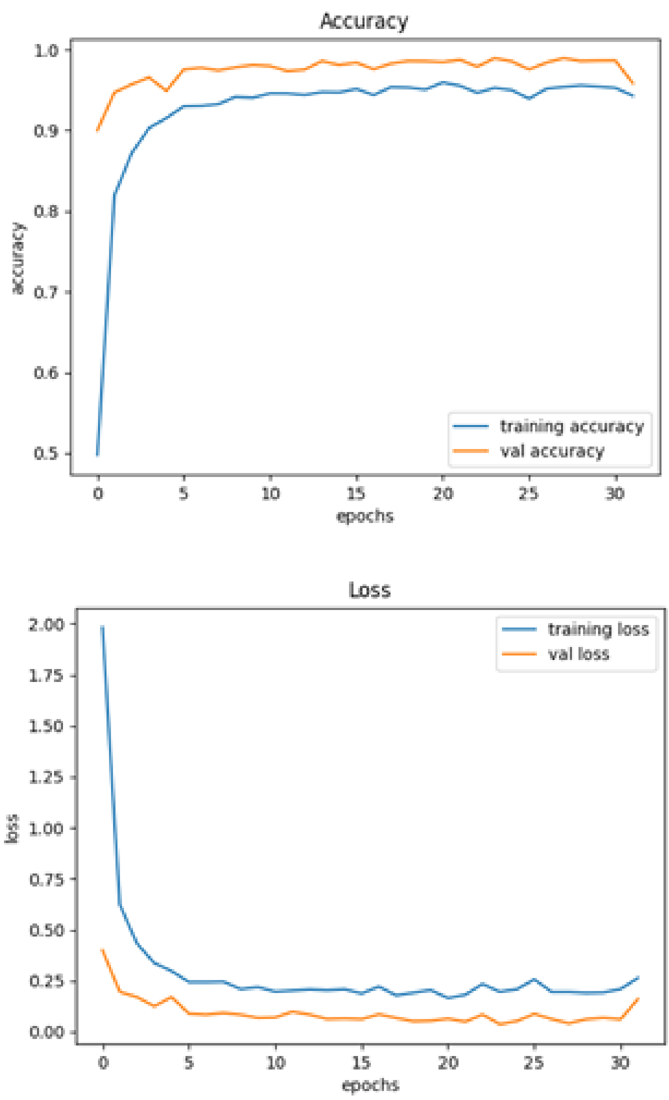

Fig. 6. Loss and Accuracy of CNN model for TSR

Power System Technology
ISSN:1000-3673
Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

