IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

HAND WRITTEN DIGIT RECOGNITION USING CONVOUTION NEURAL NETWORK

[1]Ms. Punya M.P,[2]Ms. Baby,[3]Prof. Lokesh B

[1]Student,[2]Student,[3]Head of the Department

[1]Department of Electrical and Electronics Engineering

[1]Srinivas Institute of Technology, Valachil, Mangaluru, Karnataka, India

Abstract: Digit recognition is a suitable model issue for learning about the Neural Networks, and it will pave the way for sophisticated Deep Learning techniques.Manymethods are found in literature to recognize and classify the digits which are written by hand (handwritten digits).ThispaperexplainshowtorecognizeandclassifyhandwrittendigitsusingCNNandMATLABforenhancedperformance. A deep convolutional neural network (CNN) is a type of neural network that is used to recognise images. This method determines how important the deep layer improvements are for processingthe image byusing MATLAB, we provide the implementationof necessary for constructing and applying Convolution Neural Network to a high-quality data set known as MNIST which is a collection of more than 60,000 handwritten digits dataset for training purpose and 10,000 digitsdatasetfortesting purpose.Whentestedthedevelopedmodelforclassification, we got 99.60% accuracy and prove to be betterthanotherclassifier. Many hidden convolution layers and more hidden artificial neurons could improve the accuracy of the outcome.

Index Terms - Convolution, Padding, Stride, Softmax, Maxpooling, ReLU.

I. INTRODUCTION

Handwritten digit classification problem is standard data set which is used in vision of computer and deep learning. In this study, we show how to use MATLAB to create a convolutional neural network for handwritten digit categorization. MNIST is a dataset of 10,000 small28×28pixel images which is of grayscaleimagesofhandwrittensingledigitsbetween 0 and 9. To begin, we'll split the dataset into training and validation datasets, with each category in the training set including 750 photos and the remaining images from each classification in the validation dataset.Thedevelopedmodelcontainsthreeconvolutionlayersforfeaturesextractionand1fully connected layer for classification. Each convolution layer is followed by a max-pool layer and the ReLu activation function.RelU introduces non-linearity and the max-pooling reduces featuremapsize.ActivationfunctionSoftmaxisusedintheclassification layer. Model gives 99.60% accuracy during thevalidation process.When tested with new images, it accuratelyclassifiesthenewtestdata tocorresponding classes.

II. THEORY

A.MachineLearning

Artificial intelligence (AI), machine learning (ML), and deep learning (DL) are all different types of AI. They are, nonetheless, linked in the following way.: “Deep Learning is a kind of Machine Learning,andMachineLearningisakindofArtificialIntelligence”. Machine Learning is a method of constructing a model using training data. Dependingonthetrainingmethod, machine constructing a model usingtrainingdata. Dependingonthetrainingmethod, The input data will be classified by a machine, which will determine which group it belongs to.Regression willpredict thevaluesandtakesthe value for the correct output in training data. Machine Learning's success will be determined by how well the generalisation process is implemented. We will require a significant amount of unbiased training dataset to prevent performance degradation due to the difference between the training dataset and the real input dataset.

B. ArtificialNeural Network

An artificial neural network is a node-based network that mimics the neurons in the human brain. The weighted sum of the input signals and output signals, as well as the outcome of the activation function with the weighted sum, will be calculated by the nodes. The layer of nodes is used to build the majority of neural networks. The signal enters the layered neural network from the input layer, passes through the hidden layer, and emerges out the output layer. In the case of the neural network, supervised learning will be used to change the weights and narrow the gap between the proper desired output and the neural network's output. The method which is used to adjustthe weight according to the training data is called as learningrule.The delta rule is the learningrule of the neural network. The delta rule is an iterative process that leads to a solution over time. As a result, the neural network should be trained multiple times with the training dataset until the error is as low as possible. There are two types two types of neural network. The single layer neural network is only useful for certain types of problems, and the multilayer neural network was created to solve the single layer neural network's limitations.

