International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 4 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT195386

Optical Character Recognition Using Deep Learning and OpenCV Techniques

Shallu Juneja, Dipti Nayan, Rishabh Bhardwaj
Computer Science Engineering, Maharaja Agrasen Institute of Technology, Rohini, New Delhi, India

ABSTRACT

The problem of image to text-based conversion is persisting in many areas of applications. This project seeks to classify an individual handwritten character so that handwritten text can be translated to a digital form. We used two main approaches to accomplish this task: classifying digits directly and character segmentation. For the former, we use Convolutional Neural Network (CNN) with various architectures to train a model that can accurately classify characters. For the latter, we use Long Short-Term Memory networks (LSTM) with convolution to construct bounding boxes for each character.

Keywords : Character Segmentation, Convolutional Neural Network, Long Short-Term Memory Networks, Classification.

I. INTRODUCTION

Despite the abundance of technical writing tools, many people still choose to take their notes traditionally: with pen and paper. However, there are drawbacks to handwriting text. It's troublesome to store Associate in Nursing access physical documents in an economical manner, search through them with efficiency and to share them with others. Thus, a lot of important knowledge gets lost or does not get reviewed because documents never get transferred to digital format. We have thus decided to tackle this problem in our project because we believe the significantly greater ease of management of digital text compared to written text will help people additional effectively access, search, share and analyze their records, while still allowing them to use their preferred writing method. The aim of this project is to any explore the task of classifying written text and to convert written text into the digital format. Note that even though there need to be some added layers on top of our model to create a fully functional deliverable for an end user, we believe that the most interesting and challenging part of this problem is the classification part, which is why we decided to tackle that instead of segmentation of lines into characters, documents into lines, etc.

We approach this problem with complete images because CNN's tend to work better on raw input pixels rather than features or parts of an image. Given our findings using entire characters images, we sought improvement by extracting characters from each character's image and then classifying each character independently to reconstruct a whole character. In summary, in both of our techniques, our models take in an image of characters and output the name of the characters.

OCR could be a advanced downside as a result of the variability of languages, fonts and styles in which text can be written, and the complex rules of languages etc. Hence, techniques from different disciplines of computer science (i.e. image process, pattern classification and natural language processing etc. are

employed to address different challenges. This paper introduces the reader to the problem. It enlightens the reader with the historical views, applications, challenges and techniques of OCR.

II. LITERATURE REVIEW

Character recognition isn't a brand-new downside however its roots are often derived back to systems before the inventions of computers. The earliest OCR systems weren't computers however mechanical devices that were ready to acknowledge characters, however terribly slow speed and low accuracy. In 1951, M. Sheppard fabricated a reading and automaton GISMO that may be thought-about because the earliest work on fashionable OCR. GISMO will browse musical notations in addition as words on a written page one by one. However, it will solely acknowledge twenty-three characters.

1. Early Scanners

The first thrust behind written text classification was for digit classification for communication mail. Jacob Rainibow early communication readers incorporated scanning instrumentality and hardwired logic to acknowledge mono-spaced fonts. Alum improved this by creating a classy scanner that allowed for a lot of variations in however the text was written in addition as cryptography the knowledge onto a barcode that was written directly on the letter. Several researches are planned over the years for character recognition. A noble plan was conceptualized method back in 1999 by (Sural and Das). the flexibility of a software or application like this one needs the ability to apply pattern recognition, pattern interpretation and learning. This wanted the employment of multi-layer perception model that may be a feed forward model. They used the model to develop associate OCR(Optical Character Recognition) for associate Indian Language named Bengali. The model that was used belong to a category of artificial neural network. The advantage of exploitation perception model is its potency of learning. It are often wont to train and acknowledge any reasonably knowledge set than at the start outlined. Moreover, their approach left an excellent scope of developing OCR that concentrate on any language aside from Bengali. Another work comes from Deepayan Sarkar from University of Wisconsin. He enforced OCR as AN add on package for a MATLAB-like programming atmosphere known as R. Although results weren't that smart, they were not bad either suggesting this technique is not flawed.

2. To the digital age

The first prominent piece of OCR software was invented by Ray Kurzweil in 1974 as the software allowed for recognition for any font. This software system used a a lot of developed use of the matrix technique (pattern matching). Essentially, this could compare bitmaps of the guide character with the bitmaps of the scan character and would compare them to see that character it most closely matched with. The downside was this software was sensitive to variations in sizing and the distinctions between everyone's way of writing.

3. Optical Character Recognition

Algorithms designed to acknowledge written characters are still less advanced than that for written characters, due chiefly to the hardships in handling the variety in written characters' shapes and forms. Characters segmentation to modify the popularity method is another downside.

The attention-grabbing truth concerning Deep Neural Networks is that by creating an oversized network with several layers it becomes a lot of capable to find a lot of options mechanically. In the work by Jaderberg, the authors used CNN not only to detect characters regions in the image but also to recognize the characters within these regions. They reported good performance on a few benchmark datasets such as 91% accuracy in the character

classification of ICDAR 2003. To solve the matter of latency in process the information, GPUs are used as suggested by Ciresan et al. The one who trained and tested the CNN network using committee of classifiers and reduced the error rate of MNIST dataset to 2.7%.

III. MAJOR PHASES IN OCR

a) *Numeral Size*

As can be seen in the Data section of our paper, we had many unique digits in our dataset. However, some characters images appeared in our dataset only a few times, which made it very hard for us to train on these images. This issue, along with the fact that our dataset already had a large vocabulary, encouraged us to cut off some of the characters in our dataset and not include those characters in our training/validation/test dataset that we were going to use with our models. We, therefore, limited our data with characters images that appeared at least 20 times in our dataset (before splitting into train and validation sets). Our models and algorithms are not dependent on a hardcoded number of images and would thus have worked with any number of examples, but we decided to narrow the number of characters down for efficiency purposes.

b) Classification

We trained our characters classifier with multiple CNN architectures: VGG-19, RESNET-18, and RESNET-34 By moving from the traditional 3-7 layers of previous CNN to 16-19 layers for different iterations of their model, their model not only obtained first and the second places in the localization and classification tracks, respectively, but also was found to generalize well to other computer vision tasks.

