Handwritten Character and Digit Recognition

Sanya Tyagi, Simran Sagar and Shreya Agarwal

May 18, 2023

HANDWRITTEN CHARACTER AND DIGIT RECOGNITION

Sanya Tyagi[1], Simran Sagar[1], Shreya Agarwal[1]
[1]Department of Information Technology, KIET Group OF Institutions Ghaziabad

Abstract

This research paper focuses on the development and implementation of a machine learning system for handwritten character and digit recognition. With the increasing use of digital devices, there is a growing need for accurate recognition of handwritten characters and digits. In this paper, we present a comprehensive study of various techniques and algorithms used in the field of handwritten character and digit recognition. We discuss the challenges associated with this task and present a detailed analysis of the performance of different machine learning models in recognizing handwritten characters and digits. Our experimental results show that deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) outperform traditional machine learning models. Additionally, we explore the use of data augmentation techniques to improve the accuracy of the models. The results of our study demonstrate the potential of machine learning in the field of handwritten character and digit recognition, and its potential applications in various domains, including document analysis, handwriting recognition, and digitization of historical documents.

Keywords: Image Classification, Machine Learning Model, Character Recognition

1. Introduction:

Handwritten character and digit recognition is the process to provide the ability for machines to recognize human written digits and characters.It is a process of conversion of handwritten text into machine readable form. The major problem in the handwritten character and digit recognition system is the variation of handwriting styles which can be completely different for different writers. The field of handwritten character and digit recognition has seen significant advancements in recent years, thanks to the development of powerful machine learning which has seen significant advancements in recent years, algorithms and the availability of large datasets for training and testing. This has led to the creation of systems that can accurately and efficiently recognize handwritten characters and digits, paving the way for a range of new applications and opportunities.

Handwritten character and digit recognition refers to the process of identifying and interpreting handwritten characters and digits using computer algorithms.With the increasing amount of handwritten data in today’s world , such as handwritten notes, forms, and signatures, the ability to automatically recognize and extract information from handwritten text is becoming more and more important . Handwritten character and digit recognition has numerous applications, including document analysis , form processing, automatic check processing, and postal automation.

The objective of the handwritten character recognition system is to implement user friendly computer assessed character representation that will allow successful abstraction of characters from handwritten documents and to digitize and translate the handwritten text to machine readable format.

2. Methodology

In this section we introduced the methods we used for this research, we have used different machine learning methods for our image classification, now we will explain each term

2.1 Supervised and unsupervised learning

Supervised learning basically is when we train the machine using data that is well-labelled that means some data is already tagged with the correct answer.The machine is provided with a new set of data. So that the supervised learning algorithm analyzes the training data and produces a correct outcome from labeled data.

Supervised learning is classified into you types of algorithms:

1. Classification
2. Regression

Unsupervised learning is the training of a machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.The task of the machine is to group the unsorted information according to similarities,patterns,and differences without any prior training of the data.

The handwritten character and digit recognition uses supervised learning.In this , we feed the learning algorithm a set of labeled data inputs which is taken from MNIST dataset that contains numbers.The learning algorithm infers a model representing the relationship between the vectorized pixel matrix and the character pixels represent.It is done using neural network and gradient descent.

2.2 Classification

Classification is a fundamental task in handwritten character and digit recognition. The goal of classification is to assign a label or class to an input image based on its features. In handwritten character and digit recognition, the classes correspond to the different characters or digits that the recognition algorithm is trained to recognize.

There are many classification algorithms that can be used in handwritten character and digit recognition, such as Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), Random Forests, Naive Bayes, and Neural Networks. In recent years, Convolutional Neural Networks (CNNs) have become the state-of-the-art approach for handwritten character and digit recognition due to their ability to learn hierarchical representations of the input features.

There are several types of classification methods that can be used in handwritten character and digit recognition. Here are some of the most common ones:
Binary classification,Multiclass classification,Hierarchical classification,Online classification,Offline classification,Batch classification.

Handwritten character and digit recognition commonly use multiclass classification. In this type of classification, the recognition algorithm is trained to classify an input image into one of several classes, where each class corresponds to a different character or digit. For example, in recognizing handwritten digits, the classes are the digits 0 to 9. The goal of multiclass classification is to identify the correct class label for each input image, even if the image contains variations due to different writing styles or variations in handwriting.