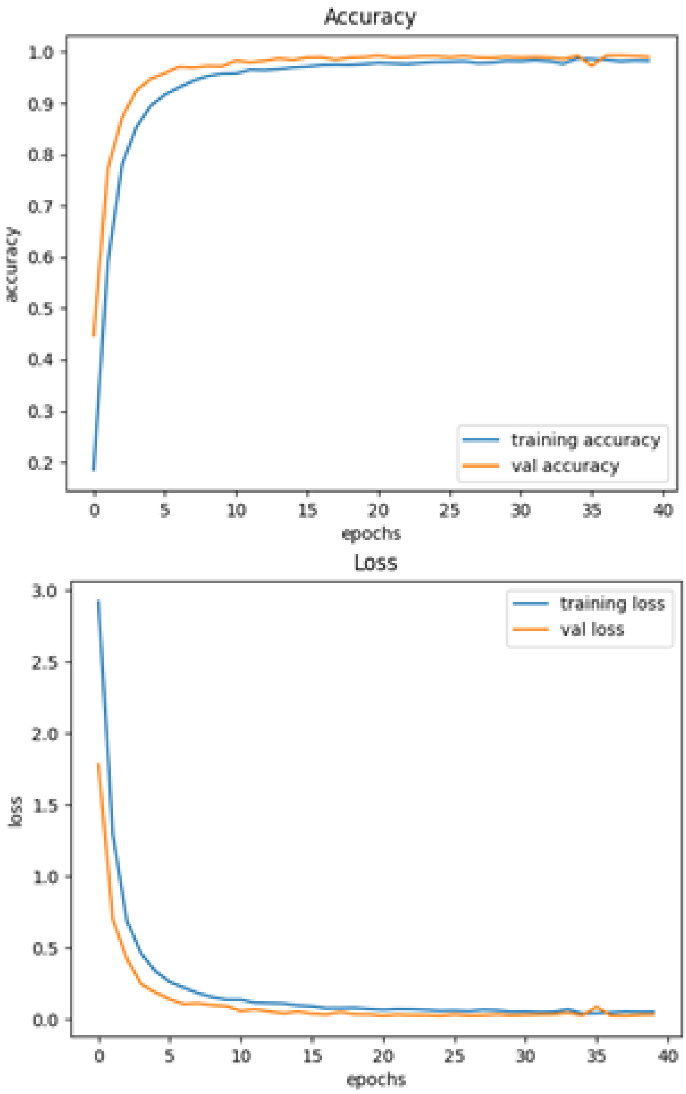

Fig. 7. Loss and Accuracy of wCwNN model for TSR

Also, compared to latest articles [30] [31] [32], better results have been obtained, which is shown in Table 2 comparing the latest results obtained on GTSRB.

TABLE 2. Comparison with the latest accuracy results

Model	Epochs	Accuracy (%)
CNN	50	98.44
Multi-Scale CNN		98.31
YOLOv5	200	97.7
wCwNN	32	98.67

Power System Technology
ISSN:1000-3673
Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

5. Discussion

As we seen before, in different states of using sigmoid function in activation part of convolution layers instead of ReLU we reached different quality accuracy and loss. Best condition is using sigmoid activation function in four first convolution layers and one convolution layer of second section instead of ReLU. In generals, it was predictable that using wavelet and sigmoid activation function in CNN will have better results because of benefits of using better resolution by wavelet. Furthermore, by using wavelet we can separate useful and useless data in a more accurate manner. Finally, by using wavelet in these model we reach to about 4 percent better accuracy than primary ones by continuing learning procedure for 11 additional epochs.

Refrences

[1] H. Emami, R. Shaghaghi Kandowan and S. A. Hosseini, "Centroid Distance Shape Recognition for Real Time Low Complexity Traffic Sign Recognition", Majlesi Journal of Telecommunication Devices, Vol.9, No. 4, pp. 157-160, 2020.

[2] Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. Neural Comput 18(7):1527–1554.

[3] LeCun Y, Bengio Y, Hinton G (1988) Deep learning. Nature 521(7553), 436–444 (2015) 15.

[4] LeCun, Y., Touresky, D., Hinton, G., Sejnowski, T.: a theoretical framework for backpropagation. In: proceedings of the 1988 connectionist models summer school, vol. 1, pp. 21–28. CMU, Pittsburgh, Pa: Morgan Kaufmann.

[5] Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. Liu J (2014) research of adaptive wavelet neural network (awnn) and ann based control system intelligent applications.

[6] Chang J, Sitzmann V, Dun X, Heidrich W, Wetzstein G (2018) Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. Sci Rep 8(1):1–10.

[7] Khan A, Sohail A, Zahoora U, Qureshi AS (2019) A survey of the recent architectures of deep convolutional neural networks. arXiv preprint arXiv:1901.06032.

[8] Sifuzzaman M, Islam M, Ali M (2009) Application of wavelet transform and its advantages compared to fourier transform.

[9] Bateux Q, Marchand E, Leitner J, Chaumette F, Corke P (2017) Visual servoing from deep neural networks. arXiv preprint arXiv: 1705.08940.

[10] Pati YC, Krishnaprasad PS (1993) Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations. IEEE Trans Neural Netw 4(1):73–85.

[11] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, arXiv: 1511.07122. [Online]. Available: https://arxiv.org/abs/1511.07122.

Power System Technology
ISSN:1000-3673

Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

[12] Pengju Liu, Hongzhi Zhang, Wei Lian, Wangmeng Zuo “Multi-Level Wavelet Convolutional Neural Networks”, vol.7, 2019.

