A NOVEL METHOD FOR HANDWRITTEN DIGIT RECOGNITION SYSTEM

LITERARURE SURVEY

[1] Handwritten digit recognition is the ability of a computer to recognize the human handwritten digits from different sources like images, papers, touch screens, etc, and classify them into 10 predefined classes (0-9). The models with low accuracy are not suitable for real-world applications. Methods like SVM, MLP, CNN and KNN are some machine learning and deep learning algorithms. In SVM, all the features are plotted and classification is performed using hyperplanes that separate the classes correctly. The MLP algorithm uses input layer, hidden layer and output layer. Each layer consists of several nodes that are also formally referred to as neurons and each node is interconnected to every other node of the next layer. The number of hidden layers may increase according to the problem with no restrictions to the number of nodes. The CNN is widely used for image processing where input image is given in small chunks rather than pixels at a time, to detect uncertain patterns more efficiently. CNN contains 3 layers namely, an input layer, an output layer, and multiple hidden layers which include Convolutional layers, Pooling layers (Max and Average pooling), Fully connected layers (FC), and normalization layers. After implementing all the three algorithms that are SVM, MLP and CNN, it is found that SVM has the highest accuracy while training data and CNN has the utmost accuracy while testing the data. SVM took the minimum time for execution while CNN took the maximum running time.

[2] Md. Anwar Hossain and Md. Mohan Ali state that CNN is an important tool when it comes to learning deep learning with a NN. Convolutional Neural Networks are a special kind of multi-layer neural networks designed to recognize visual patterns directly from pixel images with minimal pre-processing. Convolution is filtering the image with a smaller pixel filter to decrease the size of the image without losing the relationship between pixels. When we apply convolution to the 5x5 image by using a 3x3 filter with 1x1 stride (1-pixel shift at each step), we will end up having a 3x3 output (64% decrease in complexity). Inserting pooling layers is common which reduces overfitting problems. Max Pooling is one of the most common pooling techniques. A fully connected network is in any architecture where each parameter is linked to one another to determine the relation and effect of each parameter on the labels. All these convolution layers, pooling layers and fully connected layers are stacked. The pre-processed data is fed to the model. CNN transforms the original image layer by layer from the original pixel values to the final class scores. The parameters in

the convolutional / fully connected layers will be trained with stochastic gradient descent algorithm so that the class scores are consistent with the labels in training set for each image. The algorithm will prepare the trained model which will be used to classify the digits present in the test data.

[3] Alexander K. Seewald talks about a general weakness in present intelligent image analysis systems, calling it 'brittleness' in AI terminology. He has investigated two new aspects of such systems, they are essential training set size i.e., the relation between training set size and accuracy/error rate so as to determine the number of labelled training samples that are essential for a given performance level and dataset-independence i.e., how well models trained on one sample dataset for handwritten digit recognition perform on other sample datasets for handwritten digit recognition after comprehensive normalization between the datasets. For the first aspect, all the datasets considered gave similar performance but for the second aspect, none of the considered learning systems were able to transfer their expertise to other datasets. He says that small differences in the pre-processing methods which have not been documented in sufficient detail may be responsible for this effect. Another explanation might be that idiosyncrasies of the specific dataset used for training are learned as well and hamper the generalization ability of the underlying learning algorithm. This effect is observed independently of learning algorithm or feature representation.

[4] An enormous number of CNN classification algorithms have been proposed in the literature. Nevertheless, in these algorithms, appropriate filter size selection, Based on Convolution al Neural Network Approach data preparation, limitations in datasets, and noise have not been taken into consideration. As a consequence, most of the algorithms have failed to make a noticeable improvement in classification accuracy. To address the shortcomings of these algorithms, the paper presents the following contributions: Firstly, after taking the domain knowledge into consideration, the size of the effective receptive field (ERF) is calculated. Calculating the size of the ERF helps us to select a typical filter size which leads to enhancing the classification accuracy of the CNN. Secondly, unnecessary data leads to misleading results and this, in turn, negatively affects classification accuracy. To guarantee the dataset is free from any redundant or irrelevant variables to the target variable, data preparation is applied before implementing the data classification mission. Thirdly, to decrease the errors of training and validation, and avoid the limitation of datasets, data augmentation has been proposed. Fourthly, to simulate the real-world natural influences that can affect image quality, they propose to add an additive white Gaussian noise with $\sigma = 0.5$ to the MNIST dataset. As a result, the CNN algorithm

achieves state-of-the-art results in handwritten digit recognition, with a recognition accuracy of 99.98%, and 99.40% with 50% noise.

[5] Handwritten digit recognition has recently been of much interest among the researchers because of the evolution of various Machine Learning, Deep Learning and Computer Vision algorithms. They compare the results of some of the most widely used Machine Learning Algorithms like CNN- convolution neural networks and with Deep Learning algorithm like multilayer CNN using Keras with Theano and Tensorflow. MNIST is a dataset which is widely used for handwritten digit recognition. The dataset consists of 60,000 training images and 10,000 test images. The artificial neural network can almost mimic the human brain and are a key ingredient in image processing field. For example, Convolution Neural network with back propagation for image processing. The applications where handwritten digit recognition can be used is banking sector where it can be used to maintain the security pin numbers, it can be also used for blind peoples by using sound output.

[1] Handwritten Digit Recognition using Machine and Deep Learning Algorithms by Ritik Dixit of Computer Science and Engineering, Rishika Kushwah of Computer Science and Engineering, Samay Pashine of Computer Science and Engineering, Acropolis Institute of Technology & Research, Indore, India, 23 June 2021

[2] Recognition of Handwritten Digit using Convolutional Neural Network (CNN), By Md. Anwar Hossain & Md. Mohon Ali, Pabna University of Science & Technology, 2019

[3] On the Brittleness of Handwritten Digit Recognition Models, Alexander K. Seewald, 30 Nov 2011

[4] A Novel Handwritten Digit Classification System Based on Convolution al Neural Network Approach, Ali Abdullah Yahya, 18 September 2021

[5] Handwritten Digit Recognition Using Machine Learning Algorithms, S. M. Shamim, Md Badrul Alam Miah, Angona Sarker, Masud Rana, Abdullah Al Jobair, March 2018