2.3 Naive bayes

Naive Bayes is a classifier which uses Bayes Theorem. It calculates the probability for membership of a data-point to each class and assigns the label of the class with the highest probability. Naive Bayes is one of the fastest and simple classification algorithms and is usually used as a baseline for classification problems.
The problem involves building a Naive Bayes classifier on MNIST dataset.Results include confusion matrix,accuracy of each digit,and over accuracy.It also assumes that probability of each pixel is a Gaussian distribution and the probability of each digit is equal. Each image in MNIST dataset is a 28x28 pixels but for our purpose we are converting the images in a single flat array of 784 pixels.We know that each pixel from the array can take values between 0-255,it appears like continuous data.To ease the computation we use Gaussian to get the probabilities of each pixel given in a class.

The equation used:

$$P(x|c) = \frac{1}{\sqrt{(2\pi)^D \ |\Sigma|}} \exp\left(\frac{-1}{2}[(x - \mu)^T \Sigma^{-1} (x - \mu)]\right)$$

Naive Bayes Classifier does not perform well for MNIST dataset and gives an low accuracy comparing to advance cmachine learning methods.

2.4 Support Vector Machine

Handwritten digit recognition: Support vector classifiers can be applied to the recognition of isolated handwritten digits optically scanned.Text Categorization in information retrieval and then categorization of data using labels can be done by SVM SVMs are effective when the number of features is quite large.

SVM is a supervised ML algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.

Support Vectors are simply the coordinates of individual observation. The SVM classifier is a frontier that best segregates the two classes (hyperplane/ line) The data points closest to the hyper-plane are called support vectors, and they influence the hyper-position planes and orientation.

2.5 CNN

Convolutional neural networks are a popular type of neural network that are often used in handwritten characters and digits to recognition tasks. Here’s how CNN’s are typically used in the context.

2.1.1 Preprocessing: The first step is to preprocess the image of the handwritten character or digit. This typically involves converting the image to grayscale, resizing it to a standardized size, and applying normalization techniques to improve the contrast and reduce noise.

2.1.2 Convolutional layers: The preprocessed image is then passed through one or more convolutional layers, which are designed to extract features from the image. Each convolutional layer applies a set of filters to the input image to produce a set of feature maps that highlight different aspects of the image.

2.1.3 Pooling layers: After each convolutional layer, a pooling layer is typically used to reduce the spatial dimensions of the feature maps while preserving the most important information. This helps to make the model more computationally efficient and less prone to overfitting.

2.1.4 Fully connected layers: The output of the convolutional and pooling layers is then flattened and passed through one or more fully connected layers, which perform the final classification of the digit or character.

2.1.5 Softmax output: The final layer of the CNN typically uses a softmax function to produce a probability distribution over the possible classes (i.e., digits or characters).

Multilayer Perceptions A neural network based classifier, called Multi-Layer perception (MLP), is used to classify the handwritten digits. Multilayer perceptron consists of three different layers, input layer, hidden layer and output layer. Each of the layers can have a certain number of nodes also called neurons and each node in a layer is connected to all other nodes to the next layer. For This reason it is also known as a feed forward network.

The number of nodes in the input layer depends upon the number of attributes present in the dataset. The number of nodes in the output layer relies on the number of apparent classes that exist in the dataset. The convenient number of hidden layers or the convenient number of nodes in a hidden layer for a specific problem is hard to determine. But in general, these numbers are selected experimentally. In a multilayer perceptron, the connection between two nodes consists of a weight. During the training process, it basically learns the accurate weight adjustment which corresponds to each connection. For the learning purpose, it uses a supervised learning technique named as Back propagation algorithm.

