Handwritten Text Recognition App using Python

Abhay Srivatava[1], Mrs. Pooja Khulbe[2], Deeksha M Gautam[3,] Vikas Yadav[4], Virat Raj [5]
[2]*Assistant Professor, Babu Banarasi Das Institute of Technology and Management, Lucknow, India*
[1,3,4,5]*Babu Banarasi Das Institute of Technology and Management, Lucknow, India*

Abstract **- The aim of this research is to offer a new solution or to improvise with the various traditional handwriting recognition techniques. Text recognition is the one in all the emerging fields within the Computer Vision and Deep learning.**
Handwritten text Detection is a technique or ability of a Computer to receive and interpret the handwritten input from a variety of sources such as paper documents, touch screen, photos, etc. The goal of handwriting recognition is to identify input characters or image correctly then analyze to many automated process systems. This will be applied to detect the writings of different format. This automatic recognition of handwritten text can be extremely useful in many applications which currently exists for reading postal addresses, bank check amounts, and forms, etc. where it is necessary to process large volumes of handwritten data and majorly its applications can be far more.
The aim of the project is to improve existing handwritten character recognition problem and making it more precise.

Index Terms **- CNN, Handwritten Text Recognition, MNIST, Keras, Normalization.**

I.INTRODUCTION

Handwritten digits and character recognition are becoming increasingly important in today's digitized world as their practical applications in various day-to-day activities are seemingly increasing. Systems that are used to recognize Handwritten letters, characters, and digits help people to solve some complex task which can otherwise turnout to be time-consuming and even costly.

As of now, the Handwritten text recognition system is a technology becoming a mandatory need in this world.

The handwritten recognition systems can be inspired by biological neural networks, which allow humans and animals to learn and model non-linear and complex relationships. When reading handwriting in a reduced form, neural networks help to imitate how the human brain operates. It enables machines to read handwriting at a level that is comparable to, if not superior to, that of humans. Humans write in a variety of ways, some of which are difficult to decipher.

The challenge of visual pattern recognition is only apparent to develop a computer system to read handwriting. The artificial neural networks approach is considered as one of the best ways to develop systems for recognizing handwriting.

Besides, for a human reading handwriting may be time-consuming and tedious, especially when individuals are required to read several Handwriting documents by different individuals. A neural network is the most appropriate for the proposed system due to its ability to derive meaning from complex data.

Thus, during this project it's determined to tackle the downside as a result of considerably larger easy management of digital text compared to written language can facilitate individuals additional effectively access, search, share, and analyze their records, whereas still permitting them to use their most popular writing technique. This project's goal is to investigate the challenge of identifying written material and converting it to a digital representation.

Before proper implementation of this technology, we were relied on writing texts with our own hands which can result in errors. It's challenging to efficiently store and access physical data. The use of automatic processing systems in banks to handle bank checks is a good example. Without automated bank check processing systems, the bank would be forced to hire a large number of staff who might not be as productive as the computerized system.

The main aim of this paper is to develop a model that will be used to read Handwritten digits, characters, and words from the image using the concept of Convolution Neural Network

II. LITERATURE REVIEW

One of the research fields in computer vision, artificial intelligence, and pattern recognition is handwriting character recognition.

In the development of this project we have used PYTHON version 3 as language to develop the software in order to meet the project requirements as its language constructs. The goal of the language, as well as its object-oriented approach, is to assist programmers in writing clear, logical code for both small and large-scale projects. For the development and exploration, we have used PYCHARM which supports Python and makes the task completion ease.
This paper focuses on Machine learning, Deep learning, Convolutional Neural Networks. Various pre-processing techniques are being involved within the character recognition with different reasonable images ranges from easy handwritten form-based documents and documents containing colored and complicated background and varying intensities.

The proposed system consists of four phases. The first phase is creation of model, wherein sequential model which allows to build a model layer by layer. The layers used are Conv2D, Max pooling, Batch normalization. Feature extraction is a technique for identifying the most essential traits that are not dependent on others. Classification, the process of assigning labels to unseen observation. For classifying, fully connected layers are used.

These paradigms were useful in successful training models and to increase the accuracy of the result. The output will be determined by how much training the model receives in order to improve its output.

III. WORKING MODEL OVERVIEW

Handwritten Text Recognition systems consist of handwritten text in the form of scanned images. We have built a Neural Network to train on word images from the MNIST dataset because the input layer (and therefore also all the opposite layers) are often kept small for word images, NN-training is possible on the CPU.

