IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 09, Issue 5 (May. 2019), ||S (IX) || PP 33-37

Implementation of Handwritten Character Recognition Using Machine Learning

Neha Chauhan[1], Prof. Sanjay Balwani[2], Prof. Mayuri Chawla[3]
[1,2,3] Department of Electronics & Telecommunication
[1,2,3]Jhulelal Institute of Technology, RTMN University
Nagpur, Maharashtra,

Abstract: *In this paper, a new analytic scheme, which uses a sequence of segmentation and recognition algorithms, is proposed for handwriting recognition problem. We propose a new neural network architecture for state-of-the-art handwriting recognition, alternative to multi-dimensional long short-term memory (MD-LSTM) recurrent neural networks. The model is based on a convolutional encoder of the input images, and a bidirectional LSTM decoder predicting character sequences.*

Keywords: *Machine learning, Handwriting recognition, neural networks, MD-LSTM, Character sequences, segmentation and recognition algorithms*

I. Introduction

Handwriting recognition has been one of the most fascinating and challenging research areas in field of image processing and pattern recognition in the recent years. It contributes immensely to the advancement of automation process and improves the interface between man and machine in numerous applications. In general, handwriting recognition is classified into two types as off-line and on-line handwriting recognition methods. The on-line methods have been shown to be superior to their off-line counter parts in recognizing handwritten characters due to the temporal information available with the former. However, in the offline systems, comparably high recognition accuracy level isobtained.

In on-line recognition multi-dimensional long short-term memory recurrent neural networks (MDLSTM-RNNs) became established as the state-of-the-art model for handwriting recognition with the help of neural networks. They were involved in all winning systems in international evaluations and are now a standard component of industrial systems. Yet, they have conceptual drawbacks. Most notably, the features computed by an MDLSTM layer at a given position may depend on quite “distant” input features. In other words, there is no explicit control of the input context. We observed that this lack of control of context dependency makes the features learnt on text lines images not generalizable to paragraph images. In the scope of handwriting recognition, we think that the context of the whole image is not relevant to predict a single character. Instead, an area covering the character, and maybe surrounding ones, should be sufficient.

II. Related Work

A. *Renuka Kajale,* [1] In this paper, we propose a method for intelligent character recognition using classifiers and transferring the data to excel sheet. The results show such methods can produce accuracies close to 95% for alpha numerals and special characters.

B. *Subodh L. Wasankar,* [2] The prime objective of this Research is the development of effective reading skills in Machines. After reading the text and comprehending the meaning, it would self-program itself and according to the program it would implement the instructions. The current investigation presents an algorithm and software which detects, recognizes text and character with specific protocol in a live streaming video and programs itself according to the text. This technology can prove to be immensely important to various human activities in day to day life.

C. *Mithili N Mayekar,* [3] In today's world, with the advancement in the technology, everything is become digitized. This has led to converting all the important documents in digital format. It is at this point where OCR (Optical Character Recognition) plays a very important role. Lots of researchers have been working tirelessly to improve the accuracy of the recognition rate. Taking into consideration the nature of this application, compute speed is one of the most important factor that should be improved. So this increase in speed can be achieved by making use of the computer parallel processing power. The GPU (Graphic Processing Unit) is known for its parallel computing capability which helps in increasing the computation speed.

III. Character Dataset

For applying machine learning algorithms, we need a dataset. To generate a dataset, we need to go through various stages (Fig. 1). As mentioned.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

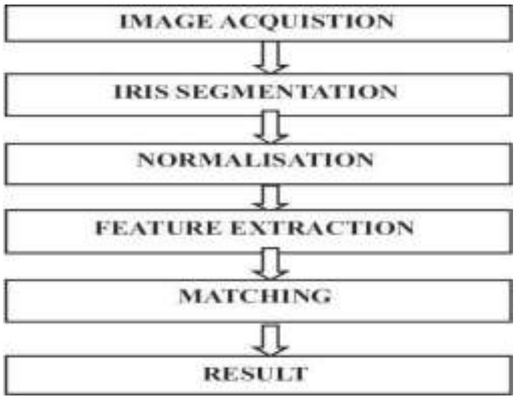

Fig 1. Overall procedure of implementation

1. Image acquisition

The first stage here is to obtain a clear image to an extent such that its zoomed view should not appear blurred. Image enhancement should not be needed, as an unclear image may not aid our task.

