International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 07 | July 2021 www.irjet.net p-ISSN: 2395-0072

HAND WRITTEN CHARACTER RECOGNITION SYSTEM

Kunal Ingle[1], Vedika Sonulkar[2], Shrawani Rangari[3], Riddhehwar Akare[4], Amruta Ingle[5], Prof. Manjusha Talmale[6]

[1,2,3,4,5]Final Year Student, Department of Computer Science And Engineering, GNIT Nagpur, Maharashtra, India
[6]M.Tech in Computer Engineering, Assistant Prof. GNIT, Nagpur, Maharashtra, India

---***---

Abstract - *Automatic handwritten character recognition is a subject that receives much attention at present. One of the main drivers behind current research is the capacity to quickly interpret tiny handwriting samples such as those seen in checks and envelopes. Many (offline) hand-written Text Recognition (HTR) systems research have been conducted to create state-of-the-art models for small enterprise line recognition. But it presents significant problems to add HTR capabilities to a multiple OCR system. This article deals with three issues related to systems such as data, efficiency and integration.*

The Project is a computer based programme that minimises effort in converting the handwritten script photographs into text documents. The problems are addressed through the use of online handwriting data for the online recognition system for a large-scale manufacturing. We present our pipeline of picture data creation and investigate how HTR models may be built using online data. We show that the data considerably enhance models in the circumstance that just a few number of actual pictures, generally the case with HTR models, are available. It allows us to considerably decrease the costs of a new script. Secondly, we present a model for line recognition based on non-recurring connectivity neural networks. We are exploring this approaches in order to develop an excellent English written word recognition system based on the recognition of character. Lexicon post-processing is used to increase the overall accuracy of recognition. There are several approaches accessible for the extraction and training, each with its own superiorities and limitations, of CR systems in literature. With the LSTM models, the model achieves equivalent precision while enhancing parallelism in training and inferences. Finally, we are offering an easy approach of integrating HTR models into OCR. This is a solution for bringing HTR into a wide-ranging OCR.

Key Words: Character, Picture, Text, Word, Image, visualization.

1.INTRODUCTION

The processing of pictures might be a computer vision modification of images. There are various ways for manipulating pictures with the use of technology. In many fields, text recognition has an enormous significance. But such a task by a machine is hard to try. We must guide the system in order to recognise the text. The acquisition, extraction, categorization and recognition of character comprises multiple phases. The capacity to receive and analyse the handwritten information from an external source is handwriting recognition. The main purpose of this research is to design a system that can acknowledge the real format character of a neural network effectively.

Neural computing may be a fairly new area, and consequently style components are less specific than other architectures. Neural computers use parallel data. A neural computer works considerably differently from the functioning of a normal computer. Neural computers are taught such that the provided data is compared to the system learned and the user receives appropriate output text. A handwriting recognition system manages formatting, segments correctly and identifies the most likely words. The automatic translation of the text of an image into letter codes, used within your computer and text-processing programmes, requires offline handwriting identification. The information collected in this way is seen as a static handwriting image.

1.1 Objective

Most companies utilise papers to obtain consumer information. These papers are usually written by hand. These may include paperwork, inspections and so on. Documents are converted and saved in digital versions for quicker retrieval or data gathering. Common procedure to manage this data is that the same data is input into the computer manually. The handling of such documentation is tireless and time-consuming. Therefore the demand for a specific handwritten recognition software is made, which recognises texts automatically from images of documents. Hand-written Character Recognition (HCR) Software makes it easier to extract and save data from hand-written documents in *electronic forms. Banking, healthcare and many similar organisations, which routinely employ handwritten papers. The newly developing fields where manual writing data input is necessary, such as electronic library development, multimedia database etc, are also being identified in HCR systems.*

1.2 Literature servery

In his work, Anuj Dutt[2] showed that he was able to achieve extremely great precision utilising deep learning techniques. By utilising Keras and Theano as backend, the convolutional Neural Network, he achieved 98.72 percent accuracy. In addition, CNN implementation using Tensorflow results even better by 99.70%. While it looks more difficult than typical machine learning algorithms, the precision of processes and codes has been more apparent.

A technique of classifying age, sex and nationality by writing was proposed by S. Al-ma'adeed and A. Hassaine[3] in another article. Use of Random Forest Classifier and Kernel Discriminant analysis for classification using spectral regression. For the situations of age, sex and nationality, they created the random forest classifications using the Random Forest Library. You explain your experiments in the QUNI manual writing database. A random ranking would forecast about 50 percent, because it is a gender two-classification.Thus, for the age classification a random classification would predict about 14 per cent. A random categorization would thus forecast for nationality predictions just around 12%.

Abdul Hamid and Amin Sjarif[4] in their suggested system have employed three classification methods, which are the Support Vector Machine (SVM), the Nearest Neighbor (KNN) and the Multilayer Perceptron Network (MLP). SVM and KNN predict the data set accurately but MLP Neural Network can't predict number 9 among these methods. The reasons behind this case have also been shown. KNN and SVM directly anticipate the extraction of features, but the MLP function is nonlinear. It is thus more suited for non-linear model learning. In addition, MLP with hidden layers function non-convex losses if more than one local minimum is present.. Various random weight initializations can give rise to various accuracy validations. But Keras can help enhance it by employing Convolutional Neural Networks.

2. Proposed method

A. Hand written text recognition:

HTR systems comprise of handwritten text in scanned pictures as seen in figure 1. Handwritten text recognition systems We'll create a neural network (NN) trained using IAM dataset word-images. Since the input layer is generally kept short for word-images (and thus all the opposing layers), CPU NN-Training is achievable (of course, a GPU would be better). The minimal prerequisite for implementing HTR is TF.

