IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A Survey on Deep Learning For Hand Written Recognition by Digit & Character

[1]Sana Pavan Kumar Reddy,[2]V Manikyala Rao

[1]Assistant Professor,[2]Assistant Professor,
[1]Computer Science and Engineering,
[1]Malla Reddy Insistitute of Technology ,Hyderabad,India

Abstract: In Deep Learning Hand Writing Recognition got lot of attention. Optical Character Recognition (OCR) and Handwritten Character Recognition (HCR) has specific domain to apply. Disparate techniques have been proposed to for character recognition in handwriting recognition system. Despite the fact, sufficient studies and papers describes the techniques for converting textual content from a paper document into machine readable form. In coming days, character recognition system might serve as a key factor to create a paperless environment by digitizing and processing existing paper documents. To classify an individual handwritten word so that handwritten text can be translated to a digital form. We used two main approaches to accomplish this task: classifying words directly and character segmentation. First, we use Convolutional Neural Network (CNN) with various architectures to train a model that can accurately classify words. Next, we use Long Short Term Memory networks (LSTM) with convolution to construct bounding boxes for each character. We then pass the segmented characters to a CNN for classification, and then reconstruct each word according to the results of classification and segmentation. This paper presents a detailed review in the field of Handwritten Character Recognition.

Index Terms: Deep *Learning, CNN, LSTM, Handwritten Character Recognition, Optical Character Recognition.*

I. Introduction

Handwritten Text Recognition is a technology that is much needed in this world as of today. Before proper implementation of this technology we have relied on writing texts with our own hands which can result in errors. It's difficult to store and access physical data with efficiency. Manual labor is required in order to maintain proper organization of the data. Throughout history, there has been 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 the use of Handwritten Text Recognition software, it's easier to store and access data that was traditionally stored. Furthermore, it provides more security to the data. One such example of Handwritten text Recognition software is the Google Lens. The aim of our project is to make an application for mobile devices that can recognize the handwriting using concepts of deep learning. We approached our problem using CNN as they provide better accuracy over such tasks.

Deep Learning

Deep learning is the acronym for Neural Networks, the network connected with multilayers. It is a set of learning methods attempting to model data with complex architectures combining different non-linear transformations. The elementary bricks of deep learning are the neural networks, that are combined to form the deep neural networks. These techniques have enabled significant progress in the fields of sound and image processing, including facial recognition, speech recognition, computer vision, automated language processing, text classification (for example spam recognition). The layers are composited form nodes. A node is just a perception which takes an input performs some computation and then passed through a node's activation function, to show that up to what context signal progress proceeds through the network to perform classification.

CNN

Convolution Neural Network is one of the most important forms of deep learning. The Convolutional Neural Networks (CNN), particularly adapted for image processing. The Convolutional Neural Network (CNN) has shown excellent performance in many computer vision, machine learning, and pattern recognition problems. Many solid papers have been published on this topic, and quite a number of high quality open source CNN software packages have been made available..CNNs are useful in a lot of applications, especially in image related tasks. Applications of CNNs include image classification, image semantic segmentation, object detection in images, etc. CNNs are comprised of three types of layers. These are convolutional layers, pooling layers and fully-connected layers.

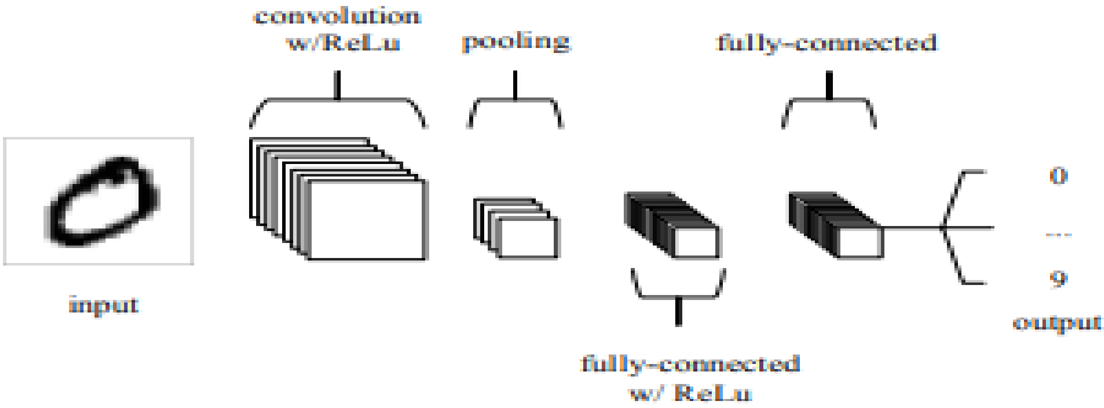

Fig. 1: An simple CNN architecture, comprised of just five layers

The basic functionality of the example CNN above can be broken down into four key areas.

