LITERATURE SURVEY

PROJECT TITLE : A NOVEL METHOD FOR HANDWRITTEN DIGIT RECOGNITION SYSTEM

STUDENTS

A.ARCHANA [422719106001]

D.DHARINI [422719104008]

K.GOWRI [422719106003]

M.PUNITHA [422719104022]

M.SUBALAKSHMI [422719104031]

JOURNAL PAPERS AND INFERENCE

TITLE **: Handwriting Digit Recognition for Banking System , International Journal of Engineering Research & Technology (IJERT) V.Gopalakrishnan ,R.Arun ,L.Sasikumar, Mrs.K.Abirami , 2021.**

A handwriting digit recognition system's goal is to transform handwritten digits into representations that computers can understand. The major goal of this work is to provide efficient and trustworthy methods for handwritten digit recognition and make banking activities simpler and error-free. The Handwritten Digit Recognition System (HDR) is designed to read and understand handwritten input on paper documents or in the form of images. Traditional handwriting recognition algorithms have depended heavily on existing information and customised features. It is difficult to train an optical character recognition (OCR) system using these requirements. Convolutional neural networks (CNNs) are the most efficient at understanding the structure of handwritten symbols and words, which enables automatic feature extraction.

TITLE : **A Novel Handwritten Digit Classification System Based on Convolutional Neural Network Approach A Novel Handwritten Digit Classification System Based on Convolutional Neural Network Approach. Ali Abdullah Yahya , Jieqing Tan and Min Hu , 2021.**

There have been a tonne of CNN classification algorithms put forth in the literature. However, these algorithms do not take into account the proper filter size selection, data preparation, dataset restrictions, or noise. As a result, few algorithms have been able to significantly increase classification accuracy. Our research makes the following improvements to overcome the inadequacies of existing algorithms: First, the size of the effective receptive field (ERF) is determined after taking the domain knowledge into account. The size of the ERF is taken into account while choosing a typical filter size, which improves the classification accuracy of our CNN. Secondly, excessive data produces inaccurate results, which has a impact on categorization accuracy. 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, we propose to add an additive white Gaussian noise with $\sigma = 0.5$ to the MNIST dataset. As a result, our 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 . In our experiments, batch normalization has been used to improve the training performance and enhance the stability of our model. With the usage of batch normalization, we can speed up the training, reduce training and testing time, in addition to lowering the sensitivity initialization. In order to avoid overfitting and underfitting, an early stopping technique has used to determine the optimal number of training epochs.

TITLE **: NOVEL FRAMEWORK FOR HANDWRITTEN DIGIT RECOGNITION THROUGH NEURAL NETWORKS ,Savita Ahlawat,Amit Choughary,Anand Nayyar,Saurabh Singh ,Byungun Yoon .Sensors (Basel) , 2020.**

Accurately identifying and categorising the hand-written characters presents the biggest barrier for natural language processing algorithms. Since each person has a different handwriting style, size, and other handwriting characteristics, accurately reading handwritten characters is a difficult challenge for humans as well. Even though this machine vision task is rather simple, greater accuracy compared to current methods is still preferred. An innovative neural network-based framework for handwritten character recognition is proposed in this manuscript. In order to achieve image flattening, the suggested neural network-based framework converts the raw data set to a NumPy array and feeds the same into a pixel vector

before feeding the network. The activation function is used in the neural network to transfer the outcome value to hidden layer where it is further minimized through the use of minimized mean square and back propagation algorithms before applying a stochastic gradient on the resultant mini–batches.

