Ideation Phase

Literature Survey

Date	3 September 2022
Team ID	PNT2022TMID06143
Project Name	A Novel Method for Handwritten Digit Recognition System

Introduction

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. Human can visually sense the world around them by using their eyes and brains.

Computer vision works on enabling computer and process images in the same way that human vision does. Several algorithms developed in the area of computer vision to recognise images. The goal of this work is to create a model to identify and determine the handwritten digits from its database with better accuracy and aim to complete this by using the concept of Convolutional Neural Network and MNIST dataset.

A method called soft max regression is used for assigning the probabilities that to handwritten digits. Though the goal is to create a model which can recognise the digits but can extend for letters and then a person's handwriting. Through this work, people can learn under practically apply the concept of a Convolutional Neural Networks.

Basic Terms:

Image Acquisition

Digitized/Digital Image is initially taken as input. The most common of these devices is the electronic tablet or digitizer. These devices use a pen that is digital in nature. Input images for handwritten characters can also be taken by using other methods such as scanners, photographs or by directly writing in the computer by using a stylus

Segmentation

Segmentation is done by separation of the individual characters of an image. Generally, document is processed in a hierarchical way. At first level lines are segmented using row histogram. From each row, words are extracted using column histogram and finally characters are extracted from words.

Pre-processing

The process of extraction of text from the document is called pre-processing or document analysis. Pre-processing includes background noise reduction, filtering, original image restoration etc.

Feature Extraction

Feature extraction is the name given to a family of procedures for measuring the relevant shape information contained in a pattern so that the task of classifying the pattern is made easy by a formal procedure. Among the different design issues involved in building a recognizing system, perhaps the most significant one is the selection of set of features. Feature extraction for exploratory data projection enables high dimensional data visualization for better data structure understanding and for cluster analysis. In feature extraction for classification, it is desirable to extract high discriminative reduced-dimensionality features, which reduce the classification computational requirements. Here, individual image glyph is considered and extracted for features. Each character glyph is defined by the following attributes: (1) Height of the character. (2) Width of the character. (3) Numbers of horizontal lines present short and long. (4) Numbers of vertical lines present short and long. (5) Numbers of circles present. (6) Numbers of horizontally oriented arcs. (7) Numbers of vertically oriented arcs. (8) Centroid of the image. (9) Position of the various features. (10) Pixels in the various regions.

Classification

Template matching or matrix matching, is one of the most common classification methods. Here individual image pixels are used as features. Classification is performed by comparing an input character with a set of templates (prototypes) from each character class. If the input characters match with the (prototypes) templates character's identity is made equivalent to most similar template and similarity of measure increases; and if the input characters does not match with the template similarity of measure decreases [3]. All the above phases are applied in a pipelined fashion i.e., the success of each phase depend on the success of previous phase and all phases depend on the success and accuracy of each other phase as output of previous phase will be used as input to next phase. Handwritten data is captured and stored in its digital format either by scanning the handwritten document by scanner or by writing with a stylus pen on an electronic surface such as a digitizer/PDA or tablet with a liquid crystal display. The two approaches may be distinguished as offline or online handwriting recognition [15]. A scanned paper may consist of either a machine printed/hand printed characters or handwritten characters. It may consist of images, graphics or any form kind data. Scanned paper forms included typewritten or computer generated fonts, handwritten characters, check boxes and bubbles, bar codes, and signatures. A Data is to be extracted from these forms.

A novel method for Handwritten Digit Recognition with Neural Networks - MALOTHU NAGU, N VIJAY SHANKAR, K. ANNAPURNA [1]

Character recognition plays an important role in the modern world. It can solve more complex problems and makes humans' job easier. An example is handwritten character recognition. Pattern recognition system consists of two-stage process (Feature Extraction and Classification). Feature extraction is the measurement on a population of entities that will be classified. This assists the classification stage by looking for features that allows fairly easy to distinguish between the different classes.