It is defined as the process of classifying a character into its appropriate category. The structural approach to classification is based on relationships present in image components. The statistical approaches are based on use of a discriminate function to classify the image.

Fig1: Digit recognition using deep learning

c) Character-Level Classification

The character-level classification model was quite like the characters-level classification model. The main differences included passing in character images instead of characters images, utilizing a character-level vocabulary instead of a character-level vocabulary, and training a different parametrization of each of the variants of our very deep learning models. The architectures of these models were otherwise the same as our characters-level models.

d) Skew Correction

Camera captured pictures fairly often suffer from skew and perspective distortion as mentioned . These occur because of uneven axes and/or planes at the time of capturing the image. Therefore, the image can't be de-skewed at one pass. These values are the distances in terms of constituent from a aspect to the primary gray/black constituent of the text region. Among these four profiles, the one that is from rock bottom aspect of the text region is taken into thought for estimating skew angle.

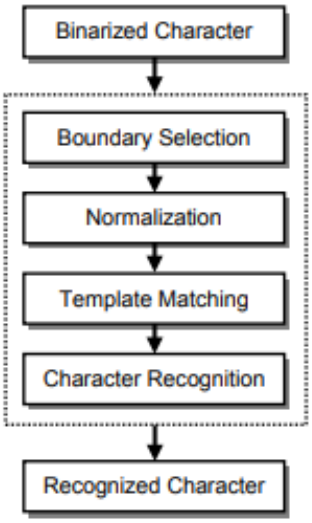

Fig 2:Phases of OCR

IV. CONCLUSION

A complete OCR system has been given during this paper. Because of the computing constraints of hand-held devices, we've got unbroken our study restricted to light-weight and computationally economical techniques. Compared to Tesseract, nonheritable recognition accuracy (92.74%) is nice enough. Experiments shows that the popularity system given during this paper is computationally economical that makes it applicable for low computing architectures like mobile phones, personal digital assistants (PDA) etc.

V. ACKNOWLEDGMENT

The authors would like to thank Maharaja Agrasen Institute of Technology for their support in the completion of this research work.

VI. REFERENCES

[1]. Alex Graves, Santiago Fernandez, Faustino Gomez, Juergen Schmidhuber, Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks, Proceedings of the 23rd international conference on Machine learning. 2006

[2]. Convolutional Neural Network Benchmarks: https://github.com/jcjohnson/cnn-benchmarks

[3]. Elie Krevat, Elliot Cuzzillo. Improving Off-line Handwritten Character Recognition with Hidden Markov Models.

[4]. Fabian Tschopp. Efficient Convolutional Neural Networks for Pixelwise Classification on Heterogeneous Hardware Systems George Nagy. Document processing applications.

[5]. Mail encoding and processing system patent: https://www.google.com/patents/US5420403

[6]. Kurzweil Computer Products. http://www.kurzweiltech.com/kcp.html

[7]. H. Bunke1, M. Roth1, E.G. Schukat-Talamazzini. Offline Cursive Handwriting Recognition using Hidden Markov Models.

[8]. K. Simonyan, A. Zisserman Very Deep Convolutional Networks for Large-Scale Image Recognition arXiv technical report, 2014

[9]. Lisa Yan. Recognizing Handwritten Characters. CS 231N Final Research Paper.

[10]. Oivind Due Trier, Anil K. Jain, Torfinn Text. Feature Extraction Methods for Character Recognition–A Survey. Pattern Recognition. 1996

[11]. Satti, D.A., 2013, Offline Urdu Nastaliq OCR for Printed Text using Analytical Approach. MS thesis report Quaid-i-Azam University: Islamabad, Pakistan. p. 141.

[12]. Mahmoud, S.A., & Al-Badr, B., 1995, Survey and bibliography of Arabic optical text recognition. Signal processing, 41(1), 49-77.

[13]. Bhansali, M., & Kumar, P, 2013, An Alternative Method for Facilitating Cheque Clearance Using Smart Phones Application. International Journal of Application or Innovation in Engineering & Management (IJAIEM), 2(1), 211-217.

[14]. Qadri, M.T., & Asif, M, 2009, Automatic Number Plate Recognition System for Vehicle

Identification Using Optical Character Recognition presented at International Conference on Education Technology and Computer, Singapore, 2009. Singapore: IEEE.

[15]. Shen, H., & Coughlan, J.M, 2012, Towards A Real Time System for Finding and Reading Signs for Visually Impaired Users. Computers Helping People with Special Needs. Linz, Austria: Springer International Publishing.

[16]. Bhavani, S., & Thanushkodi, K, 2010, A Survey On Coding Algorithms In Medical Image Compression. International Journal on Computer Science and Engineering, 2(5), 1429-1434.

[17]. Bhammar, M.B., & Mehta, K.A, 2012, Survey of various image compression techniques. International Journal on Darshan Institute of Engineering Research & Emerging Technologies, 1(1), 85-90.

Cite this article as :

Shallu Juneja, Dipti Nayan, Rishabh Bhardwaj , "Optical Character Recognition Using Deep Learning and OpenCV Techniques", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 4, pp. 05-09, July-August 2019. Available at doi : https://doi.org/10.32628/CSEIT195386
Journal URL : http://ijsrcseit.com/CSEIT195386