1. As found in other forms of ANN, the input layer will hold the pixel values of the image.
2. The convolutional layer will determine the output of neurons of which are connected to local regions of the input through the calculation of the scalar product between their weights and the region connected to the input volume. The rectified linear unit (commonly shortened to ReLu) aims to apply an elementwise' activation function such as sigmoid to the output of the activation produced by the previous layer.
3. The pooling layer will then simply perform downsampling along the spatial dimensionality of the given input, further reducing the number of parameters within that activation.
4. The fully-connected layers will then perform the same duties found in standard ANNs and attempt to produce class scores from the activations, to be used for classification. It is also suggested that ReLu may be used between these layers, as to improve performance.

Character Recognition

Character recognition is a fundamental, but most challenging in the field of pattern recognition with large number of useful applications. It has been an intense field of research since the early days of computer science due to it being a natural way of interactions between computers and humans. More precisely Character recognition is the process of detecting and recognizing characters from the input image and converts it into ASCII or other equivalent machine editable form. The technique by which a computer system can recognize characters and other symbols written by hand in natural handwriting is called handwriting recognition system. Handwriting recognition is classified into offline handwriting recognition and online handwriting recognition [3]. If handwriting is scanned and then understood by the computer, it is called offline handwriting recognition. In case, the handwriting is recognized while writing through touch pad using stylus pen, it is called online handwriting recognition. From the classifier perspective, character recognition systems are classified into two main categories i.e. segmentation free (global) and segmentation based (analytic). The segmentation free also known as the holistic approach to recognize the character without segmenting it into subunits or characters. Each word is represented as a set of global features, e.g. ascender, loops, cusp, etc. Whereas segmentation based approach [4]; each word/ligature is segmented into subunits either uniform or non-uniform and subunits are considered independently. Handwritten character processing systems are domain and application specific, like it is not possible to design a generic system which can process all kinds of handwritten scripts and language. Lots of work has been done on European languages and Arabic (Urdu) language. Whereas domestic languages like Hindi, Punjabi, Bangla, Tamil, Gujarati etc. are very less explored due to limited usage.

Digit Recognition

Handwritten digit recognition has gained so much popularity from the aspiring beginner of machine learning and deep learning to an expert who has been practicing for years. Developing such a system includes a machine to understand and classify the images of handwritten digits as 10 digits (0–9). Handwritten digits from the MNIST database are already famous among the community for many recent decades now, as decreasing the error rate with different classifiers and parameters along with preprocessing techniques from 12% error rate with linear classifier (1 layer NN) to achieving 0.23% error rate with hierarchy of 35 convolution neural networks [Yann LeCun, MNIST database of handwritten digits]. The scope of this article is to compare the different classifiers with different parameters and try to achieve near-human performance.

Digit recognition system is the working of a machine to train itself or recognizing the digits from different sources like emails, bank cheque, papers, images, etc. and in different real-world scenarios for online handwriting recognition on computer tablets or system, recognize number plates of vehicles, processing bank cheque amounts, numeric entries in forms filled up by hand (say — tax forms) and so on.

OCR

Optical Character Recognition (OCR) also known as Optical Character Reader, which is conferring the mechanized recognition of photographic document for many languages. Practically the OCR reader tool transfigures the script, which is written by hand or typewritten, printed script format into machine copy-edit text or computer processable format (ASCII), whether from a resemble scan document or a camera image. The OCR is used in following applications such as automatic data entry, baking, voice synthesizer, the reading device for visually challenged people.Pattern recognition includes two types, they are online and offline as shown in Fig.