There are Several Pattern Recognition Methods, they are: Bayesian decision theory, Nearest Neighbour rule, Linear Classification Discrimination

The Bayesian decision theory is a system that minimizes the classification error. This method is based on priority basis, it classifies using priority information about something that we would like to classify. We can use Baye's formula, which states the following: $P(wj|x) = p(x|wj) P(wj)/p(x)$.

The Nearest Neighbour (NN) rule is used to classify handwritten characters. The distance measured between the two-character images is needed in order to use this rule. The goal of Linear Classification is to assign observations into the classes. This can be used to establish a classifier rule so that it can assign a new observation into a class. In another words, the rule deals with assigning a new point in a vector space to a class separated by a boundary. Linear classification provides a mathematical formula to predict a binary result. This result is a true or false (positive or negative) result. The following is an equation that can be stated as the discriminator: $a_1 x_1 + a_2 x_2 + ... + a_n x_n > x_0$

Shape describes a spatial region. Most shapes are a 2-D space. Shape recognition works on the similarity measure so that it can determine that two shapes correspond to each other. The recognition needs to respect the properties of imperfect perception, for example: noise, rotation, shearing, etc. One of the techniques used in shape recognition is elastic matching distance.

The difficult task is there are some handwritten digits that often run together or not fully connected. Numeral 5 is an example. But once these tasks have been carried out, the digits are available as individual items. But the digits are still in different sizes. Therefore, a normalization step has to be performed so we can have to have digits in equal sizes. After the digits are normalized, they are fed into the ANN. This is a feed-forward network with three hidden layers.

The input is a 16 x 16 array that corresponds to the size of a normalized pixel image. The first hidden layer contains 12 groups of units with 64 units per group. Each unit in the group is connected to a 5 x 5 square in the input array and all 64 units in the group have the same 25 weight values. The second hidden layer consists of 12 groups of 16 units. This layer operates very similar to the first hidden layer, but now it seeks features in the first hidden layer. The third hidden layer consists of 30 units that are fully connected to the units in the previous layer. The output units are in turn fully connected to the third hidden layer.

Analytical Review of Pre-processing Techniques for Offline Handwritten Character Recognition - K. Gaurav and Bhatia P. K. [2]

OCR has enabled scanned documents to become more than just image files, turning into fully searchable documents with text content that is recognized by computers. With the help of OCR, people no longer need to manually retype important documents when entering them into electronic databases. Instead, OCR extracts relevant information and enters it automatically. The result is accurate, efficient information processing in less time.

Binarization (Thresholding)

Document Image Binarization converts the image into bi-level form in such a way that foreground information is represented by black pixels and background by the white pixels

Noise Reduction Techniques:

The major objective of noise removal is to remove any unwanted bit-patterns, which do not have any significance in the output. Noise reduction techniques are filtering, morphological operations and noise modelling. Filters can be designed for smoothing, sharpening, thresholding, removing slightly textured background and contrast adjustment process. Various morphological operations can be designed to connect broken strokes, decomposed the connected strokes, smooth the contours, clip the unwanted points, thin the characters and extract boundaries. Smoothing can be done by filters [6]. Types of filters available are: Linear Filters (Averaging mask filter), Non-Linear Filters

It is found that major activities in a typical pre-processing system are image enhancements, noise removal, contrast adjustment, binarization, normalization, segmentation. All these techniques are necessary in a pre-processing system; no technique alone cannot be said as complete activity for pre-processing e.g., only using noise removal we cannot completely pre-process an image. Therefore, all image enhancement techniques have to be applied for accuracy in pre-processing system. Pre-processed image can be used in feature extraction phase and then in neural network phase. Even applying all these techniques, we cannot obtain the 100% accuracy in a pre-processing system.

Limitations:

OCR is unable to achieve a 100% recognition rate. Because of this, a system which permits fast and accurate recognition is a major requirement. The success of any OCR device to read accurately is the responsibility of the hardware manufacturer as well as depends on the quality of the items to be processed.