3. Performance evaluation and Results

In this section, we first introduce our experiment set-up After a brief study of related research work, this work compares the proposed model with other state of the art methods. The performance matrices that we selected are the most relevant in estimating the performance of machine learning methods and Neural Network

We conducted an experiment to evaluate the performance of our proposed algorithm for handwritten character and digit recognition. The experiment was carried out on a dataset of 10,000 handwritten characters and digits from the MNIST database.We implemented our proposed algorithm using Python programming language and the Keras deep learning library. The model consisted of a convolutional neural network (CNN) with two convolutional layers, each followed by a max pooling layer. The output of the second max pooling layer was then flattened and fed into a fully connected layer with 128 neurons, followed by a softmax activation layer for classification.We experimented with different combinations of these techniques and found that a combination of Canny edge detection and LBP performed the best in terms of accuracy and efficiency

Data Description: The image dataset we used to perform to estimate the performance of our model is downloaded from Kaggle[Cite]. It is a labeled dataset, it has 3410 image of hand written characters, these are ten digits and twenty-six english alphabets. Each image is a grayscale scale image of 28 X 28 pixels.

DataSet	MNIST dataset
Type	Images
Classes	36
Size	3410 images, 29 MB
Feature	28 X 28, Grey Scale

Table 1. Dataset statistic

Data Preprocessing The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image rendering it suitable for segmentation. The various tasks performed on the image in pre-processing stage. Binarization process converts a gray scale image into a binary image using global thresholding technique. Detection of edges in the binarized image using sobel technique, dilation the image and filling the holes present in it are the operations performed in the last two stages to produce the pre-processed image suitable for segmentation. The various tasks performed on the image in pre-processing stage are noise removal, binarization, skew correction etc.

Noise removal : It is a process of removing noise from scanned image by using appropriate filter for example smoothing linear filter, order statistic filter etc. Smoothing is used for blurring and reducing noise, and removal of small details from the image extracting large objects.

Binarization : It converts a gray scale image into a binary image using global thresholding technique like otsu's method of thresholding. Otsu's provide optimum value of threshold.

Skew correction : It is removal of skew in scanned document for its proper further segmentation. It is not necessary that handwritten documents are perfectly horizontally aligned thus skew correction methods are required to be performed. For example projection profile analysis, Hough transforms, nearest neighbour clustering, cross-correlation, piece-wise covering by parallelogram.

Image segmentation is a sub-domain of computer vision and digital image processing which aims at grouping similar regions or segments of an image under their respective class labels.

Since the entire process is digital, a representation of the analog image in the form of pixels is available, making the task of forming segments equivalent to that of grouping pixels. Image segmentation is an extension of image classification where, in addition to classification, we perform localization. Image segmentation thus is a superset of image classification with the model pinpointing where a corresponding object is present by outlining the object's boundary. In computer vision, most image segmentation models consist of an encoder-decoder network as compared to a single encoder network in classifiers.The encoder encodes a latent space representation of the input which the decoder decodes to form segment maps, or in other words maps outlining each object's location in the image.

Evaluation Metrices: In our proposed work, we classify the images according to the different image features. To evaluate the performance of our proposed model, we considered model accuracy and entropy as our primary metrics. Most of the existing research work used similar metrics to evaluate their models.

Results: We test our model in a number of experiments: supervised image classification in citation for MNIST datasets, extracted the image features from the dataset an evaluation of various image classification models and a run-time analysis on machine learning methods. For Training and test, we have randomly divided our character image dataset to training plus validation set and testing set. Training plus validation set contained 80% of the images that is 2728 image samples and remaining 20% images which is 682 image samples were treated as testing set.

First we conducted the experiments using Naive Bayes, Support Vector Machine(SVM) and Convulation Neural Network(CNN), and we obtained the results as shown in Table 2. We can observe from the table that using CNN we have significant improvement in the accuracy of the obtained results.

Classifier	Accuracy
Naïve Bayes	68.940
Support Vector Machine	83.620
CNN	98.583

Table 2. Comparison of classification accuracy of different classifier

During the training process of CNN model, the value of the active layer's learning rate was set to 0.0001 for achieving remarkable accuracy in extracting invariant features from character image and different combination of epochs and batch-size was used to get minimum training time with optimal training accuracy.

The plot of training and validation accuracy curve of the CNN retrained with the fixed user defined parameters is presented in Fig. 1. Fig. 2 represented the plot of training and validation loss curve of the CNN. The training and validation accuracy curve shows that the training and validation accuracy is above 96% during the last four epochs while from the training and validation loss curve, it is observed that the train and validation losses were below 1% during the last four epochs.

