



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Hand Written Text Recognition Using Machine Learning

Mahajan Rachana Dhananjay¹, Chandwani Reema Sushil², Patil Yogita Devidas³, Bhavsar Pratik Shamkumar⁴, Phulwani Harsh Nanakram⁵

U.G. Student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India¹

U.G. Student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India²

U.G. Student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India³

U.G. Student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India⁴

U.G. Student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Bambhori, Jalgaon, India⁵

ABSTRACT: The Purpose of the Handwritten Text Recognition is to recognize the handwritten text and to convert handwritten text into digital format efficiently and with accuracy. The problems are faced by several peoples in recognizing hand-written text and hence, the aim of the system is to convert the handwritten text into digital format. The system takes image as an input and with the help of CNN (Convolution Neural Network) system is trained to produce expected output. The character to character segmentation helps in generating output. The proposed system is algorithm-based and fully functional model helping user to solve the problem of converting handwritten text into digital format. The system provides highest accuracy possible to avoid any errors in digital text generation.

KEYWORDS : Handwritten Text Recognition, Digits, Manual, Error-free, Accuracy, CNN, Recognition, MNIST dataset

I. INTRODUCTION

The main objective of this proposal is to recognize the hand-written text and to convert handwritten text into digital format efficiently and display digital text on computer where input given will be in a form of image containing text in English cursive or block writing. The system is algorithm-based and fully functional model helping user to solve the problem of converting handwritten text into digital format. The aim of the text recognition can be achieved using various techniques which helps in translation of images into machine editable form, CNNs [1] algorithm. The system is supposed to extract characters individually and classifies them. It provides highest accuracy possible to avoid any errors in digital text generation. Though various Writing tools and editors are available but handwritten text have their own importance in the field of medical, Banks and Post-office, etc. To ensure readability and efficient understanding of text it is necessary to convert handwritten text into digital format with the help of proposed project which will generate Error-free information.

II. HAND WRITTEN TEXT RECOGNITION SYSTEM

In the paper the emphasis is on machine learning technique to predict the Hand written Text using the CNN Algorithm. The system takes image as an input and with the help of CNN (Convolution Neural Network) system is trained to produce expected output [8]. The input is given in form of image on which Pre-processing step is performed. It essentially enhances the image rendering it suitable for segmentation. The role of pre-processing is to segment the

interesting pattern from the background. Generally, noise filtering, smoothing and normalization should be done in this step. The pre-processed input image is segmented into isolated characters by assigning a number to each character using a labelling process[7]. This labelling provides information about the number of characters in the image. Each individual character is uniformly resized into pixels. In the feature extraction stage each character is represented as a feature vector, which becomes its identity[4]. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements. Handwritten recognition is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices. To acquire correct, error-free and enhanced data, necessity is to develop a system which can convert manual hand-written data into digital format. Users should be able to easily provide input and get the output with utmost accuracy[6]. Handwritten text recognition system is a key to fetch data or information effectively. Also, it saves physical storage, provides quality of data and saves time to understand data.

III. RELATED WORK

The research aims over recognition[2] of Hand written text which helps people to recognise the text. The effectiveness of the method that uses feature extraction using character geometry and gradient technique from scanned images containing handwritten characters is presented. The feature extraction methods have performed well in classification when fed to the neural network and pre-processing of image using edge detection and normalization are ideal choice for degraded noise image[3]. The method of training neural network with extracted features from sample images of each character has detection accuracy to greater extent.

IV. METHODOLOGY

The proposed system recognizes handwritten digits and characters using Convolution Neural Network [CNN][5]. Convolution Neural Network is an algorithm used to analyze image. A mathematical operation called Convolution works on two functions that produces a third function that shows how the shape of one is modified by the other. Convolution provides filter to an input resulting in an activation. In project Convolution Neural Network is designed with 2-dimensional image.



Fig 1 : Convolution Neural Network

Steps Involved:

- Acquisition: The first step is to load the dataset. After loading the dataset, retrieval of the image from the source is known as acquisition. Then the image is used for further processing. After this image is preprocessed to remove noisy data.
- Segmentation: Segmentation is used to divide image into number of segments. It provides labels to the data which helps in easy location of image and also provide object type of the data. In Semantic segmentation pixels of image are mapped
- Feature extraction: Feature extraction is an important step in construction of model. Features are various patterns of objects. Feature extraction works on pre-processed image and then finds the different variability in patterns.
- Prediction: After performing all the above steps, image is reduced, noisy data is removed and the data is organized. After building a sequential model using Convolutional Neural Network, accuracy is also measured. The model is then used to predict the handwritten characters and digits.