C. Back–propagationAlgorithm

Themulti-layerneuralnetworkcan’tbetrainedusing delta rule and it should be trained using the back-propagationalgorithm,whichiscalledasthelearningruleofDeepLearning.Back-propagationalgorithm will define the error of the hidden layer as it propagatesthe output error backward from output layer. After obtaining the error of the hidden layer, the weights of each layerswill be adjusted using the delta rule. The importance of thebackpropagationalgorithmisthatit providesasystematicmethod to find the error of the hidden node. The number of output nodes and activation function for a neural network classifier are usually determined by whether it is for a binary classification (two classes) or a multiclass classification (three or more classes). A single output node and sigmoid activation function are used to build the neural network for binary classification.The activation function's maximum values are transformed from the training data's proper output. The neural network has output nodes equal to the number of classes in multi-layer categorization.The softmax function is used to activate the output node's function. The one-hot encoding method is used to turn the right output of the training data into a vector. The crossentropy function is used in the learning rule's cost function.

D. Deep Learning

DeepLearningisdefinedastheMachineLearning technique that employs the deep neural network. The deep neural network's poor performance is due to a lack of sufficient model training.The vanishing gradient, overfitting, and computing burden are the three key issues. Using the ReLU activation function and the cross entropy-driven learning rule, the vanishing gradient problem is improved. The deep neural network is more fit tooverfitting.Thisproblem is solved usingthedropregularization or dropout.Due of the complex calculations, more training time is required. Using the GPU and different methods, this was alleviated to a considerable amount.

E. Convolution NeuralNetwork

Instead of providing the original image, the feature map, which emphasises the unique features, should be provided to improve Machine Learning's image recognition performance. Previously, the feature extractor had been created by hand.For the feature extractor, CNN uses a unique sort of neural network, the weights of which are determined during the training phase. CNN's feature extractor is made up of alternating stacks of convolution and pooling layers.The convolution layer generates images that enhance the features of the input image using convolution filters. This layer produces the same number of output images as the number of convolution filters in the network.In reality, the convolution filter is nothing more than a two-dimensional matrix. The image size is reduced by the pooling layer. It binds adjacent pixels and substitutes a representative value for them. The greatest or mean value of the pixels is the representative value.

III. IMPLEMENTATIONSTYLING

In this section, we'll show you how to use Deep Learning to design and train a simple convolutional neural network. Convolutional neural networks are a type of Deep Learning technology that is used to recognise images. The procedure is as follows:

a) Loadandexploring theimage data.

Input thedigitsampledataasimagedatastore imds. The command imageDatastoreautomatically will labeltheimagesbasedon the foldernames and stores the data. The imageDatastore allows you to store enormous amounts of image data, including data that won't fit in memory, and read batches of photos quickly during convolution neural network training. It will divide the data into training and validation datasets, with each class including 750 photos in the training set and the remaining imagesfrom each class in the validation set. Split EachLabel is used to do thisfunction.imds=imageDatastore(digitDatasetPath,'IncludeSubfolders‟,true,'LabelSource','foldernames');[imdsTrain,imdsValidation]=splitEachLabel(imds,750,„‚randomize');

b) Definethe networkarchitecture.

It specifies the architecture of Convolutional Neural Networks. There are three convolutional layers and one fully connected layer in this rudimentary convolutional neural network. Each convolution layer has a max-pool layer and a ReLU activation function. Each convolution layer will generatenewimagescalledfeaturemaps.It will contain filters that convert images. It generates same number offeature maps as the convolution filters. ReLU introducesnonlinearityandmax-poolingreducesthesizeoftheimageas it combines neighbouring pixels of a certain area of theimage into a single maximum value. MATLAB codes given below showshow to define each layer.layers =[imageInputLayer([28281])convolution2dLayer(3,8,'Padding','same')batchNormalizationLayerreluLayermaxPooling2dLayer(2,'Stride',2)

Picture Input Layer: This layer defines the image size, which is 28-by-28-by-1 in this case. These figures represent the channel's height, width, and length.