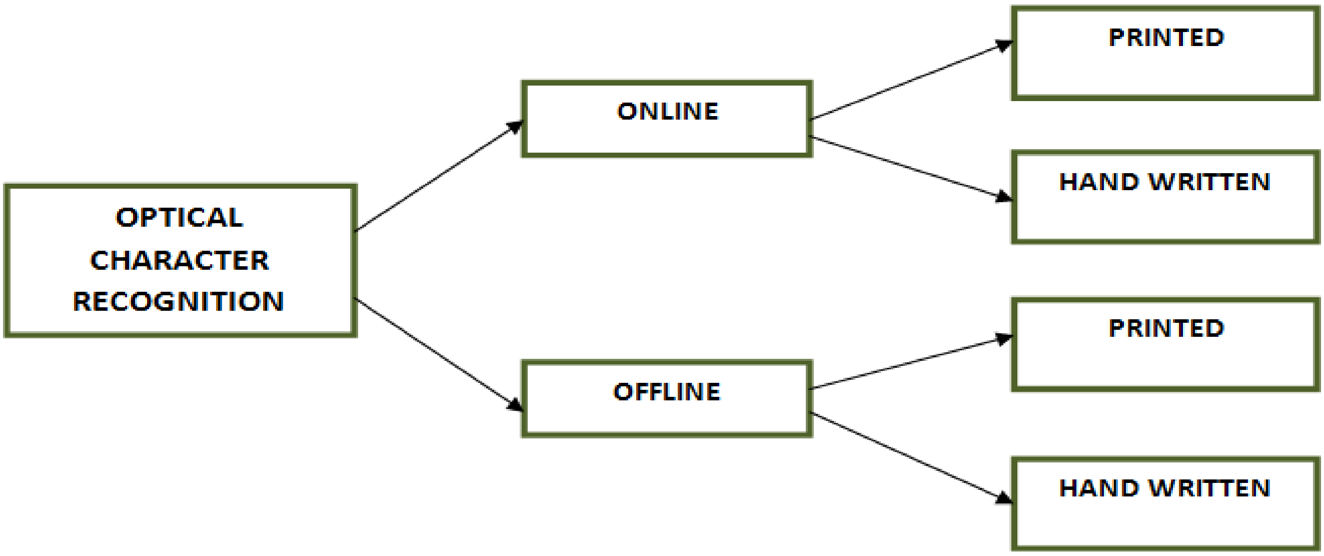

Fig 2: OCR Pattern Recognition

LSTM

Long Short Term Memory networks usually just called "LSTMs" are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hochreiter & Schmidhuber (1997), and were refined and popularized by many people They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. Remembering information for long periods of time is practically their default behavior, not something they struggle to learn!

Literature Survey

A. *Prior Works*

We have used the dataset of EMNIST , there have been several accomplishments that has been achieved using this dataset. Even before using Deep learning, Handwritten text recognition has been made possible, however their accuracies were really low or they had a relatively small dataset as said by Line Eikvil . In this paper, usage of OCR has been discussed such as in Speech Recognition, Radio Frequency, Vision systems, Magnetic Stripes, Bar Code and Optical Mark Reading. A popular machine learning task is classifying the MNIST dataset, which is dataset of numbers. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis by Simard, Steinkraus and Platt is a valuable paper for understanding usage of convoluted neural networks (CNNs). For word recognition, a Paper by Pham et al., used a 2-layer CNN which fed into a bidirectional recurrent neural networks (RNN) with Long Short-Term Memory (LSTM) cells . The best model implemented, according to us, is by Graves and Schmiduber with a multidimensional RNN structure. Another paper on Handwritten Text Recognition' by M.J. Castro-Bleda dealt with dataset with slanted words as well and corrected them during pre-processing. Development of English Handwritten Recognition Using Deep Neural Network by Teddy Surya and Ahmad Fakhrur uses a Deep Neural Network model having two Encoding layer and one SoftMax layer on the EMNIST dataset. Their accuracy using DNN was way better than the earlier proposed pattern net and feed forward net ANN (Artificial Neural Networks). Handwritten text recognition can also be achieved on basis of Relaxation Convolutional Neural Networks(R-CNN) and alternatively trained relaxation convolutional neural networks (ATRCNN) as done by ChunPeng Wu and Wei Fan. Our model achieved accuracy over 87 percent using Convolutional Neural Networks from Keras library.

