

Ideation Phase Literature Survey

Date	02 September 2022
Team ID	PNT2022TMID32324
Project Name	Project - A novel method for handwritten digit recognition System

Abstract

Handwritten Digit recognition is becoming more and more important in the modern world. It helps humans ease their jobs and solve more complex problems. An example is handwritten character recognition which is widely used in the world. This system is developed for zipcode or postal code recognition that can be employed in mail sorting. This can help humans to sort mails with postal codes that are difficult to identify. For more than thirty years, researchers have been working on handwriting recognition. Over the past few years, the number of companies involved in research on handwriting recognition has continually increased. The handwritten digit recognition is the capability of computer applications to recognize the human handwritten digits. It is a hard task for the machine because handwritten digits are not perfect and can be made with many different shapes and sizes. The handwritten digit recognition system is away to tackle this problem which uses the image of a digit and recognizes the digit present in the image. Number recognition has numerous operations like number plate recognition, postal correspondence sorting, bank check processing, etc. The goal of our work is to create a model that will be able to recognize and classify the handwritten digits from images by using concepts of Convolution Neural Network. Though the goal of our research is to create a model for digit recognition and classification, it can also be extended to letters and an individual's handwriting. The existing methods in current image recognition use as inputs all the pixels of the image. The purpose of this work is to minimize the number of pixels by using as input the data extracted and calculated from the initial image. This will be used for many fields.

Paper:1 Handwritten Digit Recognition using Machine Learning Algorithms By S M Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana & Abdullah Al Jobair

Mawlana Bhashant Science and Technology Loiversity Abstract 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 cigits 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 machine learning algorithm namely, Multilayer Perceptron, Support Vector Machine, Naive Bayes, Bayes Net, Random Forest, J48 and Random Tree has been used for the recognition of digits using WEKA

Keywords: pattern recognition, handwritten recognition, digit recognition, machine learning, WEKA off-line handwritten recognition, machine learning algorithm, neural network, classification algorithm.

GJCST-D Classification: 17.5, 127, 15.m

Paper 2: Novel Deep Neural Network Model for Handwritten Digit Classification and Recognition

Year: 2021

Authors: Ayush Kumar Agrawal and Vineet Kumar Awasthi

An artificial neural network has one hidden layer between the input and output layers, whereas a deep neural network has numerous hidden layers with input and output layers. Deep neural networks use several hidden layers to increase model performance and achieve higher accuracy compared to accuracy of machine learning models. Most researchers do their research in the area of pattern recognition. In the field of pattern recognition, there are many patterns that can be used, including handwritten numbers, characters, pictures, faces, sounds, and speech. This study focuses on the classification and recognition of handwritten digits. 1000 were utilized as test samples and 1000 were training samples. 10000 picture samples make up the USPS dataset, of which 7291 serve as training samples and 2007 serve as testing samples. We've used the proposed deep neural network technique in this paper to classify and identify data from the ARDIS and USPS datasets. The suggested model consists of six layers with softmax and relu activation functions. After model implementation, accuracy for ARDIS samples reached 98.70% testing and 99.76% training, which is greater than accuracy from prior research. Additionally, using the USPS samples dataset, 98.22% training accuracy and 93.01% testing accuracy were attained. When compared to earlier methodologies, the data show that deep neural networks perform incredibly well.

Paper 3: A Novel Handwritten Digit Classification System Based on Convolutional Neural Network Approach

Year: 2021

Authors: Ali Abdullah Yahya, Jieqing Tan, Min Hu

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. The paper makes the following contributions to solve these methods'

drawbacks: First, the size of the effective receptive field (ERF) is determined after taking domain knowledge into account. They choose a typical filter size with the aid of the ERF calculation, improving the classification accuracy of our CNN. Second, excessive data produces inaccurate results, which has a detrimental impact on classification accuracy. Before carrying out the data classification task, data preparation is conducted to ensure that the dataset is devoid of any redundant or irrelevant variables to the goal variable. Thirdly, data augmentation has been suggested as a way to reduce training and validation errors and prevent dataset limitations. Fourthly, the paper suggests adding an additive white Gaussian noise with a threshold of 0.5 to the MNIST dataset in order to imitate the natural factors that can affect image quality in the real world. With a recognition accuracy of 99.98% and 99.40% with 50% noise, our CNN algorithm achieves state-of-the-art performance in handwritten digit recognition.

Paper 4: Handwritten Character Recognition using Neural Network and TensorFlow

Year : 2019

Authors : Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta

The offline handwritten character recognition in this study will be carried out using Tensorflow and a convolutional neural network. a process known as using SoftMax Regression, one may assign probabilities to one of the many characters in the handwritten text that offers the range of values from 0 to 1, summed to 1. The objective is to create software that is extremely accurate and that has a minimum level of spatial and temporal complexity.

It was determined that strategies for feature extraction like diagonal and direction are significantly better at producing high accuracy

Outcomes in comparison to other conventional vertical and horizontal techniques moreover use the best Neural network tried layers provides the benefit of a higher accurate outcome by having a high noise tolerance. The feed forward model in neural networks is the back-propagation algorithm that was primarily used to classify the characters, recognise them, and receive training continually more. In addition to these, normalizing along with feature extraction, the results were better and more effective. Character recognition is the outcome of accuracy.

The paper will describe the best approach to get more than 90% accuracy in the field of Handwritten Character Recognition (HCR).