TITLE : **Handwritten Digit Recognition Using Machine Learning Algorithms, S.M.Shamim, Md Badruln Alam Miah, Angona Sarkar, Masud Rana March 2018.**

Handwritten character recognition is one of the practically important issues in pattern recognition applications. The applications of digit recognition includes in postal mail sorting, bank check processing, form data entry, etc. The heart of the problem lies within the ability to develop an efficient algorithm that can recognize hand written digits and which is submitted by users by the way of a scanner, tablet, and other digital devices. This paper presents an approach to off-line handwritten digit recognition based on different machine learning technique. The main objective of this paper is to ensure effective and reliable approaches for recognition of handwritten digits. Several machines learning algorithm namely, Multilayer Perceptron, Support Vector Machine, Naïve Bayes, Bayes Net, Random Forest, J48 and Random Tree has been used for the recognition of digits using WEKA . The main objective of this investigation is to find a representation of isolated handwritten digits that allow their effective recognition This work is carried out as an initial attempt, and the aim of the paper is to facilitate for recognition of handwritten numeral without using any standard classification techniques .

TITLE **: A NOVEL METHOD FOR THE RECOGNITION OF ISOLATED HANDWRITTEN ARABIC CHARACTERS , A Sahlol ,C Suen , 2014.**

There are many difficulties facing a handwritten Arabic recognition system such as unlimited variation in human handwriting, similarities of distinct character shapes, interconnections of neighbouring characters and their position in the word. The typical Optical Character Recognition (OCR) systems are based mainly on three stages, preprocessing, features extraction and recognition. This paper proposes new methods for handwritten Arabic character recognition which is based on novel preprocessing operations including different kinds of noise removal also different kind of features like structural, Statistical and Morphological features from the main body of the character and also from the secondary components. Evaluation of the accuracy of the selected features is made. The system was trained and tested by back propagation neural network with CENPRMI dataset. The proposed algorithm obtained promising results as it is able to recognize 88% of our test set accurately.

In Comparable with other related works we find that our result is the highest among other published works**.**

TITLE **: A NOVEL METHOD FOR HAND WRITTEN DIGIT RECOGNITION USING DEEP LEARNING**, **Rohini. M , Dr.D.Surendran ,.INTERNATIONAL JOURNAL OF CURRENT ENGINEERING AND SCIENTIFIC RESEARCH (IJCESR) , 2019.**

Handwritten digit recognition has recently been of very interest among the researchers because of the evolution of various Machine Learning, Deep Learning and Computer Vision algorithms. In this report, We 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 consist of 60,000 training images and 10,000 test images . The artificial neural networks can all most mimic the human brain and are a key ingredient in image processing field . For example Convolution Neural networks with back propagation for image processing .The applications where these handwritten digit recognition can be used are 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.

TITLE **: A Machine Learning and Deep Learning Approach for Recognizing Handwritten Digits Ayushi Sharma , Harshit Bhardwaj , Arpit Bhardwaj , Aditi Sakalle , Divya Acharya and Wubshet , 2017**.

Optical character recognition (OCR) can be a subcategory of graphic design that involves extracting text from images or scanned documents. We have chosen to make unique handwritten digits available on the Modified National Institute of Standards and Technology website for this project. The Machine Learning and Deep Learning algorithms are used in this project to measure the accuracy of handwritten displays of letters and numbers. Also, we show the classification accuracy comparison between them. .The results showed that the CNN classier achieved the highest classification accuracy of 98.83%.In this paper, we applied machine learning and deep learning techniques to predict the handwritten digits. Popular algorithms such as KNN, SVM, RFC, DECISION TREE, GNB, GP, and CNN were tested to analyze the differences between them. We are using keras the backend and tensorflow as the software library. The CNN classier outperforms the other classier with a classification accuracy of 98.83% for the recognition of handwritten digits.

TITLE : **Survey of Handwritten Character Recognition with MNIST and EMNIST Alejandro Baldominos A , Yago Saez and Pedro Isase , 2019 .**