Devnagari numeral recognition by combining decision of multiple connectionist classifiers - Reena Bajaj, Lipika Dey, and S. Chaudhury [3]

Classifier combination

The classifiers used in this approach are based upon different representations of the input pattern. These representations, since they encode different types of property – style-dependent and style-invariant, cannot be combined into a single monolithic feature vector. Individual classifiers dealing with these representations output the class labels depending on the features used. These class labels have been combined using a meta-pi network because it can devise a combination scheme on the basis of consistency and accuracy of the individual classifiers. Each output unit of the meta-pi net modulates output of individual classifiers. The network is called the meta-pi network owing to the multiplicative function that its output units perform. This function serves to combine the outputs of sub-networks (or "modules"), independently trained to classify numerals based on different features. Hence, there are three output nodes in the present meta-pi net. The network has been trained with the samples used for training the individual nets along with newer examples. Through the training process the meta-pi net learns to choose any one of the valid classifier outputs or a combination of these valid outputs to produce the correct global output.

The initial stage of the proposed architecture consists of connectionist modules for style-based categorisation of the input. Since style groups are characteristic of each character, for each character there exists a style categorisation module which acquires knowledge about style categories of the corresponding character through unsupervised learning. Output of this stage would indicate similarity of an unknown input with style categories of the different characters (including the correct one). An interesting feature of this stage is that the classifiers are not forced to classify distinct looking samples of one character into one monolithic class. The classifiers can self-organise themselves to categorise them into distinct style categories.

The novel feature of this work is the approach followed for identification and integration of style specific information in the recognition scheme. Use of multiple classifiers using the meta-pi network is another significant feature of this work. A complete hierarchical recognition architecture has been suggested in this work.

Rohan Sethi, Ila Kaushik [4]

This paper is to demonstrate and represent the work which is related to hand-written digit recognition. The hand-written digit recognition is a very exigent task. In this recognition task, the numbers are not accurately written or scripted as they differ in shape or size; due to which the feature extraction and segmentation of hand-written numerical script is arduous. The vertical and horizontal projections methods are used for the purpose of segmentation in the proposed work. SVM is applied for recognition and classification, while convex hull algorithm is applied for feature extraction.

L. Bottou, C. Cortes, J.S. Denker, H. Drucker [5]

This paper compares the performance of several classifier algorithms on a standard database of handwritten digits. We consider not only raw accuracy, but also training time, recognition time, and memory requirements. When available, we report measurements of the fraction of patterns that must be rejected so that the remaining patterns have misclassification rates less than a given threshold.

Eva Tuba, Nebojsa Bacanin [6]

This paper describes an algorithm for handwritten digit recognition based on projections histograms. Classification is facilitated by carefully tuned 45 support vector machines (SVM) using One Against One strategy. Their proposed algorithm was tested on standard benchmark images from MNIST database and it achieved remarkable global accuracy of 99.05%, with possibilities for further improvement.

U. Ravi Babu, Y. Venkateswarlu, Aneel Kumar Chintha [7]

This paper presents a new approach to off-line handwritten digit recognition based on structural features which does not require thinning operation and size normalization technique. This paper uses four different types of structural features namely, number of holes, water reservoirs in four directions, maximum profile distances in four directions, and fill-hole density for the recognition of digits. The digit recognition system mainly depends on which kinds of features are used. The main objective of this paper is to provide efficient and reliable techniques for recognition of handwritten digits. A Euclidean minimum distance criterion is used to find minimum distances and k-nearest neighbour classifier is used to classify the digits. A MNIST database is used for both training and testing the system. 5000 images are used to test the proposed method a total 5000 numeral images are tested and get 96.94% recognition rate.