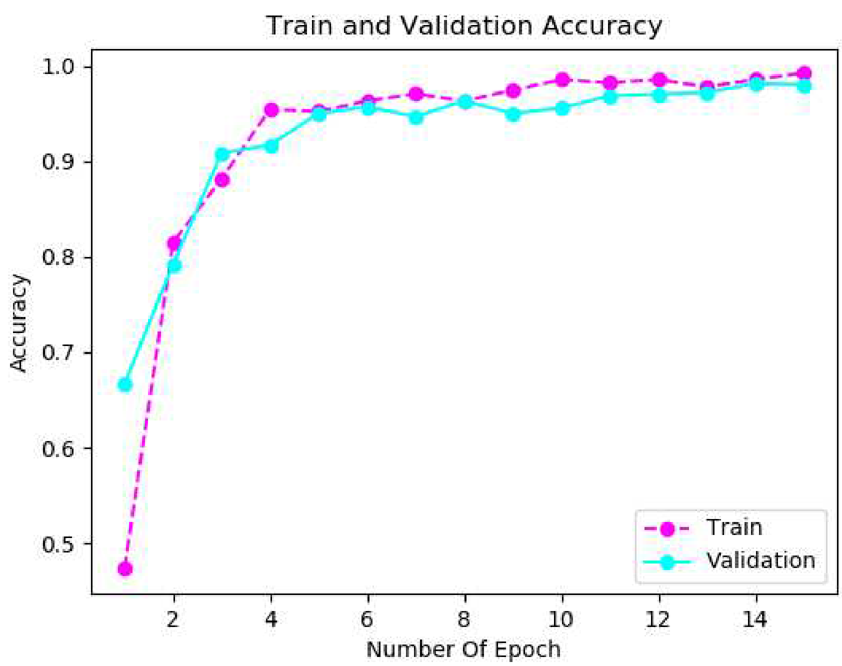

Figure 1. Train and validation accuracy with respect to number of epochs.

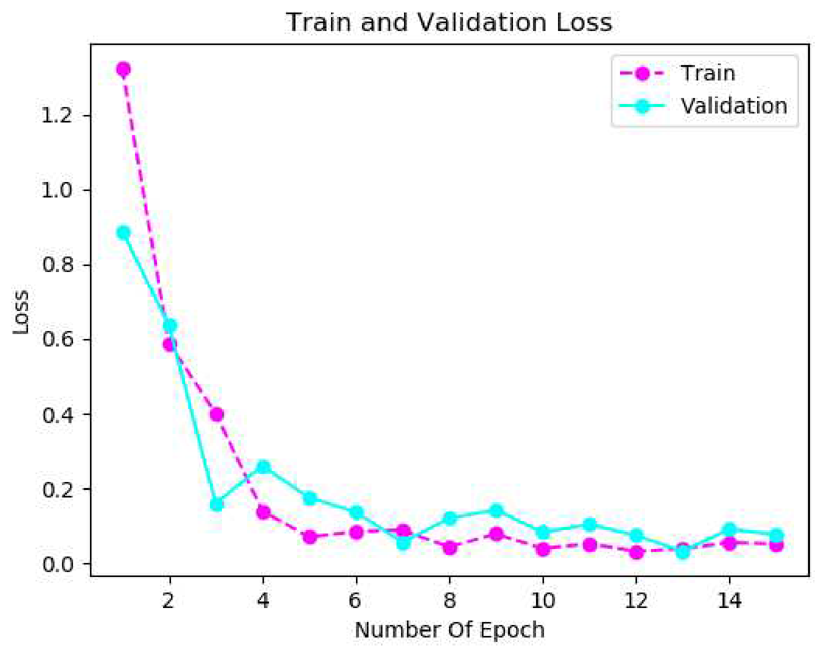

Figure 2. Train and validation loss with respect to number of epochs.

