DOI: https://doi.org/10.46632/daai/2/4/5

Data Analytics and Artificial Intelligence
Vol: 2(4), 2022
REST Publisher; ISBN: 978-81-948459-4-2
Website:http://restpublisher.com/book-series/data-analytics-and-artificial-intelligence/

Handwritten Character Recognisation Using Machine Learning

***KalathurBharath, R. Naven Kumar, P. Vetheshwaran, V. Harishkumar**
Vel Tech HightechDr.RangarajanDr.Sakunthala Engineering CollegeChennai,Tamil Nadu, India.
*Corresponding author Email: bharath2k22@gmail.com

Abstract. This electronic document is a “live” template and already define the components of your paper [title, text, heads, etc.] in its style sheet. For people, it is common and easy to recognize a character, but it is challenging to build a machine that can do the same. One of those abilities that some individuals just naturally have is the ability to recognize characters. Computers and other contemporary technology fall short of human brain power. The three types of algorithms used in character recognition are feature extraction, classification, and image pre-processing. They are frequently employed because feature extraction is easier with image pre-processing, and good classification depends on feature extraction. identifying handwritten letters and the English alphabet from A to Z. Making a neural network that needs to be trained on a dataset of letter images will be used to do this. First the input image is converted into a gray-scale image and normalized in such a way that it represents the same resolution (28 x 28) as that of EMNIST dataset. CNN is trained using EMNIST dataset and use it as a classifier for better results. The feature vectors are extracted from the input image and provided to the trained model Convolution Neural Network is a type of neural network that recognizes and produces the required output.
Keywords: EMNIST, CNN, Machine Learning

1. Introduction

The Convolutional Neural Network automatically extracts the features since it has been programmed to do so. A normalised output layer and several hidden layers will both be present. The input is transferred many times to the hidden layer, which also normalises the output layer. Machine learning offers several ways to recognise manually typed characters with less effort from humans. Deep Learning is a machine learning technique that teaches computers to do what comes naturally to people: learn by doing. Human efforts in seeing, learning, recognising, and many other areas can be reduced with the use of deep learning techniques. The computer learns to perform categorization tasks from images or content from any document using deep learning. Character recognition in handwriting has been a thing since the 1980s. The task of reading handwritten characters using a classifier has a wide range of applications, including online character recognition on PC tablets, reading mail, and processing bank check amounts in word. The difficulties in trying to solve this problem are numerous. The size, thickness, orientation, and placement of the handwritten characters in relation to the margins are not always the same.Ease of UseDeep-Learning Handwritten Character Recognition At the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), which took place on September 27, were Rohan Vaidya, Darshan Trivedi, SagarSatra, and Professor MrunaliniPimpale. The amount of computational power needed to train a neural network has increased due to the accessibility of GPUs and other cloud-based services like Google Cloud Platform and Amazon Web Services that provide resources to do so. On the basis of picture segmentation, we developed a method for reading handwritten characters. In our system, we trained neural networks using TensorFlow and used OpenCV for image processing. Python was used as the programming language to build this system. Maintaining the Integrity of the Specifications. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.A method that uses deep learning to recognise handwritten text when there is text that has been struck out, HiqmatNisa, James A. Thom, Vic Ciesielski, and RuwanTennakoon will present at the 2020 International Conference on Image and Vision Computing in New Zealand on January 16. The effectiveness of a popular handwritten text recognition system, Convolutional Recurrent Neural Network, was tested on lines with words that had been struck out (CRNN). On the words on a text line, some common types of struck-out strokes were used for this. An algorithm trained on the IAM line database was put to the test on lines with words that were struck out. Character Error Rate (CER) increased, going from 0.09 to 0.11. This model was retrained on a dataset including text that was struck out. The model performed superbly in terms of identifying struck-out text.Using deep learning to recognise handwritten characters produced by machines, 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA), 20 July 2017, Kian Peymani, Mohsen Soryani.Character recognition is one computer vision application where Deep Convolution Neural Networks have excelled. However, these networks (in certain cases) lack generality and need a lot of data to learn properly. To solve this problem, we offer a special dataset and a precisely crafted model that can only be rtheless detect handwritten characters with a respectable degree of accuracy.Handwritten Text Recognition Platform Based on Machine Learning Tensor Flow? Classifiers use convolutional neural networks to enhance the feature of the input image. To disseminate information, a recutrained on photos of machine-generated characters in different typefaces, yet neve rrent neural network is employed. The probability distribution of the character at each location in the image is contained

in the recurrent neural network outputs matrix. The connectionist temporal categorization function is used to decode the matrix into the final text. Spelling mistakes are present in the text that has been decoded and are carried over into the processed text. As a result, handwritten writing is more accurate and can be used to other areas, such security systems. Almost all institutions and governments produce a sizable amount of handwritten paperwork each day.

