

Survey on Handwritten Character Recognition Machine Learning

Shweta Changan, Gauri Mahajan, Ketan Bobade, Yash Darekar, Sonali Patil, Chandrakant Kokane

Nutan Maharashtra Institute of Engineering and Technology, Pune, Maharashtra, India

ARTICLE INFO

Article History :

Accepted: 10 Nov 2023

Published: 30 Nov 2023

Publication Issue :

Volume 10, Issue 6

November-December-2023

Page Number :

239-242

ABSTRACT

In the fields of pattern recognition and image processing, handwriting recognition has emerged as the most captivating and difficult subject. For word recognition, a number of strategies were used. The majority of approaches are used for simple papers. The suggested system offers methods for identifying handwritten English words from photos of real-world scenes. The CNN classifier-based holistic word recognition approach was employed by the system. Since CNN is taught to recognize words as a whole, it does not need to recognize individual characters. The suggested system is trained and tested using the VGG synthetic word dataset.

Keyword: Hand written, Pattern detection, ML

I. INTRODUCTION

To create a model capable for classifying and recognizing handwritten numbers. With the help of their eyes and brains, humans are able to see and perceive the visual environment. The goal of computer vision is to give machines the same sense of perception and processing capabilities as human vision. Numerous methods for picture recognition were created in the field of computer vision. Our project aims to develop a model with improved accuracy for recognizing and identifying handwritten digits from their images.

The goal of this work is to gain knowledge and experience with Convolutional Neural Network ideas. In the field of example order, transcribed digit acknowledgment has long been a contentious issue. A

few tests have shown that a neural network performs incredibly well when it comes to information organization. This paper's main goal is to provide robust and efficient methods for recognizing transcriptions by examining various existing arrangement models. This paper discusses the Convolutional Neural Network (CNN) exhibition.

The results show that CNN classifier formed Neural Network with significantly higher computational effectiveness without sacrificing performance. Machine Learning's Convolutional Neural Network can be used to recognize handwritten digits. In essence, several libraries including NumPy, Pandas, TensorFlow, and Keras were needed to run the model. The significance of the convolutional neural network is the topic of this discussion. The division of dataset

into training and test datasets was also covered. To predict handwritten digits from 0 to 9, a dataset was used. The dataset underwent shaping, scaling, and cleaning. A CNN model was developed with TensorFlow and trained on the training dataset.

II. PROBLEM STATEMENT

For nearly four decades, the focus of study has been on handwriting recognition. In order to forecast a digit, this research study examines how classification algorithms (CNN) behave in a sizable handwriting dataset. Machine-learning approaches have become more and more significant in design of these recognition systems, especially when applied to Neural Networks such as CNN/ANN. For nearly past 40 years, handwriting recognition has been primary focus of study. The goal of this research project is to predict a digit by analyzing the behavior of classification algorithms (CNN) on huge handwriting dataset. The design of these recognition systems has become more and more dependent on machine-learning techniques, especially when applied to Neural Networks such as CNN/ANN.

Our effort aims to develop a model employing CNN ideas that can identify and categorize handwritten numbers from photographs. While developing model for digit identification and classification is the main objective of our work, it may also be applied to letters and a person's handwriting. The primary objective of the suggested system is to comprehend Convolutional Neural Networks and use them to system for recognising handwritten numbers by utilizing the generated dataset.

III. RELATED WORK

A support vector machine (SVM)-based real-time handwritten digit classification system was created by Ahmed Hafiz (2018), however it is not appropriate for large data sets.

A Handwritten Digits Classification using Back Propagation Network was proposed by Anuj Dutt (2017). The fundamental process of neural network training is backpropagation. This technique involves adjusting a neural network's weights according to the error rate recorded in the preceding epoch. It is less effective, though, by using Keras and Theano as the backend for a convolutional neural network.

In an effort to minimize error rates in handwriting recognition, Denker J.S. (2019) designed Handwritten Digits Recognition. In 1 study, 3-NN trained and tested on MNIST yielded an error rate of 1.19%. The multimodal neural architecture known as the Coherence Recurrent Convolutional Network (CRCN) is used. It is employed that we should retrieve sentences from image. To get around shortcomings of convolutional layers, some academics are working for developing novel methods. Using MNIST datasets, one strategy that may be used for improved performance is NCFM (No combination of feature maps).

Haider (2020) created a brand-new, difficult Arabic dataset that was gathered from various school study levels. After putting in great effort to distribute and gather digital forms from hundreds of elementary, high school, and college students, a sizable dataset was gathered. He put lot of effort into creating a difficult Arabic digit dataset after seeing that there were few and undemanding datasets available.

Mukesh N. (2018) carried out a Handwritten Character Classification. Using k-means clustering machine learning algorithms, it has high baseline error and low accuracy, making it unsuitable for large data sets.

An artificial neural network-based method for handwritten digit recognition was proposed by Nitin Kali Raman (2021). Artificial neurons, that resemble neurons in biological brain somewhat, are group of interconnected units/nodes that form the foundation of an ANN. Similar to the synapses in a living brain, every link has ability to communicate with other neurons. After processing a signal, an artificial neuron

is able to communicate with other neurons that are connected to it. It was a High Baseline Error, though.