Cheng-Lin Liu, K. Nakashima, H. Sako, H. Fujisawa [8]

This paper presents the latest results of handwritten digit recognition on well-known image databases using the state of-the-art feature extraction and classification techniques. The tested databases are CENPARMI, CEDAR, and MNIST. On the test dataset of each database, 56 recognition accuracies are given by combining 7 classifiers with 8 feature vectors. All the classifiers and feature vectors give high accuracies. Among the features, the chain-code feature and gradient feature show advantages, and the profile structure feature shows efficiency as a complementary feature. In comparison of classifiers, the support vector classifier with RBF kernel gives the highest accuracy but is extremely expensive in storage and computation. Among the non-SV classifiers, the polynomial classifier performs best, followed by a learning quadratic discriminant function classifier. The results are competitive compared to previous ones and they provide a baseline for evaluation of future works.

Chao Zhang, Zhiyao Zhou, Lan Lin [9]

This paper proposes a new type of handwritten digit recognition system based on convolutional neural network (CNN). In order to improve the recognition performance, the network was trained with a large number of standardized pictures to automatically learn the spatial characteristics of handwritten digits. For model training, according to the loss function, the convolutional neural network continuously updates the network parameters with the data set in MNIST, which contains 60,000 examples. For model tests, the system uses the camera to capture the pictures composed of the images generated by the test data set of MNIST and the samples written by different people, then continuously processes the captured graphics and refreshes the output every 0.5 seconds.

M. Revow, C.K.I. Williams, G.E [10]

Hinton describes a method of recognizing handwritten digits by fitting generative models that are built from deformable B-splines with Gaussian "ink generators" spaced along the length of the spline. The splines are adjusted using a novel elastic matching procedure based on the expectation maximization algorithm that maximizes the likelihood of the model generating the data. This approach has many advantages: 1) the system not only produces a classification of the digit but also a rich description of the instantiation parameters which can yield information such as the writing style; 2) the generative models can perform recognition driven segmentation; 3) the method involves a relatively small number of parameters and hence training is relatively easy and fast; and 4) unlike many other recognition schemes, it does not rely on some form of pre-normalization of input images, but can handle arbitrary scaling, translations and a limited degree of image rotation. We have demonstrated that our method of fitting models to images does not get trapped in poor local minima. The main disadvantage of the method is that it requires much more computation than more standard OCR techniques

Shadman Sakib [11]

This paper is to observe the variation of accuracies of CNN to classify handwritten digits using various numbers of hidden layers and epochs and to make the comparison between the accuracies. For this performance evaluation of CNN, we performed our experiment using the Modified National Institute of Standards and Technology (MNIST) dataset. Further, the network is trained using stochastic gradient descent and the back-propagation algorithm.

V.N. Manjunath Aradhya, G. Hemantha Kumar [12]

In this paper, we propose a novel system based on radon transform for handwritten digit recognition. We have used a radon function which represents an image as a collection of projections along various directions. The resultant feature vector by applying this method is the input for the classification stage. A nearest neighbour classifier is used for the subsequent recognition purpose. A test performed on the MNIST handwritten numeral database and on Kannada handwritten numerals demonstrate the effectiveness and feasibility of the proposed method.

Savita Ahlawat, Amit Choudhary [13]

This paper is to develop a hybrid model of a powerful Convolutional Neural Networks (CNN) and Support Vector Machine (SVM) for recognition of handwritten digits from MNIST dataset. The proposed hybrid model combines the key properties of both the classifiers. In the proposed hybrid model, CNN works as an automatic feature extractor and SVM works as a binary classifier. The MNIST dataset of handwritten digits is used for training and testing the algorithm adopted in the proposed model. The MNIST dataset consists of handwritten digits images which are diverse and highly distorted. The receptive field of CNN helps in automatically extracting the most distinguishable features from these handwritten digits. The experimental results demonstrate the effectiveness of the proposed framework by achieving a recognition accuracy of 99.28% over MNIST handwritten digits dataset.