2. Existing System

You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar. The most widely adopted standard for handwritten digit recognition is MNIST. For numbers, recognition is performed. Background subtraction is used to reduce noise during pre-processing. The feature vectors from the handwritten images are extracted using the local gradient feature descriptors. K-nearest neighbor is used for classification. There is no doubt that this is a highly challenging topic given the wide variation in handwriting amongst individuals. Even while this distinction does not cause any problems for individuals, teaching computers to understand standard handwriting is becoming more and more challenging. Understanding how information is represented on images is crucial for the image recognition problem, such as handwritten classification. Scientists are familiar with Handwritten Recognition from the MNIST dataset because, by using different classifiers for different parameters, the error rate has been reduced. For instance, a board of 35 convolution neural systems was able to reduce the error rate from a linear classifier (1-layer NN) with 12% to 0.23%. The goal of this is to construct a Handwritten Character Recognition framework and consider various classifiers and strategies in order to achieve performance that is as close to human as possible.

3. Proposed System

Identifying handwritten letters and the English alphabet from A to Z. This will be accomplished by simulating a neural network that must be trained on a dataset of photographs of alphabets. The input image is first turned into a grayscale version and normalized so that it has the same resolution (28 x 28) as the EMNIST dataset. For better outcomes, CNN uses the EMNIST dataset, which is used to train its classifier. The trained model of the convolution neural network receives the feature vectors from the input image and uses them to recognize and produce the desired results.

4. System Requirements

Generally speaking, these are the conditions for completing the project. We cannot complete this job without using the tools and software mentioned. Therefore, the project has two requirements. need for hardware the hardware specifications should be an exhaustive and consistent description of the entire system since they may form the basis of a contract for the system's implementation. they serve as the foundation for system design for software engineer. ram: 2gb processor: dual core intel i3 os:windows10 Software requirements are outlined in the system's pacification document. Both a definition and a list of prerequisites should be included. The software requirements specification is built on top of the software requirements. Python 3.8: Python is a high-level programming language that is widely used on a global scale. Software engineers may convey concepts in less code thanks to its language structure, which was primarily designed for prominence on code. Python is a programming language that enables you to work more quickly and efficiently coordinate frameworks.

