Optical Character Recognition for Handwritten Text using Region- based CNN

Article · May 2022

CITATIONS
0

READS
46

5 authors, including:

Swati Nitin Shekapure
Marathwada Mitra Mandal's College of Engineering
22 PUBLICATIONS **18** CITATIONS

SEE PROFILE

WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Optical Character Recognition for Handwritten Text using Region-based CNN

[1]Sanskruti Chilvery, [2]Rutuja Kulkarni, [3]Aishwarya Pardeshi, [4]Snehal Solankar, [5]Swati Shekapure

[1,2,3,4] Department of Computer Engineering, Marathwada MitraMandal's College of Engineering, Pune

[5] Faculty at Department of Computer Engineering, Marathwada MitraMandal's College of Engineering, Pune

[3]aishwaryapardeshi2001@gmail.com

ABSTRACT

Optical Character Recognition is a rapidly expanding technology in the field of Machine Learning and Image Processing due to its numerous applications in different fields including Banking, Healthcare, Finance and Legal sector. However, there are very few tools available for Handwritten Text Recognition. It is difficult to store and access physical data with efficiency. Manual labor is required to maintain the proper organization of the data which is a tedious task. Throughout history, there has been a severe loss of data because of the traditional method of storing data. Modern-day technology is letting people store the data over machines, where the storage, organization, and accessing of data is relatively easier. Adopting to the use of Handwritten Text Recognition software, it is easier to store and access data digitally that was traditionally stored physically. Furthermore, it provides more security to the data. In this project we are making an effort to create an application that would be able to recognize handwritten text and give the output in machine readable format which would be analyzable, editable and searchable. For this, we are using Deep learning algorithm Region-based Convolutional Neural Network (RCNN).

Keywords - Deep Learning, Handwritten Text Recognition, Image-preprocessing , Machine Learning, Optical Character Recognition, RCNN

1.Introduction

Handwritten character recognition is becoming increasingly important in today's computerized environment as their practical application in many daily activities appears to be expanding. People can use such technologies to help them do tedious jobs that would otherwise be time-consuming and costly. Handwritten recognition systems can be inspired by biological neural networks, which allow people and animals to learn and simulate non-linear and complex interactions. Neural networks help to imitate how the human brain operates while reading reduced handwriting. It enables machines to read handwriting at a level comparable to, if not superior to, that of humans. Humans use a variety of writing styles, some of which are difficult to understand. Because of its ability to infer meaning from complicated input, a neural network is the best choice for the suggested system. The purpose of this study is to look into the difficulty of locating textual material and transferring it to digital(machine-readable) format. A good example of an OCR system would be the use of an automatic processing system used by banks for depositing cheques. The aim of this project is to develop a system for recognition of handwritten alphabets and words. Deep learning algorithm, Region based CNN is used to achieve this aim.

2.Literature Survey

The algorithm sets two steps: First, preprocessing the handwriting character based on Faster RCNN, second, character recognition based on the Convolutional Neural Networks. The correctness of this method is better than the traditional OCR by the testing data-Handwriting Text Recognition Based on Faster R-CNN paper by J. Yang, P. Ren and X. Kong.[15]

In 'A Survey on various Optical Character Recognition Techniques, March 2018', paper various OCR techniques are given that are used for various character recognition. Sabu, Abin M; Das, Anto Sahaya [12]. The survey includes various stages required to include pre-processing, Classification, Segmented Processing, Feature Extraction.& Recurrent Neural Network.

In"A complete optical character recognition methodology for historic Documents" In this paper is used for identifying historical documents, either printed or handwritten without any knowledge of the font, is presented. & G. Vamvakas's,B. Gatos's. [8] In this paper, a complete OCR method is used for recognising historical documents and paper.

An OCR is not an atomic process but comprises various phases such as acquisition, pre-processing, segmentation, fea-ture extraction, classification, and post-processing. This is stated by Noman Islam, Zeeshan Islam, Nazia Noor in 'A Survey on Opti-cal Character Recognition System, December 2016 [14].

In"Improved Handwritten Digit Recognition using convolutional Neural networks"Amit Choudhary, Anand Nayyar, Saurabh Singh, and Byungun Yoon [15]. In this paper, we have proposed work is to achieve comparable accuracy using a pure CNN architecture through extensive investigation of the learning parameters in CNN architecture for MNIST digit recognition & CNN.

3.System Architecture

1.1.System Architecture

Figure 1 shows the proposed system architecture of the system.