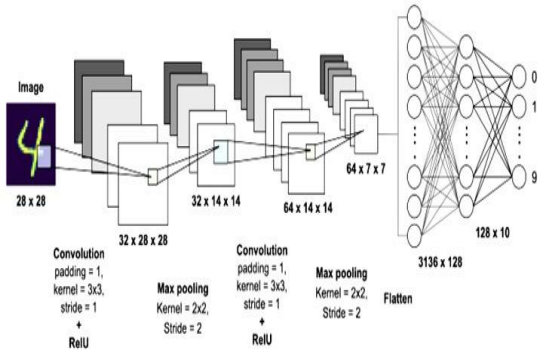

Figure-1MODELING OF CONVOLUTIONAL NEURAL NETWORK TO CLASSIFY HANDWRITTEN DIGITS

Our project aims to make an interface that can be used to recognize users' handwritten characters. We approached our problem using Convolutional Neural Networks to get higher accuracy. The main focus will be providing a GUI that can be used to easily predict characters.
To begin, we'll construct a model that will be trained on the Emnist dataset, which comprises over 690,000 train images, and then validated on the Emnist test dataset. We will work on python to segment characters from an image of a word and predict each character using our model.

Data Pre-Processing
Before training our sequential model with our trained images, we were required to apply some pre-processing techniques to our data to make it more fitting for our model.
Normalization is the process of adjusting the range of pixel intensity values in image processing.
Input image filtering - Here, we discuss various pre-processing techniques that we have applied to get an efficient classification. Firstly, we resize an image in the ratio of 4:5 (using inbuilt python libraries) and convert it into grayscale. This increases the spacing between the printed characters, allowing us to distinguish between characters that were written very close together. Next, we perform thresholding over the image. We could segment an image into foreground and background this way. After this, the image is converted to binary format because contours can be found easily in binary images. We apply the cv2 threshold function for this. We then dilate the image.

This expands its interior objects and increases the chances of efficient character segmentation.

Further, we will apply a gausian blur to our image. As a result, a gaussian kernel is multiplied by each pixel in our image. This reduces noise and detail. Now our image is ready for segmentation.

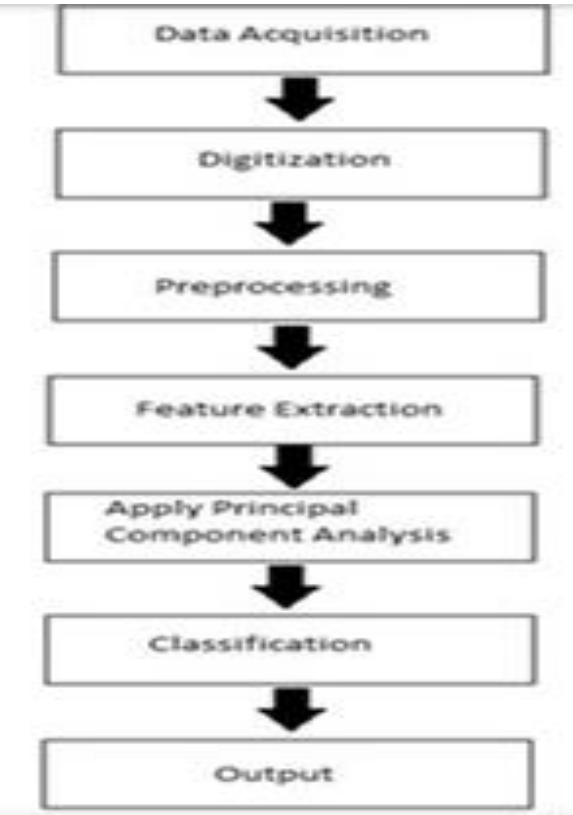

Figure -2 A basic representation in pictorial format on which our project implementation and working depends.