To analyze the performance of the proposed model, we used two matrices namely accuracy and loss, we can observe from Figure 1 demonstrating the accuracy of the model with respect to the number of epoces that accuracy is above 95% and that can be considered as an efficient result. We can observe in Figure 2 demonstrating the loss, loss signifies how well the model is doing in these train and test datasets, it is interpreted from the figure that in our model the loss is around 1% that proves the efficiency of ur model. From the results, we can conclude that that CNN is the best fit method for image classification dataset.

5.Conclusion

In this project, The Handwritten Digit Recognizer has been implemented which is capable of recognizing the number digits of different types of handwritings. The CNN is one of the most widely used machine learning algorithms which has been trained and tested on the dataset in order to compare and analyze.With this deep learning technique, a high amount of accuracy can be obtained. The main objective of this investigation is to find a representation of isolated handwritten digits that allow their effective recognition. In this paper,we used different machine learning algorithm for recognition of handwritten numerals. In any recognition process, the important problem is to address the feature extraction and correct classification approaches. The proposed algorithm tries to address both the factors and well in terms of accuracy and time complexity. The overall highest accuracy 90.37% is achieved in the recognition process by Multilayer Perceptron. 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.

6.References

[1] F. Kimura and M. Shridhar, "Handwritten numerical recognition based on multiple algorithms," Pattern Recognit., vol. 24, no. 10, pp. 969–983, Jan. 1991.

[2] S. Impedovo, L. Ottaviano and S. Occhinegro, "Optical character recognition", International Journal Pattern Recognition and Artificial Intelligence, Vol. 5(1-2), pp. 1-24, 1991.

[3] R.G. Casey and E.Lecolinet, "A Survey of Methods and Strategies in Character Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No.7, July 1996, pp. 690-706.

[4] U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts," Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2007.

[5] Rafael C. Gonzalez, Richard E. woods and Steven L.Eddins, Digital Image Processing using MATLAB, Pearson Education, Dorling Kindersley, South Asia, 2004.

[6] C. L. Liu, H. Fujisawa, "Classification and Learning for Character Recognition: Comparison of Methods and Remaining Problems", Int. Workshop on Neural Networks and Learning in Document Analysis and Recognition, Seoul, 2005.

[7] N. K. Garg, D. L. Kaur, and D. M. Kumar, "Segmentation of handwritten hindi text," Int. J. Comput. Appl., vol. 1, no. 4, pp. 22–26, Feb. 2010.

[8] N. Arica and F. Yarman-Vural, "An Overview of Character Recognition Focuse8 on Off-line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2001, 31(2), pp. 216 - 233.

[9] Anshul Gupta, Manisha Srivastava and Chitralekha Mahanta," Offline Handwritten Character Recognition International Conference on Computer Applications and Industrial Electronics (ICCAIE), 2011.

[10] Gupta, J., Chanda, B. Novel Methods for Slope and Slant Correction of Off-line Handwritten Text Word. Third International Conference on Emerging Applications of Information Technology. 2012

[11] D. K. Patel, T. Som and M. K Singh," Improving the Recognition of Handwritten Characters using Neural Network through Multiresolution Technique And Euclidean Distance Metric", International Journal of Computer Applications (0975 – 8887) Volume 45– No.6 May 2012.

[12] Kruse, R., Borgelt, C., Klawonn, F., Moewes, C., Steinbrecher, M., & Held, P. Multi-Layer Perceptrons. In Computational Intelligence (pp. 47- 81). Springer London, 2013.

[13] S. AL-Mansoori, "Intelligent handwritten digit recognition using artificial neural network," International Journal of Engineering Research and Applications, vol. 5, no. 5, pp. 46–51, 2015.

[14] Wu, Ming & Zhang, Zhen. (2019). Handwritten Digit Classification using the MNIST Data Set

[15] Ahmed Mahdi Obaid, IIHazem M. El Bakry, IIIM.A. Eldosuky, IVA.I. Shehab "Handwritten text recognition system based on neural network." Int. J. Adv. Res. Comput. Sci. Technol.(IJARCST) 4.1 (2016): 72-77

[16] S. M. Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana, Abdullah Al Jobair "Handwritten digit recognition using machine learning algorithms." Global Journal Of Computer Science And Technology (2018).