1.2. System Design

Handwritten Text Recognition (HTR) systems consist of handwritten text in the form of scanned images as shown in figure we are going to build a Neural Network which is trained on word images from the IAM 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 (of course, a GPU would be better). For the implementation of HTR, the minimum requirement is TF. Image
of the word taken from IAM Dataset.The proposed recognition technique relies on a Regionl-based convolutional neural network model (RCNN) with a feature-mapped output layer.

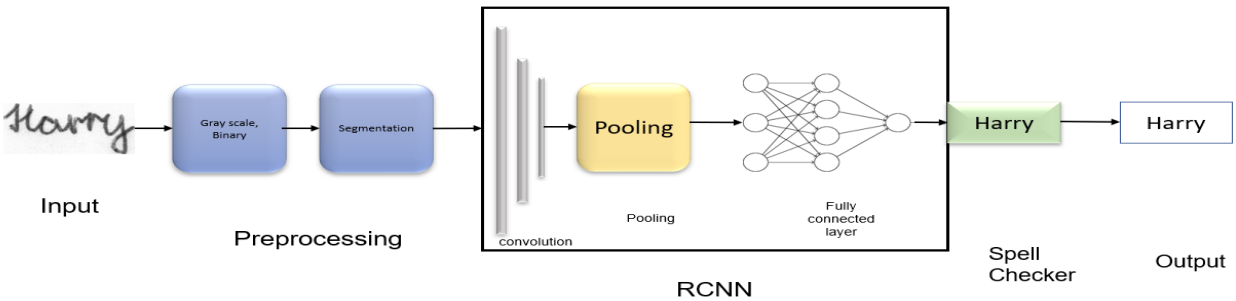
4.Methodology

We have used Region-based Convolutional neural networks (RCNN) for reading and converting input handwritten text into machine readable format. First we take the image as input, then we convert the input image to gray-scale image and then convert the gray-scale image to binary image. This binary image is then passed to the RCNN model and the model processes it by segmenting the input into words and further segmenting the words into alphabets and then giving the output in machine readable form. We are also using a spell-checker, so even if some alphabets are misclassified,
the spell checker would correct it and we will get a more accurate output. The time taken for processing the image is also displayed on screen.

Region-based Convolutional neural network:

RCNN stands for Region-based Convolutional Neural Network. It's a type of neural network that works with data with a grid-like architecture, like an image. Each neuron is linked to other neurons so that the complete network is covered. The Selective search technique is used in the R-CNN architecture to create many regional ideas.

The CNN architecture then uses these region proposals to compute CNN features. These characteristics are then fed into an SVM model, which classifies the object in the region proposal. In an RCNN, each neuron analyses data alone in its own receptive field. The layers are designed so that simpler patterns are detected first, followed by more complicated patterns. A convolutional layer, a pooling layer, and a fully connected layer are the three layers of a standard RCNN.



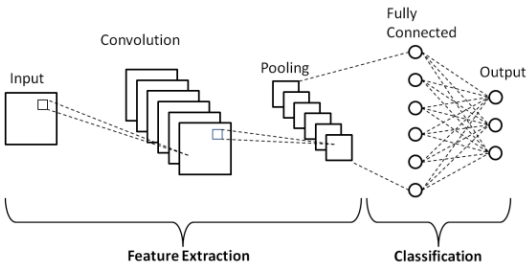


Fig. 2 RCNN architecture

Convolution Layer:

CNN's main building block for extracting features from input images is the convolutional layer. The dot product between two matrices is performed by this layer. As an output, we obtain the Feature Map, which contains information about the image such as its corners and edges. This Feature Map is then given to the next layer, which learns the image's other features.

Max-Pooling Layer:

The goal of this layer is to shrink the convolved feature map in order to reduce computational expenses. This is accomplished by reducing the number of connections between layers. Remove unwanted data, The purpose of max pooling is to zoom out and in to improve model accuracy.

Fully Connected Layer:

The Fully Connected layer contains weights and biases the neurons. It is used to connect the neurons between two different layers. These layers are placed before the output layer and form the last few layers of a RCNN Architecture.

4.Implementation

We have implemented this system using Python language, Django framework, SqLite3 for storing login/registration information, HTML/CSS for front and VS Code IDE. Following are the steps we have followed.

4.1. Data gathering

The IAM dataset used in the study contains 13,353 images of handwritten lines of text authored by 657 writers, totaling 1,539 handwritten pages with 115,320 words classed as part of the contemporary collection.

4.2. Model Training

Our data set contains labeled data. We have used Region based Convolutional Neural Network for training the model. The model is trained on approximately thirteen thousand images. The neural network training is done using five layers. The first two layers contain convolution, max pooling and dropout operation. The last three layers contain dense-layer and dropout layers. The Convolution operation slides over the 2D input data, performing element-wise multiplication, which results in a single output pixel. The same operation is performed for every location it slides over creating a matrix of features. Then max pooling calculates the maximum value for each patch of the feature map.