B. *Dataset Description*

The EMNIST dataset is a collection of handwritten alphanumeric derived from the NIST Special Database 19. Each image is converted to a 28x28 format and dataset structure that directly matches the dataset is used. The training set has 697932 images and test set has 116323 of uppercase and lowercase alphabets and numerals from 0-9 which are mapped to their corresponding classes. The test set and training set is available in the form of list within list. Each item of outer list represents an image and inner list represents the intensity values of 784 pixels (because size of image is 28 x 28 pixels) ranging from 0-255. The test images as well as train images have a white foreground and black background. Both the test images as well as train images are horizontally flipped and rotated 90 degrees clockwise. Y train and Y test both are arrays which contain number ranging from 0 to 61 as there are 10 numerals from 0-9 and 26 uppercase and lowercase alphabets each which adds up to 62.

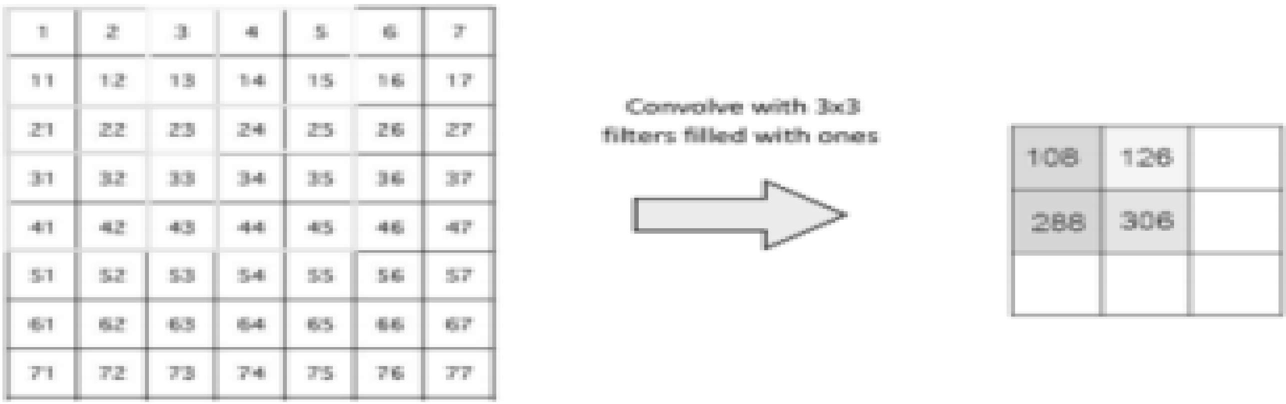

Fig 3: Max Pooling

Existing System

The handwritten Character or Digits are not always of the same size, width, orientation and justified to margins as they differ from writing of person to person, so the general problem would be while classifying the Character or Digits due to the similarity. This problem is faced more when many people write a single Character or Digits with a variety of different handwritings. Lastly, the uniqueness and variety in the handwriting of different individuals also influence the formation and appearance of the Character or Digits . Now we introduce the concepts and algorithms of deep learning and machine learning. Applying the classifiers to real scenario problems. Accuracy and speed of recognition are considered the better measure. The performance of the classifier can be measured in terms of ability to identify a condition properly (sensitivity), the proportion of true results (accuracy), number of positive results from the procedure of classification as false positives (positive predictions) and ability to exclude condition correctly.

The Optical Character identification (OCR) tool will make a prosperous solution to the many real world problems such as data entry problems and the computer vision industry. Therefore, OCR tools are being evolved for almost all major languages in the world level. Different classifiers used with different features and parametric values performing with various accuracies and error rate. Classifiers as k nearest neighbors (KNN), proximal support vector machine, and neural network with different layers can perform well, but best performing classifier on MNIST dataset is convolution neural network (part of deep learning) and the performance is best in this classifier

Proposed System

Deep learning has been widely used to recognise handwriting. In offline handwriting recognition, text is analysed after being written. The only information that can be analysed is the binary output of a character against a background. Although shifts towards digital styles for writing gives more information, such as pen stroke, pressure and speed of writing, there is still a necessity for offline methods, when online is inaccessible. It is particularly necessary for historical documents, archives, or mass digitization of hand-filled forms. Extensive research into this field has resulted in significant progress from classical methods, right up to human-competitive performance. This article serves as an overview of that journey, and the potential future of the field.