Convolution layer: Filter Size is the height and width of the filters used by the training function while scanning the images, and it is the first argument in the convolution layer. The second option is the number of filters, or the number of neurons that link to the same region of the input. This option determines the number of feature maps. Use the 'Padding' name-value combination to add padding to the input feature map. For a convolution layer with a default stride of 1,'same' padding ensures that the spatial output size is the same as the input size.

Layer normalization: Batch Normalization The activations and gradients propagating through a network are normalised by batch normalisation layers, making network training a simpler optimization task.The rectified linear unit (ReLU), which introduces nonlinearity, is the most popular activation function.The Max Pooling Layer decreases the size of feature maps by removing superfluous spatial information.

Fully Connected Layer: Following the convolution and down sampling layers are one or more completely linked layers. As the name implies, a fully connected layer is one in which all of the neurons in the preceding layer are connected. This layer collects all of the information learnt by the preceding layers throughout the image to discover the broader patterns. The features for image classification are combined in the final fully linked layer. As a result, the Output Size parameter in the last fully linked layer equals the number of classes in the target data.The In this case, the output size is ten, matching the ten classes.

SoftMaxLayer: The fully connected layer's output is normalised using the SoftMax activation function. The SoftMax layer produces a set ofpositive numbers that add up to one, which the classification layer can utilise as classification probabilities.

ClassificationLayer: The categorization layer is the final layer. This layer assigns each input to one of the mutually exclusive classes using the probabilities supplied by the SoftMax activation function, and computes the loss.

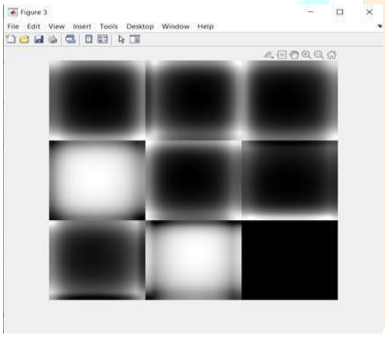
c) Options for training should be specified.

After you've defined the network structure, you'll need to choose your training settings. Train the network using a 0.01 learning rate and stochastic gradient descent with momentum (SGDM). Four epochs should be the maximum number of epochs. An epoch is a complete training cycle on the entire training data set. You can maintain track of the network's accuracy throughout training by specifying validation data and frequency. Shuffle the data every epoch. During training, the software uses training data to train the network and then uses validation data to calculate accuracy at regular intervals. The validation data is not used to update the network weights. Disable the command window output and enable the training progress show.

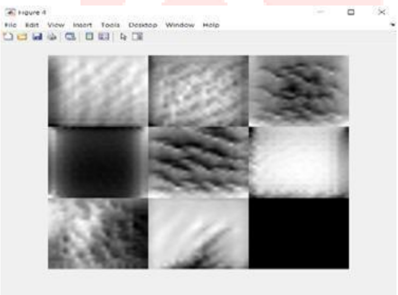
d) It is necessary to prepare the network.

The layers-based architecture, training data, and training settings are all used to train the network. Train Network will use a GPU by default if one is available; otherwise, it will use a CPU. The training progress plot displays the mini-batch loss and accuracy, as well as the validation loss and accuracy. The loss is the cross-entropy loss. The accuracy is the percentage of images successfully classified by the network.

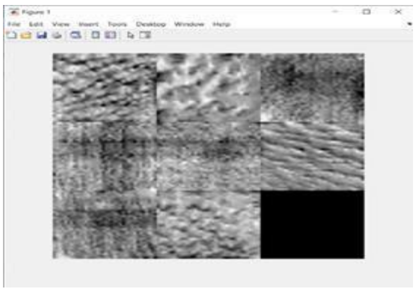
e) Predict the labels of new data and calculate theclassificationaccuracy.