The Dropout layer randomly sets input units to 0 with a frequency of rate at each step during training time, which helps prevent overfitting. Dense Layer, also known as fully connected layer, is a layer of neurons in which output from a previous neuron is considered as input for current neuron. Activation function used is ReLU. The model involves segmenting the handwritten text into words and then into alphabets.Epochs were set to 500 and after 120 epoch, results were not improving, so training was stopped. The training took around 8-9 hours to complete.

4.3. Model Saving

After the completion of training, the model is saved. This saved model is used for recognizing handwritten text. Spell checker is also used, so that even if some letters are misclassified, we would get correct output with the help of spell-checker.

4.4. Testing

After training and saving the model, it is tested.

Different types of testing are performed on the system such as Unit Testing, Integration Testing, Black box testing, etc.

4.5. Output

The following images show the output of our system.

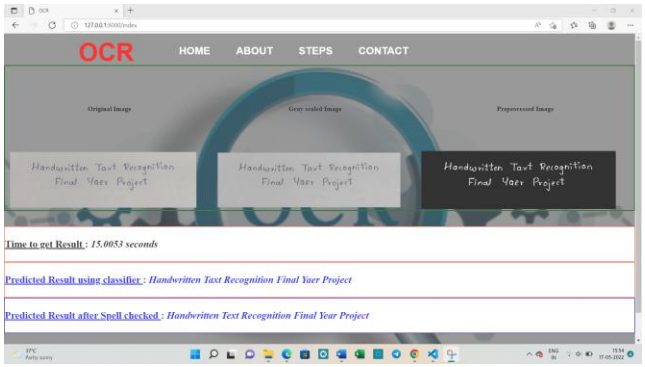


Fig. 3 Output

As you can see in the output image, handwritten text is recognized correctly, processing time for the image is also shown. Spell checker is also applied to correct the spellings.

Conclusion

Handwritten words and sentences are being successfully recognized by the system. By using the RCNN approach for recognizing handwritten text, we have achieved over 93% accuracy.

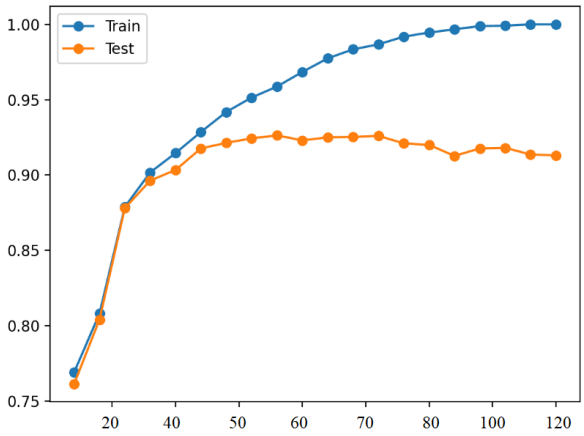


Fig. 4 Accuracy Graph

Acknowledgments

We take this to express our deep sense of gratitude towards our esteemed guide Prof.Swati Shekapure for giving us this splendid opportunity and guiding us throughout the completion of this project and for also providing facilities for successful completion. We thank Dr.K.S.Thakre, Head, Department of Computer Engineering, for opening the doors of the department towards the realization of the seminar, all the staff members, for their indispensable support, priceless suggestion and for most valuable time lent as and when required. With respect and gratitude, we would like to thank all the people, who have helped us directly or indirectly.

References

[1] G. Vamvakas's,B. Gatos's. A Complete Optical Character Recognition Methodology for Historical Documents September 2008 DOI:10.1109/DAS.2008.73 SourceDBLP

[2] Sabu, Abin M; Das, Anto Sahaya. A Survey on various Optical Character Recognition Techniques, March 2018 DOI: 10.1109/ICEDSS.2018.8544323.

[3] Noman Islam, Zeeshan Islam, Nazia Noor. 'A Survey on Opti-cal Character Recognition System, December 2016.

[4] J. Yang, P. Ren and X. Kong, "Handwriting Text Recognition Based on Faster R-CNN," 2019 Chinese Automation Congress (CAC), 2019, pp. 2450-2454, doi: 10.1109/CAC48633.2019.8997382.

[5] Shreya Mhalgi, Ketki Ganu, Prajakta Marne, Radhika Phadke and Swati Shekapure. Handwritten Character Recognition using Neural Networks for Banking Applications. International Journal of Computer Applications. 176(27):1-7, June 2020.