Conclusion

We used the EMNIST data set to train our model and tested different optimizers to finally select Adamax as it not only yielded a high accuracy with each epoch on our train data but also our test data. A further application of accurate text OCRs is to help the partially sighted and the blind in the absence of braille. By also integrating a simple text to speech module in the app the user can point his phone to any text which will then read out the text for the user. A dedicated device can also be built for this purpose with a more sophisticated image recognition system which can identify objects to tell the user how many steps to walk in which direction and even when to stop and turn. The EMNIST datasets, a suite of six datasets, considerably increased the challenge faced by employing only the MNIST dataset. Even though the structure of EMNIST dataset is similar to that of MNIST, it provides a higher number of image samples and output classes and an even more complex and varied classification task. It was thus obvious to use it as the backbone of our project. Without the use of EMNIST data set it would be practically impossible to achieve this accuracy. Our current android app requires the user to draw/ write the text on the screen and then analyses it to identify the alphanumeric character. The app can be developed further to import images from the gallery in the user's device and identify the text in present in those images. Another development can be to convert text to speech to further increase the applications of the mobile app. The android app can be developed further using googles cloud natural language API which provides natural language understanding technologies like, sentiment analysis, entity recognition, entity sentiment analysis, and text annotations to understand the text further and better by providing dictionaries that will rectify the mistakes made by the model to provide a meaningful result. Another development can be the use of googles cloud vision API to increase the accuracy of the data read and even to identify different objects.

References

[1]Kai Ding, Zhibin Liu, Lianwen Jin, Xinghua Zhu, “A Comparative study of GABOR feature and gradient feature for handwritten chinesecharacter recognition”, International Conference on Wavelet Analysis and Pattern Recognition, pp. 1182-1186, Beijing, China, 2-4 Nov. 2007.

[2]Pranob K Charles, V.Harish, M.Swathi, CH. Deepthi, "A Review on the Various Techniques used for Optical Character Recognition", International Journal of Engineering Research and Applications, Vol. 2, Issue 1, pp. 659-662, Jan-Feb 2012.

[3]Bhatia Neetu, “Optical Character Recognition Techniques”,International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 5, May 2014.

[4]Liana M. Lorigo and Venu Govindaraju, “Offline Arabic Handwriting Recognition: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 28 Issue 5, May 2006

[5]https://www.math.univtoulouse.fr/~besse/Wikistat/pdf/st-m-hdstat-rnn-deep-learning.pdf

[6]https://medium.com/@himanshubeniwal/handwritten-digit-recognition-using-machine-learning-ad30562a9b64

[7]https://cs.nju.edu.cn/wujx/teaching/15_CNN.pdf

[8]https://arxiv.org/pdf/1511.08458.pdf

[9]https://www.sciencepubco.com/index.php/ijet/article/view/25315/12909

[10] G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, “Emnist: an extension of mnist to handwritten letters, arXiv preprint arXiv:1702.05373, 2017.

[11] HANDWRITTEN TEXT RECOGNITION: with Deep Learning and Android Shubham Sanjay Mor, Shivam Solanki, Saransh Gupta, Sayam Dhingra, Monika Jain, Rahul Saxena

[12]G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, “Emnist: an extension of mnist to handwritten letters,”arXiv preprint arXiv:1702.05373, 2017.

[13]L. Eikvil, “Optical character recognition,” citeseer. ist. psu. edu/142042. html, 1993.

[14]A. Graves and J. Schmidhuber, “Offline handwriting recognition with multidimensional recurrent neural networks,” in Advances in neural information processing systems, 2009, pp. 545–552.

[15]V. Pham, T. Bluche, C. Kermorvant, and J. Louradour, “Dropout improves recurrent neural networks for handwriting recognition,” in Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on. IEEE, 2014, pp. 285–290.

[16]S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. Zamora-Martinez, “Improving offline handwritten text recognition with hybrid hmm/ann models,” IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 4, pp. 767–779, 2011.

[17]T. S. Gunawan, A. F. R. M. Noor, and M. Kartiwi, “Development of english handwritten recognition using deep neural network,” 2018.

[18]C. Wu, W. Fan, Y. He, J. Sun, and S. Naoi, “Handwritten character recognition by alternately trained relaxation convolutional neural network,” 2014 14th International Conference on Frontiers in Handwriting Recognition, pp. 291–296, 2014.

[19]https://www.ijeat.org/wpcontent/uploads/papers/v8i3S/C11730283S19.pdf