[13] I. Daubechies, ``The wavelet transform, time-frequency localization and signal analysis,'' IEEE Trans. Inf. Theory, vol. 36, no. 5, pp. 9611005, Sep. 1990.

[14] I. Daubechies, Ten Lectures on Wavelets. Philadelphia, PA, USA: SIAM, 1992.

[15] De Silva D, Vithanage H, Fernando K, Piyatilake I (2020) Multipath learnable wavelet neural network for image classification. In: Twelfth International Conference on Machine Vision (ICMV 2019), vol. 11433, p. 114331O. International Society for Optics and Photonics Res improv wavelet convolutional wavelet neural netw 35.

[16] Waibel A, Hanazawa T, Hinton G, Shikano K, Lang KJ (1989) Phoneme recognition using time-delay neural networks. IEEE Trans Acoust Speech Signal Process 37(3):328–339.

[17] Kiskin I, Orozco BP, Windebank T, Zilli D, Sinka M, Willis K, Roberts S (2017) Mosquito detection with neural networks: the buzz of deep learning. arXiv preprint arXiv:1705.05180.

[18] Zhang Q (1997) Using wavelet network in nonparametric estimation. IEEE Trans. Neural Network 8(2):227–236.

[19] McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5(4):115–133.

[20] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition, Neural Networks, Available online 20 February 2012, ISSN 0893-6080, 10.1016/j.neunet.2012.02.016. (http://www.sciencedirect.com/science/article/pii/S0893608012000457).

[21] Wali, S.B, "Vision-based traffic sign detection and recognition systems: Current trends and challenges", Sensors 2019, 19, 2093.

[22] Wang, G.Y.; Ren, G.H.; Jiang, L.H.; Quan, T.F., "Hole-based traffic sign detection method for traffic signs with red rim", journal of Visual Computations, 2014, 30, 539–551.

[23] Hechri, A.; Hmida, R.; Mtibaa, A., "Robust road lanes and traffic signs recognition for driver assistance system", International Journal of Computers, Science and Engineering. 2015, 10, 202–209.

[24] Lillo-Castellano, J.M.; Mora-Jiménez, I.; Figuera-Pozuelo, C.; Rojo-Álvarez, J.L., "Traffic sign segmentation and classification using statistical learning methods", journal of Neurocomputing 2015, 153, 286–299.

[25] Y. Han, K. Virupakshappa, E. Vitor, S. Pinto and E. Oruklu, “Hardware/Software Co-Design of a Traffic Sign Recognition System Using Zynq FPGAs,”, In Electronics journal, 2015, Vol. 4, p. 1062-1089; doi:10.3390/electronics4041062.

Power System Technology
ISSN:1000-3673
Received: 10-10-2023 Revised: 02-12-2023 Accepted: 15-12-2023

[26] L. Abdi, “Deep learning traffic sign detection, recognition and augmentation,” Proceedings of the Symposium on Applied Computing, Maroc, 2017, p. 131-136.
[27] Savareh BA, Emami H, Hajiabadi M, Azimi SM, Ghafoori M (2019) Wavelet-enhanced convolutional neural network: a new idea in a deep learning paradigm. Biomed Engin Biomedizinische Technik 64(2):195–205.
[28] Liu, JW. Zuo, FL., Guo, YX. Et al. Research on improved wavelet convolutional wavelet neural networks. Appl Intell 51, 4106–4126 (2021).
[29] https://www.github.com/hamidre1/Deep_TSR/
[30] Vincent M.A., Vidya K., Mathew S.P. Traffic sign classification using deep neural network; Proceedings of the 2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS); Thiruvananthapuram, India. 3–5 December 2020; pp. 13–17.
[31] Gökberk A., Durdu A., Nesimioğlu B.S. Accuracy Comparison of CNN Networks on GTSRB Dataset. J. Artif. Intell. Data Sci. 2022; 2:63–68.
[32] Zhu Y., Yan W.Q. Traffic sign recognition based on deep learning. Multimed. Tools Appl. 2022; 81:17779–17791. Doi: 10.1007/s11042-022-12163-0.