Model

The CNN model for recognizing handwritten characters is constructed using the KERAS library of python. The sequential model consists of a linear stack of layers. The same has been used for developing the handwritten text recognition model. Defining a model was our primary task which generally means adding different layers to the stack. The reshape layer was the first layer we added, and it received an input of (784,1) and reshaped it into (28,28,1). **Layer 2** and **Layer 3**: Convolutional layer - The function of the convolutional layer is to find out distinctive features from the given input matrix which is an image of handwritten text. The convolutional layer comprises a filter matrix made up of zeros and ones. This convolutional layer takes the input matrix of the shape of (28,28,1). This layer uses a filter of size (5,5) and a stride size of 1. It uses ReLU as the activation function. The purpose of using the convolution layer twice is to extract more features to improve the model's accuracy. **Layer 4**: MAX POOLING layer - Max pooling is done by applying a max filter to non-overlapping sub-regions of the initial representation. When the photos are too big, the number of trainable parameters must be adjusted. Pooling is done solely for the aim of shrinking the image's spatial size. **Layer 5**: FLATTEN layer - The function of the flatten layer is to convert all elements of the feature maps matrices to individual neurons which will serve as input to the next layer. **Layer 6**: DENSE layer (input layer) which accepts the output from the FLATTEN layer as input. The value held by a neuron is called the activation of that neuron. Every unit of input(neuron) has activation corresponding to the intensity of the pixel. The output of this layer is determined using the activation function (RELU). The activation function's purpose is to power on the DENSE layer's neurons. **Layer 7**: Dropout layer - The function of the dropout layer is to remove some of the neurons i.e. unwanted features that can make the model bulky and increases the training time. It is also helps in avoiding overfitting. **Layer 8**: Output layer (Last layer) - The last layer has 62 neurons corresponding to 0-9 numbers and A-Z alphabets in uppercase as well as lower case. The neurons are activated similarly, but this time SOFTMAX is used as the activation function. SOFTMAX is a logistic classification function that is similar to the SIGMOID function

Figure-3

IV. DISSCUSSION AND ANALYSIS

In this project, we will feed an image as an input and then it predicts the output by loading the model which we already created. We trained our model with

different optimizers available for Keras. The dataset has sample Handwriting digits for evaluating machine learning models on the problem of Handwritten digit and character recognition. As observed from the results of the experiment, CNN proves to be far better than other classifiers. The following steps were performed for the classification on Handwriting digits dataset:

First, Load the datasets of Handwriting digits and characters for classification. Then, Split the datasets into two sets; one for training and the other for testing. After Splitting, a recognizer was trained to predict the given image of Handwritten text. At last, Testing the accuracy of the classifier.

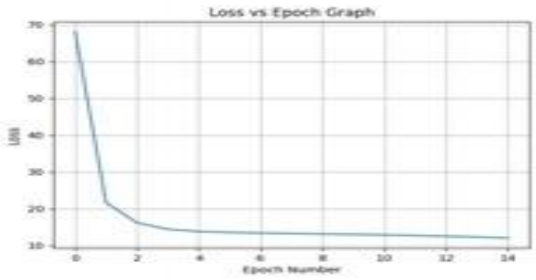

Figure-4

The text can be a single or a word. We use OpenCV to work with images for this research. we find contours (curve joining all the continuous points having the same color or intensity in the image). After finding contours, rectangular bounding boxes were created around each character in a copied image to avoid overlapping.

The dataset was split into two halves - half of the dataset as the training set and the remaining half as testing sets The system can show one or more than one-word recognition.

The result shows that the decision tree classifier turned out to be effective in the recognition of Handwritten data. It will be an indication that the system is well trained. Although it also shows that some of the digits were not recognized by the model.

S.No	Number of Observations from running software	Success attempts recorded [accuracy- 70% to above]	Failure attempts recoded [accuracy- upto 60%]	Approximation percentage(%)
1	3	2	1	66
2	5	4	1	71
3	7	5	2	80
4	9	7	1	77

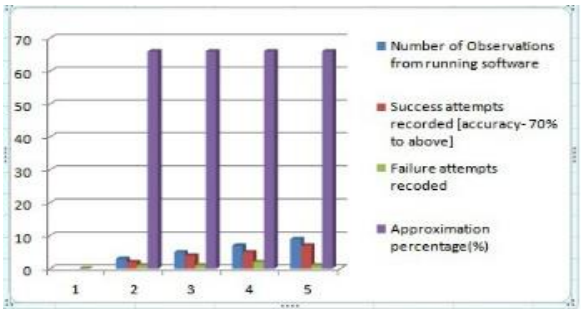

Figure-5

In the input image out of 7 characters, our model has recognized and processed 7 characters out of which 6 were correctly recognized and 1 was faulty. Hence with an accuracy of 92% for this input and overall accuracy of 75%.