This paper summarizes the top state-of-the-art contributions reported on the MNIST dataset for handwritten digit recognition. This dataset has been extensively used to validate novel techniques in computer vision, and in recent years, many authors have explored the performance of convolutional neural networks (CNNs) and other deep learning techniques over this dataset .This paper makes a distinction between those works using some kind of data augmentation and works using the original dataset out-of-the-box. Also, works using CNNs are reported separately. Nowadays, a significant amount of works have attained a test error rate smaller than 1% on this dataset, which is becoming non-challenging. By mid-2017, a new dataset was introduced . EMNIST, which involves both digits and letters, with a larger amount of data acquired from a database different than MNIST's. In this paper, EMNIST is explained and some results are survey. This paper has provided an exhaustive review of the state of the art for both the MNIST and EMNIST databases. The MNIST database of handwritten digits was introduced almost two decades ago and has been extensively used to validate computer vision algorithms, and more recently, also as a benchmark to test different convolutional neural networks architectures and approaches. some works are proposing novel developments or improvements, which are often combined with convolutional neural networks, reporting outstanding results. Although accuracy in MNIST and EMNIST is very close to 100% and will hardly increase, these novel developments might hold the key for breaking more complex computer vision problems.

TITLE **: Multi-Language Handwritten Digits Recognition based on Novel Structural Features . Jaafar M. Alghazo Ghazanfar Latif. Loay Alzubaidi Ammar ,March – April 2019 .**

Automated handwritten script recognition is an important task for several applications. In this article, a multi-language handwritten numeral recognition system is proposed using novel structural features. A total of 65 local structural features are extracted and several classifiers are used for testing numeral recognition. Random Forest was found to achieve the best results with an average recognition of 96.73%. The proposed method is tested on six different popular languages, including Arabic Western, Arabic Eastern, Persian, Urdu, Devanagari, and Bangla. In recent studies, single language digits or multiple languages with digits that resemble each other are targeted. In this study, the digits in the languages chosen do not resemble each other. Yet using the novel feature extraction method a high recognition accuracy rate is achieved. Experiments are performed on well-known available datasets of each language. A dataset for Urdu language is also developed in this study and introduced as PMU-UD. Results indicate that the proposed method gives high recognition accuracy as compared to other methods. Low error rates and low confusion rates

were also observed using the novel method proposed in this study. In this paper, we proposed a novel Local Feature Extraction method that is used to design a unified multi-language handwritten numeral recognition system. We targeted many languages even though their digits do not resemble each other.The possibility of redesigning the system in a cloud-based environment will also be part of future work in order to achieve a continuous learning curve and obtain a continuous accuracy improvement.

TITLE : **Unknown-Length Handwritten Numeral String Recognition Using Cascade of PCA-SVMNet Classifiers. SALEH ALY AND AHMED MOHAMED , Deanship of Scientic Research (DSR), Majmaah University , April 29, 2019.**

Automatic recognition of handwritten digit string with unknown length has many potential real applications. The most challenging step in this problem is how to efficiently segment connected and/or overlapped digits exhibited in the input image. Most existing numeral string segmentation approaches combine several segmentation hypotheses to handle various types of connected digits. This paper proposes a new handwritten digit string recognition without applying any explicit segmentation techniques. The proposed method uses a new cascade of hybrid principal component analysis network (PCANet) and support vector machine (SVM) classifier called PCA-SVMNet . PCANet is an emerging unsupervised simple deep neural network typically with only two convolutional layers. The proposed PCA-SVMNet model adds a new fully connected layer trained separately using SVM optimization method. Cascaded stages of PCA-SVMNet classifiers are constructed and trained to recognize various types of isolated and connected digits. Every PCA-SVMNet classifier is trained separately using combinations of real and synthetic touching digits. The first 1D-PCA-SVMNet stage is trained to recognize isolated handwritten digits (0 . . . 9) while forwarding non-isolated digits to the next stages. Each of the following stages is designed to recognize a class of connected digits and forwards the higher class to its successor. Multiple stages can be added accordingly to classify more complex touching digits. The experimental results using NIST SD19 real dataset show that the cascade of PCA-SVMNet classifier efficiently recognizes unknown handwritten digit string without applying any sophisticated segmentation methods. The proposed method achieves state-of-the-art recognition accuracy compared to other segmentation-free techniques . In the experiment using NIST SD19 dataset, where most of the strings contain only isolated digits, the method achieves state-of-the-art results compared to segmentation-free methods and comparable results with segmentation-based techniques.