V. RESULT AND DISCUSSION

Handwritten Text Recognition system is to recognize the handwritten text and to convert handwritten text into digital format efficiently and with accuracy. The problems are faced by several people in recognizing handwritten text and hence, the aim of the system is to convert the handwritten text into digital format. The system provides highest accuracy possible to avoid any errors in digital text generation. Though various Writing tools and editors are available but handwritten text has their own importance in various fields. So, this system is designed to give effective result.

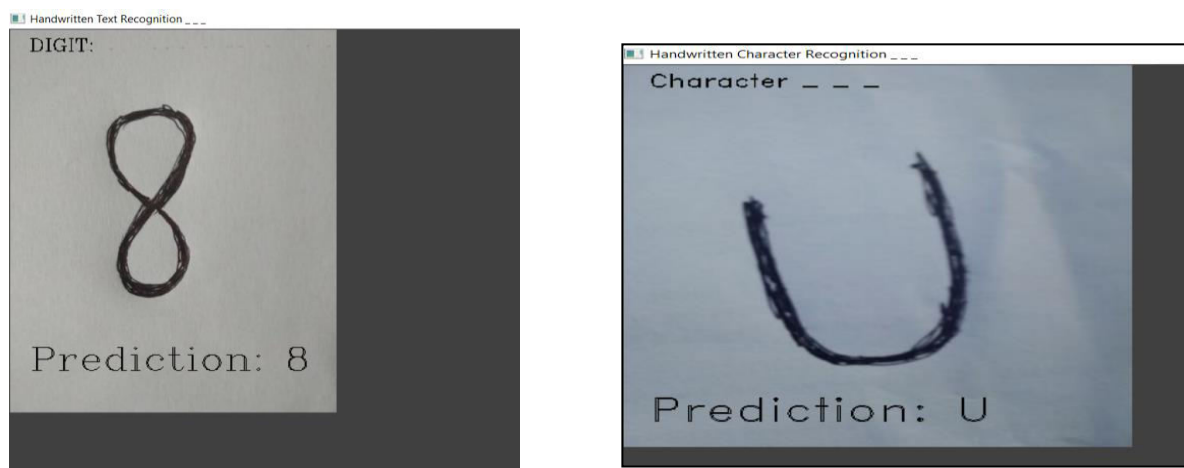


Fig 2 :Output of Text Recognition System

VI. CONCLUSION

The research aims at recognizing hand written text and convert it to digital format. The performance of CNN for handwritten recognition performed significantly. The proposed method obtained 90% accuracy. The performance of CNN for handwritten recognition performed significantly. The project provides highest accuracy possible to avoid any errors in digital text generation. User will be able to get digital text by giving image as an input and can execute further task effectively. This system is successfully developed by using Python, Tensorflow, Numpy, Pandas, Matplotlib libraries. All the tools used in this project are open source and easy to use, by which the proposed system is feasible in terms of resources, effort, and time.

VII. FUTURE WORK

The future enhancement of our system is to identify the handwritten text of other languages and text have been recognized with more accuracy. After working on digits and characters the future work will be that system should be able to predict words also. The main aim would be to make accuracy of system as high as possible.

REFERENCES

- [1] Convolutional Neural Network Benchmarks: <https://github.com/jcjohnson/cnn-benchmarks>
- [2] K. Simonyan, A. Zisserman Very Deep Convolutional Networks for Large-Scale Image Recognition arXiv technical report, 2014
- [3] U. Marti and H. Bunke. The IAM-database: An English Sentence Database for Off-line Handwriting Recognition. Int. Journal on Document Analysis and Recognition, Volume 5, pages 39 - 46, 2002.
- [4] <https://www.pyimagesearch.com/2020/08/24/ocr-handwriting-recognition-with-opencv-keras-and-tensorflow/>
- [5] <http://cs231n.stanford.edu/reports/2017/pdfs/810.pdf>
- [6] <http://cs231n.stanford.edu/reports/2017/pdfs/810.pdf>



- [7] Yellapragada SS Bharadwaj, Rajaram P, Sriram V.P, Sudhakar S, KollaBhanuPrakash ,” Effective Handwritten Digit Recognition using Deep Convolution Neural Network” International Journal of Advanced Trends in Computer Science and Engineering, 9(2), March
- [8] <http://www.warse.org/IJATCSE/static/pdf/file/ijatcse66922020.pdf>



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details