Visualizationoffeaturesextracted byconv1



Visualizationoffeaturesextracted byconv2



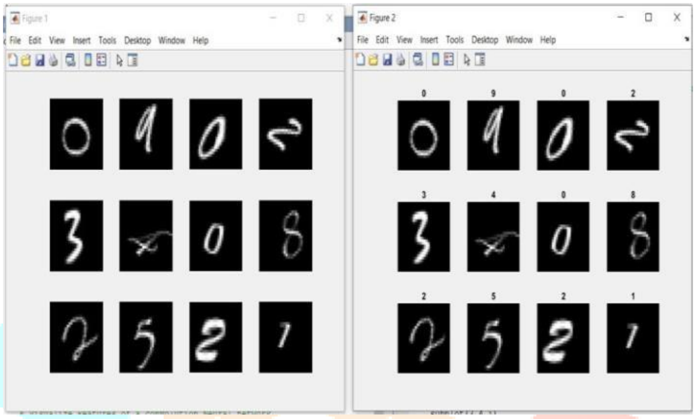
Visualizationoffeaturesextracted by conv3

IV. RESULTS

Method-1(Validation)

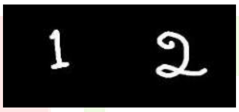
```
>> [YPred, Score] = classify(Diginet,imdsValidation); % Recognizing digits
YValidation = imdsValidation.Labels; % Getting Lables

% Output Method 1 Finding Accuracy in precentage
Accuracy = sum(YPred == YValidation)/numel(YValidation)

Accuracy =

    0.9960
```

Method-2(Validation)

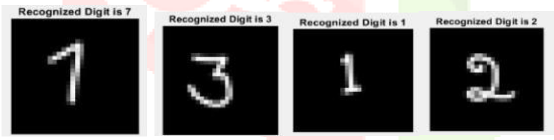


InputImagesClassified Images

After all the output methods are checked, next we need totestthemodelwithown hand written digitimages



TestImages(New Images)



ClassifiedImages

V.CONCLUSIONANDFUTURE SCOPE

The first step towards the wide field of Artificial Intelligence and Computer Vision is handwritten digitrecognition CNN outperforms alternative classifiers, as evidenced by the results of the trial. With more convolution layers and buried neurons, the findings can be made more precise. Digit recognition is an excellent model issue for learning about neural networks and provides a solid foundation for developing more complex deep learning approaches.

REFERENCES

[1]SavitaAhlawat,AmitChoudhary,AnandNayyar,SaurabhSinghandByungunYoon,“ImprovedHandwrittenDigitRecognitionUsingConvolutionalNeuralNetworks”25May2020;Accepted:9June2020;Published:12 June2020.

[2]Fathima Siddique, ShadmanSakib, Md. Abu Bakr Siddique,“RecognitionofHandwrittenDigitusingConvolutionalNeuralNetworkinPythonwithTensorFlowandComparisonofPerformanceforVariousHidden Layers”

[3]Md.AnwarHossain&Md.MohonAli“RecognitionofHandwrittenDigit using Convolutional Neural Network” Volume 19 Issue 2Version1.0Year2019,GlobalJournals,OnlineISSN:0975-4172

[4]Saqib Ali1, Zeeshan Shaukat, Muhammad Azeem, Zareen Sakhawat,Tariq Mahmood, Khalil ur Rehman, “An efficient andimprovedschemeforhandwrittendigitrecognitionbasedonconvolutionalneuralnetwork”,©SpringerNatureSwitzerlandAG2019

[5]Aditi M Joshi, Kinjal Thakar “A Survey on Digit Recognition UsingDeepLearning”,©2018IJNRD,Volume3,ISSN:2456-4148.

[6]HaiderA.Alwzwazy,HayderM.Albehadili,YounesS.Alwan,NazE.Islam,“HandwrittenDigitRecognitionUsingConvolutional NeuralNetworks”Vol.4,Issue2,February2016, ISSN(Print):2320-9798