Figure-6

Representation of a set of digits to be trained to our model

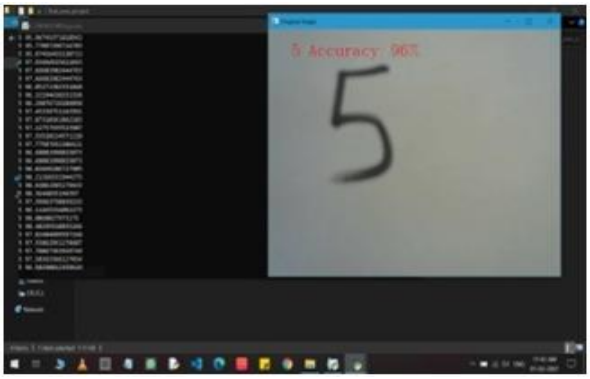

Figure -7 Represents the accuracy of the recognized digit

V. CONCLUSION

The main aim of this research was to develop a model that would be able to provide improved and unique methods for recognition of handwritten characters and digits. Handwritten text recognition is a complex problem. The model is built using modern day

techniques like neural networks to implement deep learning to analyze the written text and convert it in computer text.As seen from the results of the experiment, CNN proves to be far better than other classifiers.The MNIST dataset provides a higher number of image samples and output classes and an even more complex and varied classification task. It was thus obvious to use it as the backbone of our project. Without the use of EMNIST data set it would be practically impossible to achieve this accuracy. The final system satisfied the specified requirements of accuracy as well as recognition.This system will provide both the efficiency and effective result for the recognition.

The results can be made more accurate with more convolution layers and a greater number of hidden neurons. This app can be helpful in many health and consumer sectors.

VI. FUTURE SCOPE

Text recognition in pics is an active research subject that aims to create a computer application that can read text from images automatically. There is a much need nowadays for storing information from paper documents in a computer-readable format for subsequent use. Scanning the documents and then collecting them as images is a simple approach to store information from these paper documents in a computer system. However to reuse this information it is very difficult to read the individual contents and searching the contents form these documents line-by-line and word-by-word. The font properties of the characters in paper documents, as well as the quality of the photos, are the obstacles. The computer is unable to recognize the characters while reading them as a result of these difficulties.

There can many development in the field of text recognition and now also there are many application then also the requirements of humankind is not satisfied either by the result or to be specific the accuracy of the result. Here we are proposing that in future updates we will make this app as accurate not only for digit but also for character's and if possible we will try to introduce materialistic recognitions.

VII. ACKNOWLEDGMENT

We thank our research organization for supporting us during the course of this work. We also thank our guide Mrs. Pooja Khulbe for her remarkable guidance and support throughout. We'd like to thank our parents for their financial and ethical assistance during the study process. Finally we like to acknowledge our efforts in accompanying each other at work.

REFERENCES

[1] Jayaram Kinthali, Kowshik Kotha, Kanithi Santosh Kumar, Jagilinki Jayalaxmi "Handwritten Text Recognition using Deep Learning with TensorFlow" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV9IS050534. Published by: ijert.org Vol. 9 Issue 05, May-2020

[2] Polaiah Bojja, Naga Sai Satya Teja Velpuri, Gautham Kumar Pandala,S D Lalitha Rao Sharma Polavarapu , Pamula Raja Kumari "Handwritten Text Recognition using Machine Learning Techniques in Application of NLP" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019

[3] Kumar D.A., Sastry A.S.C.S., Kishore P.V.V., Kumar E.K., Kumar M.T.K., "S3DRGF:Spatial 3-D relational geometric features for 3-D sign language representation and recognition", IEEE Signal Processing Letters, ISSN:10709908, Vol No:26, Issue No:1, 2019, pp: 169 - 173, DOI: 10.1109/LSP.2018.2883864

[4] Ramesh C., Venkat Rao D., Murthy K.S.N., "Short note on the application of compressive sensing in image restoration", Smart Innovation, Systems and Technologies, ISSN:21903018, Vol No:105, 2019, pp:267 - 273, DOI: 10.1007/978-981-13-1927-3_28

[5] G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, "Emnist: an extension of mnist to handwritten letters," arXiv preprint arXiv:1702.05373, 2017.

[6] Manoj Sonkusare and Narendra Sahu "A Survey on Handwritten Character Recognition (Hcr) Techniques For English Alphabets" Advances in Vision Computing: An International Journal (AVC) Vol.3, No.1, March 2016.

[7] [7] Fischer, A., Frinken, V., Bunke, H.: Hidden Markov models for off-line cursive handwriting

recognition, in C.R. Rao (ed.): Handbook of Statistics 31, 421 – 442, Elsevier, 2013

[8] Bhatia Neetu, "Optical Character Recognition Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 5, May 2014.