2. Pre-processing

If image is not enhanced in the previous step, we need to do it here. We apply spatial image filtering, global image thresholding processes, etc. Once this step is over, noise removal techniques are applied here (Fig. 2). The following algorithmic steps are used in pre-processing:

i. Noise removal : Before sending the image for the further stages, we need to pre-process the image. Then we will be able to read the images clearly or else there will be more/less gaps between edges, characters, etc. Some parts of these techniques come under morphological operations.

ii. Colorization and background detection :

After noise removal, the image is sent further to detect various colors in and around the nameplate. Moreover, the background surrounding the characters is also detected and finally we have separated the characters from its surroundings.

iii. Edge detection :

In image processing, all the unnecessary data is removed and just the edges are kept. This helps in processing the given data.

iv. Plate segmentation :

This is an additional step and may not be used always. If the image is not processed in the background detection step, it is done here. As soon as characters are separated from nameplate, the job of this stage is over.

v. Character segmentation :

This is an important step because the characters are of different scripts. Some nameplates contain English, some Hindi and so on. In-depth processing is done here, as there are separate character segmentation of different languages.

vi. Erosion and dilation :

We need to increase or decrease the objects in size and that is done in this stage. The rule says that an image is made smaller by eroding the pixels on its edges while that same image is made larger by adding pixels around its edges.

vii. Skew detection/correction :
The image at this stage is skewed and rotated at an angle such that the image is viewed from a desired point. In some cases, the image appears to be laterally inverted and this stage operation is useful in such cases.

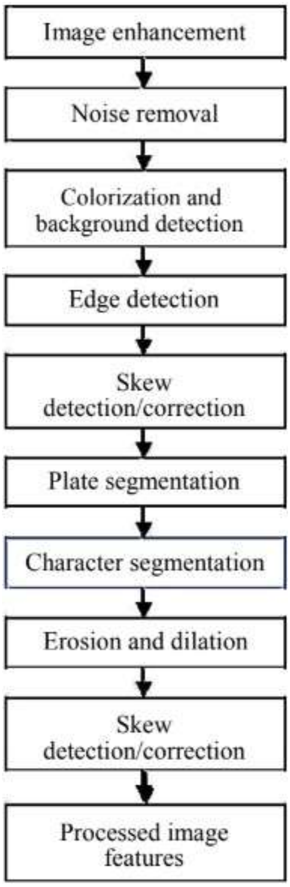

Fig 2. Preprocessing of the image(Flow chart)

IV. Visualization Of The Dataset And Classification

Now the process of applying classification algorithms comes into picture. The characters are identified using Gated Convolutional Recurrent Neural Networks.

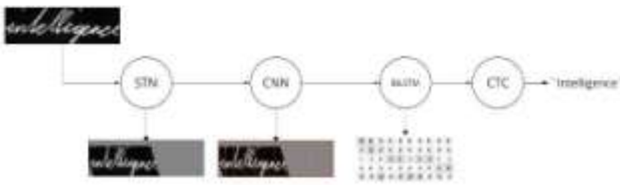

Fig 3 : Processing of the image back end process

In this section, we present the convolutional gates that we use in the encoder. The idea is depicted on Fig. 3. The gate controls the propagation of a feature to the next layer.

Basically, the gate looks at the feature value at a given position, and at neighboring values, and decides whether that feature at that position should be kept or discarded. It allows to compute generic features across the whole image, and to filter when, according to the context, the features are relevant.

V. Architecture Details

We built neural networks according to the principles previously described. To optimize different objectives (accuracy, speed and size), we designed two architectures. Both are made of a convolutional encoder, a max-pooling across the vertical dimension and a recurrent decoder. The encoder consists of a 3x3 convolutional layer with 8 (resp. 4) features, a 2x4 convolutional layer with 16 (resp. 8) features, a 3x3 convolutional gate, a 3x3 convolutional layer with 32 (resp. 16) features, a 3x3 convolutional gate, a 2x4 convolutional layer with 64 (resp. 24) features, a 3x3 convolutional layer with 128 (resp. 32) features, for the big (resp. small) network. The encoder of the small network has an additional 3x3 convolutional layer with 64 features. The decoders are made of 2 bidirectional LSTM layers of 128 (resp. 50 and 100) units, with a linear layer of 128 (resp. 100) neurons in between.

VI. Training

In this section, we present the results of the networks trained on all the available data, in all languages. We will refer to these networks as Generic models, since they are not specific to a language. We compare the four networks (Accurate, Fast, Fast Small and Faster Small) to the previous version of A2iA Text Reader.