A dataset and recognition system for handwritten digits was created by Plamondon, R. (2018) and trained on both convolutional and artificial neural networks. The accuracy evaluation metric was used to determine the average error for both networks. CNN has lower average error rate than a CPU-based artificial neural network. Nevertheless, CNN training on a CPU required less time than training on the artificial neural network. However, CNN does a superior job in image classification. In summary, the accuracy of recognition increases as the model is trained with CNN; however, training on a GPU can yield the best results for CNN classification.

A Multilayer Perceptron (MLP) Neural Network was proposed by Saeed Mansoori (2020) to identify and forecast handwritten numbers from 0 to 9. One fully connected type of feedforward artificial neural network (ANN) is the multilayer perceptron (MLP). The term "MLP" is used in an imprecise manner; it can refer to any feedforward artificial neural network (ANN) or, more precisely, to networks made up of many layers of perceptrons that are activated by threshold. A dataset obtained from MNIST was used for training and testing the suggested neural system.

Gaussian Naive Bayes was used by Shamim S.M. (2018) to introduce machine learning algorithms for handwritten digitizer recognition. A Naive Bayes variation that supports continuous data and adheres to the Gaussian normal distribution is called Gaussian Naive Bayes. This has the low precision as a downside. Via the MNIST dataset, Shyam R. (2017) performed Handwritten Digits Classification and found that deep networks perform better when trained via straightforward back-propagation. But in contrast to NORB and CIFAR10, their architecture yields the lowest error rate on MNIST.

Sonia Flora (2016) created a support vector machine (SVM)-based handwritten Digits Classification system. Support vector machines evaluate data for

regression and classification using supervised learning models and related learning methods. but achieved a lower level of accuracy than convolution neural networks (CNNs). It provides less accuracy because it is a vast dataset.

Handwritten Character Recognition Methods and Techniques were implemented by Vinjit B.M (2020). When comparing the accuracy of keras with Theano and Tensorflow to the most popular machine learning methods, the results are least favorable. Convolutional Neural Networks (CNNs) are widely employed in image classification, video analysis, and other applications despite having the lowest accuracy.

IV. CONCLUSION

The project's original goal was to provide users with a platform on which to do OCR. The OCR web application, which allows users to extract text from photographs in an edible format, is shown in the proposed project. Different papers and materials are referred to propose methodology of this project.

V. REFERENCES

- [1]. Arif R.B, Siddique A.B, M. M. R. Khan and Z. Ashrafi, "Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Neural Network Algorithm," in 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEICT), 2018, pp. 118-123: IEEE.
- [2]. Adwait Dixit , Ashwini Navghane, Yogesh Dandawate, "Handwritten Devnagari character Recognition using Wavelet Based Feature Extraction and Classification Scheme", 2014, Annual IEEE India Conference (INDICON).
- [3]. Ashutosh Aggarwal , Karamjeet Singh, kamalpreet Singh, "Use of Gradient Technique for extracting features from Handwritten

- Gurmukhi Characters and Numerals”, 2014, International Conference on Information and Communication Technologies (ICICT 2014).
- [4]. Akanksha Gaur, Sunita Yadav, “Handwritten Hindi character Recognition using K-Means Clustering and SVM”, 978-1-4799- 5532-9/15/\$31.00 ©2015 IEEE.
- [5]. Archana N.Vyas, Mukesh m.Goswami, “Classification of Handwritten Gujarati Numerals”, 978-1-4799-8792- 4/15/\$31.00 ©2015 IEEE.
- [6]. Denker J. S Gardner W. R., Graf, H. P., Henderson, D., Howard, R. E., Hubbard, W., Jackal, L. D., Baird, H. S., and Guyon, I. (1989). Neural Network Recognizer for Hand-Written Digits. In Touretzky, D., editor, Neural Information Processing Systems, volume 1, pages 323-331, Denver, 1988.
- [7]. Kokane, Chandrakant D., and Sachin D. Babar. "Supervised word sense disambiguation with recurrent neural network model." Int. J. Eng. Adv. Technol.(IJEAT) 9.2 (2019).
- [8]. Kokane, Chandrakant D., Sachin D. Babar, and Parikshit N. Mahalle. "Word Sense Disambiguation for Large Documents Using Neural Network Model." 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2021.
- [9]. Kokane, Chandrakant, et al. "Word Sense Disambiguation: A Supervised Semantic Similarity based Complex Network Approach." International Journal of Intelligent Systems and Applications in Engineering 10.1s (2022): 90-94.
- [10]. Kokane, Chandrakant D., et al. "Machine Learning Approach for Intelligent Transport System in IOV-Based Vehicular Network Traffic for Smart Cities." International Journal of Intelligent Systems and Applications in Engineering 11.11s (2023): 06-16.
- [11]. Kokane, Chandrakant D., et al. "Word Sense Disambiguation: Adaptive Word Embedding

with Adaptive-Lexical Resource." International Conference on Data Analytics and Insights. Singapore: Springer Nature Singapore, 2023.

Cite this article as :

Shweta Changan, Gauri Mahajan, Ketan Bobade, Yash Darekar, Sonali Patil, Chandrakant Kokane, "Survey on Handwritten Character Recognition Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 6, pp. 239-242, November-December 2023.

Journal URL : <https://ijsrset.com/IJSRSET2310618>