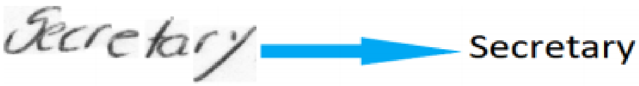

Fig. 1: Image of word taken from IAM Dataset

B. Model Overview:

We use a NN for our task. It consists of a convolutional neural network (CNN) layers, recurrent neural network (RNN) layers, and a final Connectionist Temporal Classification (CTC) layer.

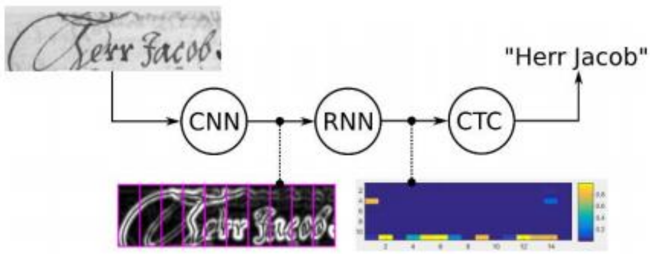

Fig. 2: Overview of HTR

We have taken five CNN (functional extraction), two RNN layers and one CTC layer in this project (calculate the loss). We must first prepare the photos so thatwe can decrease noise.

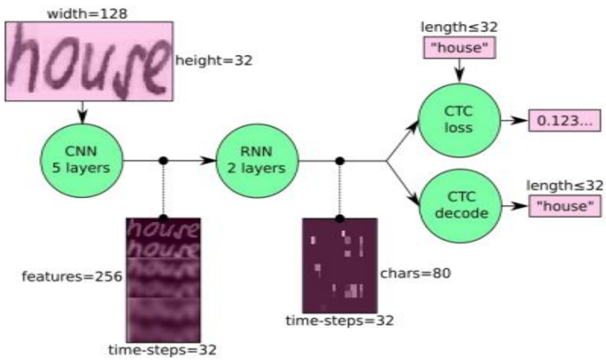

Fig. 3: Green indicates the operations of NN and Pink indicate the dataflow through NN.

We can also interpret the NN in very formal fashion (see Eq. 1), which maps an image (or matrix) M size W to a sequence of personality (c1, c2, ...) between 0 and L. (see Eq. 1). As you see, the text is identified at the level of the character, thus words or texts not in the data are recognised as such (as long because the individual characters get correctly classified).

3. CONCLUSION

The characters are classified in this project. Through the typical neural network the project is realised. The exactness we have achieved is over 90,3%. The efficiency and effective results for recognition will be achieved using this method. The project is most precise for the least noise-dependent material. The exactness depends entirely on the data set We can gain better precision if we expand the data. If the cursive writing is avoided, it is likewise the best outcome.

REFERENCES

[1] Amin, Adnan, et al. "Recognition of hand-printed Latin characters based on a structural approach with a neural network classifier." Journal of Electronic Imaging, Vol. 6(3), July 1997. pp. 303-310.

[2] Bharath, Ramachandran. Neural Network Computing.McGraw-Hill, Inc., New York, 1994. pp. 4-43.

[3] Blumenstein, M. and Verma, B. "An Artificial Neural Network Based Segmentation Algorithm for Off-line Handwriting Recognition." International Conference on Computational Intelligence and Multimedia Applications, 1998. pp. 306-311.

[4] Blumenstein, M. and Verma, B. "Neural-based Solutions for the Segmentation and Recognition of Difficult Handwritten Words from a Benchmark Database." Proceedings of the Fifth International Conference on Document Analysis and Recognition, September 1999. pp. 281-284.

[5] Cho, Sung-Bae. "Pattern recognition with neural networks combined by genetic algorithm."Fuzzy Sets and Systems, Vol. 103(2), April 1997. pp. 339-347.

[6] Gader, Paul D., et al. "Neural and Fuzzy Methods in Handwriting Recognition." Computer, February 1997. pp. 79-86.

[7] Liu, Cheng-Lin, and Nakagawa, Masaki. "Handwritten Numeral Recognition Using Neural Networks: Improving the Accuracy by Discriminative Training." Proceedings of the Fifth International Conference on Document Analysis and Recognition, September 1999. pp. 257-260.

[8] Luger, George F., and Stubblefield, William A. Artificial Intelligence: Structures and Strategies for Complex Problem Solving, (2nd Edition). Benjamin/Cummings Publishing Company, Inc., California, 1993, pp. 516-527.

[9] Molen, A.E., and Veelenturf, L.P.J. "On-line signature validation using Kohonen's Neural Network." Neural Networks: Best Practice in Europe. World Scientific Publishing Co. Pte. Ltd., Singapore. pp. 145-148.

[10] Russell, Stuart J., and Norvig, Peter. Artificial Intelligence: A Modern Approach. Prentice-Hall, Inc, New Jersey, 1995. pp. 567- 587.

[11] Skapura, David M., Building Neural Networks. ACM Press, New York. pp. 29-33.

[12] Xiao, Xu-Hong, and Leedham, Gary. "Signature Verification by Neural Networks with Selective Attention."Applied Intelligence, Vol. 11(2), Sept/Oct. 1999. pp. 213-223.

[13] Stephane Armand, Michael Blumenstein and Vallipuram Muthuk kumarasamy, "Off-line Signature Verification based on the Modified Direction Feature", 2006 International Joint Conference on Neural Networks, pages 684-691.