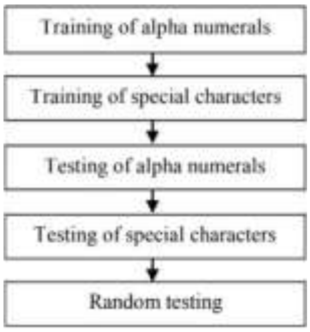

Fig 3. : Machine learning procedure

We trained the network to minimize the Connectionist Temporal Classification objective function. We performed the optimization with stochastic gradient descent, using the RMS Prop method with a base learning rate of 0.0004 and mini batches of 8 examples. Following the motivations previously exposed, we first train each of the four models on all the available data, to obtain generic Latin-script neural networks. The decoder part is subsequently adapted to each language, using the features extracted with the encoder. The encoder is not fine-tuned to maintain generic, language and collection-independent features. For comparison purposes, we also adapted to whole architecture to each language, and trained neural networks without gates.

VII. Future Scope

Going by the issues addressed in this paper, we conclude That there is a huge scope to extend this work forward. We can go further in recognizing characters in a more challenging scenario.

Machine learning has been applied to a number of applications. Some of the literatures covering these are languages other than English, namely, Latin, Cyrillic, Arabic, Hebrew, Indic, Bengali (Bangla), Devanagari, Tamil, Chinese, Japanese, Korean, etc. Highlight in 1950's, applied throughout the spectrum of industries resulting into revolutionizing the document management process. Optical Character Recognition or OCR has enabled scanned documents to become more than just image files, turning into fully searchable documents with text content recognized by computers.

VIII. Result

A. Multilingual Training

In this section, we present the results of the networks trained on all the available data, in all languages. We will refer to these networks as Generic models, since they are not specific to a language. We compare the four networks (Accurate, Fast, Fast Small and Faster Small) to the previous version of A2iA Text ReaderTM. A2iA Text ReaderTM comes with two neural networks for each language (an accurate setup and a fast one). Each network is an MDLSTM- RNN following the architecture, or variations thereof (e.g. different numbers of neurons in hidden layers, or subsampling convolutions before the first MDLSTM to get faster decoding). In most cases, the accurate version operates on original (300DPI) images, and the fast version on downscaled input images to about 70% of the original size. Each one of these networks is only trained on language specific data, although for low-resourced languages the initial weights come from the network trained on another language. The comparison is carried out on a large benchmark including the validation and test sets of all public datasets, mobile captured pictures of some images from IAM and Maurdor, some private data (not necessarily coming from the same collections as the training sets). We compare the results of the previous version of A2iA Text ReaderTM to the generic gated convolutional recurrent neural network (GCRNN). We see that for all languages except French and Russian, the generic GCRNN is better, and even the smaller and faster architectures are competitive. Note that the comparison here is between seven language-specific networks of the previous version of A2iA Text ReaderTM and a single GCRNN network. The Portuguese networks in the previous version of A2iA Text ReaderTM are actually the networks trained on French data, due to the lack of training documents in Portuguese when the systems were created. They are intended to be applied with a Portuguese language model. That explains why the character error rates of the network alone on Portuguese data is so high.

IX. Conclusion

The paper addresses and implements a useful method of character recognition. Handwritten recognition was the main aim, as printed text recognition did not require many datasets for the machine to be trained. A dataset of around 200 samples was applied for those 36 alpha numerals and further 100 for special characters. Gated Convolutional Recurrent Neural Network & machine learning proves to be useful as classification labels can be obtained from the features. Ultimately, the data is stored in excel sheet for further analysis.

References

[1]. Renuka Kajale1, "Supervised machine learning in intelligent character recognition of handwritten and printed nameplate"2017 International Conference on Advances in Computing, Communication and Control (ICAC3), 2017*)*

[2]. Subodh L. Wasankar, "Machine Learning with Text Recognition" 2010 IEEE International Conference on Computational Intelligence and

[3]. Computing Research,2010.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4]. Mithili N Mayekar, "Implementation Of Machine Learning Algorithm For Character Recognition On GPU"Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication (ICCMC), 2017.

[5]. -Dippal Israni, "Vehicle Classification and Surveillance using Machine Learning Technique"2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)..

[6]. Moriwake Ryo, "Character Recognition in Road Signs Using a Smartphone" 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI).M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7]. Introduction to machine learning by Nils .J. Nilssons

[8]. Machine learning for absolute beginner by Oliver theobald

[9]. https://www.researchgate.net/publication/320442536_A_Survey_on_ Optical_Character_Recognition_System

[10]. https://ieeexplore.ieee.org/Xplore/home.jsp

[11]. https://www.researchgate.net/publication/320442536_A_Survey_on_ Optical_Character_Recognition_System