S. Bernard, S. Adam, L. Heutte [14]

In the pattern recognition field, growing interest has been shown in recent years for multiple classifier systems and particularly for bagging, boosting and random sub-spaces. Those methods aim at inducing an ensemble of classifiers by producing diversity at different levels. Following this principle, Breiman introduced in 2001 another family of methods called random forest. Their work aims at studying those methods in a strictly pragmatic approach, in order to provide rules on parameter settings for practitioners. For that purpose, we have experimented with the forest-RI algorithm, considered as the random forest reference method, on the MNIST handwritten digits database. In this paper, we describe random forest principles and review some methods proposed in the literature. We present our experimental protocol and results. They finally draw some conclusions on random forest global behaviour according to their parameter tuning.

L.S. Oliveira, R. Sabourin, F. Bortolozzi, C.Y [15]

Suen discusses the use of genetic algorithms for feature selection for handwriting recognition. Its novelty lies in the use of multi-objective genetic algorithms where sensitivity analysis and neural networks are employed to allow the use of a representative database to evaluate fitness and the use of a validation database to identify the subsets of selected features that provide a good generalization. Comprehensive experiments on the NIST database confirm the effectiveness of the proposed strategy.

P. Gallinari, M. Gilloux, A. Bellili [16]

This paper presents an original hybrid MLP-SVM method for unconstrained handwritten digits recognition. Specialized support vector machines (SVMs) are introduced to significantly improve the multilayer perceptron (MLP) performances in local areas around the separation surfaces between each pair of digit classes, in the input pattern space. This hybrid architecture is based on the idea that the correct digit class almost systematically belongs to the two maximum MLP outputs and that some pairs of digit classes constitute the majority of MLP substitutions (errors). Specialized local SVMs are introduced to detect the correct class among these two classification hypotheses.

Reference:

[1] A novel method for Handwritten Digit Recognition with Neural Networks - MALOTHU NAGU, N VIJAY SHANKAR, K. ANNAPURNA

[2] Analytical Review of Pre-processing Techniques for Offline Handwritten Character Recognition - K. Gaurav and Bhatia P. K.

[3] Devnagari numeral recognition by combining decision of multiple connectionist classifiers - Reena Bajaj, Lipika Dey, and S. Chaudhury

[4] Hand Written Digit Recognition using Machine Learning - Rohan Sethi, Ila Kaushik (2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT))

[5] Learning algorithms for classification: A comparison on handwritten digit recognition - L. Bottou, C. Cortes, J.S. Denker, H. Drucker

[6] An algorithm for handwritten digit recognition using projection histograms and SVM classifier - Eva Tuba, Nebojsa Bacanin (2015 23rd Telecommunications Forum Telfor (TELFOR))

[7] Handwritten Digit Recognition Using Structural, Statistical Features and K-nearest Neighbour Classifier - U. Ravi Babu, Y. Venkateswarlu, Aneel Kumar Chintha

[8] Handwritten digit recognition: benchmarking of state-of-the-art techniques - Cheng-Lin Liu, K. Nakashima, H. Sako, H. Fujisawa

[9] Handwritten Digit Recognition Based on Convolutional Neural Network - Chao Zhang, Zhiyao Zhou, Lan Lin

[10] Hand-printed digit recognition using deformable models - M. Revow, C.K.I. Williams, G.E

[11] Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Convolutional Neural Network - Shadman Sakib

[12] Two-Dimensional Matrix Principal Component Analysis Useful for Character Recognition - V.N. Manjunath Aradhya, G. Hemantha Kumar

[13] Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN) - Savita Ahlawat, Amit Choudhary

[14] Dynamic Random Forests - S. Bernard, S. Adam, L. Heutte

[15] A Methodology for Feature Selection Using Multi-Objective Genetic Algorithms for Handwritten Digit String Recognition - L.S. Oliveira, R. Sabourin, F. Bortolozzi,C.Y

[16] An MLP-SVM combination architecture for offline handwritten digit recognition - P. Gallinari, M. Gilloux, A. Bellili.