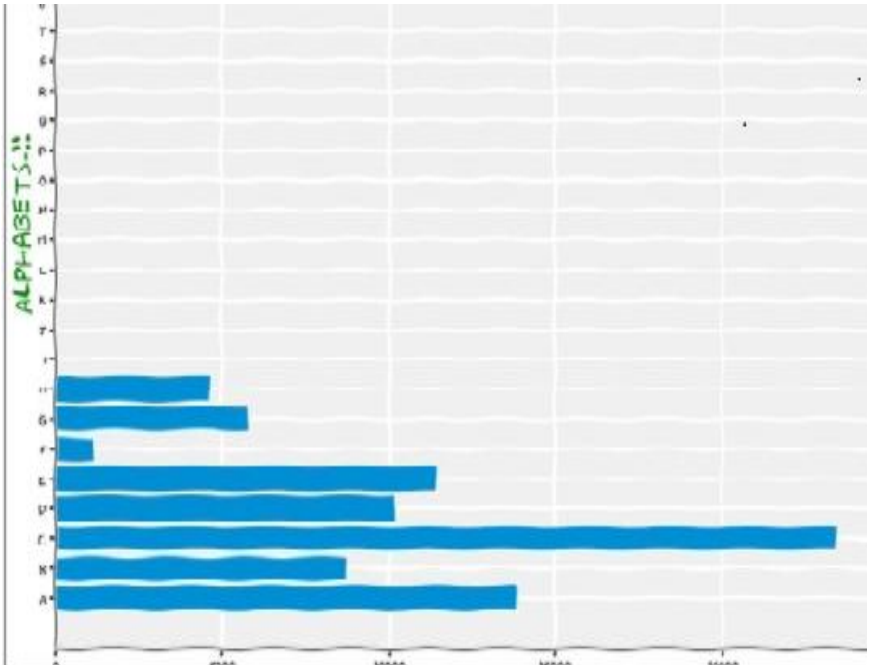
5. Keywords and Definition

System architecture: The system architecture depicted in figure 5.1 exemplifies the general workflow for cnn-based handwritten character recognition. The character is identified from the input. In cnn, the feature is retrieved from each frame and the dimension is shrunk. input dataset loading and learning The emnist dataset is a collection of handwritten character digits taken from the nist special database 19 and transformed into an image format and dataset structure that exactly matches the mnist dataset (28x28 pixels). The data can be loaded most easily with keras. Training and testing data are included in the mnist dataset. The english alphabet and numbers from the emnist dataset are used to train the neural network. 131,600 character pictures across 47 classes make up the emnist balanced dataset. Preprocessing The given image is converted into a gray-scale image as part of the pre-processing step. Typically, a regular coloured image is made up of three channels: the rgb (red, green, and blue) channels. In order to remove undesirable noise from the image, the coloured image is then transformed to a grayscale image, which consists of a single monochrome channel. The size of the given input image would vary, which could result in a loss of accuracy in prediction when the image is compared to that of a convolutional neural network that has been trained. In order for the image resolution to match that of the emnist dataset, the image is scaled and placed on a 28 × 28-pixel blank image. conventional layer for handwritten character recognition, features like hog, sift, surf, and lbp are extracted using convolutional neural networks (cnn). The features are extracted by defining filter, stride, and padding. The pre-trained model is used to initialise the cnn parameters. There are 25 layers in the model. Other than the relu (rectified linear unit) layer, normalization layer, maximum pooling layer, dropout layer, and classification layer, it primarily has five convolutional layers, three fully connected layers, and a softmax layer. Because cnns extract the features directly from the images, manual feature extraction is no longer necessary. Cnn combines input data and learned features local receptive fields, shared weights, and pooling are the guiding ideas of cnn. image input layer The raw pixel value of the image is contained in the image input size, which specifies the dimensions of the input image for a cnn. The height, width, and number of color channels of an image all contribute to its size. The image's basic dimensions are 320 x

240. The network's input is fixed at 227 x 227 x 3 during training. convolutional layer Convolutional filters are the first layer that accepts an input signal. The network attempts to label the input signal throughout this procedure. Neurons that are connected to different input image regions make up a convolutional layer. While navigating a picture, a convolutional layer learns the attributes that are localized by these regions. Generic characteristics like edge, color, and texture features make up convolutional features. A filter is a group of weights that are applied to a specific area of the image. The same computation is repeated for each region as the filter advances vertically or horizontally along the input image. It walks in steps, each of which is referred to as a stride. A feature map is formed by a filter that goes along the input and utilizes the same set of weights and bias as the convolutional.

6. Result

Here are the suggested system's implementation specifics. The results of the individual modules are included in this chapter, along with a brief summary of each result depicts the training data set loaded. identifying handwritten letters and the English alphabet from A to Z. This will be accomplished by simulating a neural network that must be trained on a dataset of photographs of alphabets.



After training, the character is recognized. Convolutional Neural Network automatically extracts the features since it has been programmed to do so. A normalized output layer and several hidden layers will both be present. The input is transferred many times to the hidden layer, which also normalize the output layer.

7. Conclusion

To get the right character recognition, many machine learning approaches are applied. This work developed handwritten recognizers and tested them on the EMNIST (Extended Modified National Institute of Standards and Technology) dataset. The results showed improvements could be made to both the recognition performance and training speed. Therefore, by repeatedly doing this, we have succeeded in correctly detecting the character and showing the correct character at the same time. Therefore, based on the formations created and the values in the dataset, the system would be able to recognize the inserted character.

Reference

[1]. Handwritten Character Recognition Using Deep-Learning, Prof. MrunaliniPimpale, Rohan Vaidya, Darshan Trivedi, SagarSatra, 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 27 September 2018.

[2]. A Survey of Character Recognition Using Machine Learning and Deep Learning, 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), August 17, Reya Sharma, Baijnath Kaushik, and Naveen Gondhi.

[3]. A deep learning method for recognizing handwritten writing when there is struck-out text HiqmatNisa, Vic Ciesielski, RuwanTennakoon, and James A. Thom will present at the 2020 International Conference on Image and Vision Computing in New Zealand on January 16.

[4]. Using a deep learning strategy, machine generated to handwritten character recognition, 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA), 20 July 2017, Kian Peymani, Mohsen Soryani.

[5].

[6]. Recognition of Hand Written Text Using Machine Learning Tensor Flow, 2021 International Conference on Computer Communication and Informatics (ICCCI), April 21, Nitin Gupta and Neha Goyal.

[7]. Yu Yanmei, Liu Hui, and Luo Daisheng. Character recognition using a neural network based on English momentum BP. Sichuanb University Journal, 2011(6):1324-05

[8]. Aibin Chen and Lina Lin. Multi-feature-based digital recognition of printed numbers [J]. 2011 (3): 105–108 Computing Technology and Automation

[9]. He Xiaohai, Gao Mingliang, HouShengwei, and TengQizhi. In the [J], automatic detection and recognition. Sichuan University Journal, 2013 (3): 522-528.

[10]. Based on handwritten Chinese character key word extraction using text-dependent handwriting recognition technology, Chen Rui and Tang Yan [J]. Sichuan University Journal, 2013(4): 719–727.

