International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 5 ǁ 2021 ǁ PP. 01-05

Handwriting Comparison and Character Analysis

[1]Sudhanva Paturkar, [2]Ankeeta Mitra, [3]Sakshi Tannirwar, [4]Varun Shiralkar
[1]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India
[2]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India
[3]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India
[4]Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India
Corresponding Author: Sudhanva Paturkar

Abstract
Traditionally handwriting authentication tools have focused mainly on several mathematical equations which aim to study and chalk out curves and analyse them, however this makes the techniques too focused on one language and is not effective for symbolic languages like Chinese. Instead, we analysed the handwriting by an ocular approach. We used CNN to analyse the text pictorially. We converted the text to patches which were then fed to the layers and the model learned from the training data about how to associate each user to his handwriting by simply comparing patterns in input text and the patterns it analysed from the training data and authenticate the handwriting accordingly. This approach is not only easier to implement but also removes any language barriers. As we are studding pictorial patterns it does not matter if the text in English or Hindi or any other language.

Keywords: *Convolutional Neural Networks, Handwriting, Analysis, Language Independent, Multiclass Classification*

Date of Submission: 27-04-2021 Date of acceptance: 11-05-2021

I. INTRODUCTION

In today's world although the shift towards digitization is increasing rapidly there are many sectors and fields where handwriting still prevails. Many of these sectors are such that handwriting is an essential tool in way the system works for example banks where signatures are indispensable or schools and colleges where handwritten work is believed to hammer in knowledge better or in sectors like Calligraphy where handwriting is everything. In most of these sectors to a certain degree it is important to know which handwriting belongs to whom. Although a trained eye might be able to distinguish handwriting of a set of say 10 students, but when the number reaches in hundreds and thousands the task of distinguishing and classifying handwriting becomes downright impossible. This problem has been approached before where Machine Learning techniques are used to study and analyse complex features of individual's handwriting using equations that study curves and strokes in handwriting. Instead, we adopted a simpler approach where handwriting can be compared optically using Convolutional Neural Networks where the model can itself extract features it requires to distinguish handwritings.

II. LITERATURE SURVEY

2.1 Traditional Approach

Methods for off-line writer identification can be categorized into two groups: text-dependent and text-independent [3]. Text-dependent methods require input image with fixed text contents and which usually compares the input with registered templates for identification. In contrast with this, text-independent methods do not make assumptions on input content and have broader applications. However, compared with text-dependent one, text-independent writer identification needs to deal with image with arbitrary texts which exhibits huge intra-category variations, therefore, and is much more challenging [3].

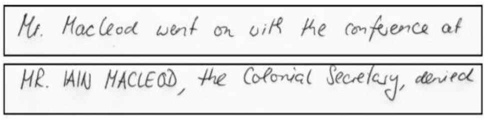

Fig. 1. Two text lines written by writer 009 from IAM dataset

Fig. 2. Two symbols from Chinese language

Fig.1 and Fig.2 shows several examples of handwritten English and Chinese [4] by different writers. As can be seen, the main difference between two handwritten images is dominated by the text contents. For writer identification, one needs to extract abstractive written style features and fine details which reflect personal writing habits. This poses a great challenge for current handcrafted features which usually capture the local shape and gradient information. These handcrafted features may include both information of written contents (text) and written styles (person), which may limit their performance on this task [4].

The traditional approach to solving this would be to extract language dependent features like curvature of different letters, spacing b/w letters etc. and then use a classifier like SVM to distinguish between writers. In the blog, I want to demonstrate a deep learning-based approach to identifying these features. We will pass small patches of handwritten images to a CNN and train it [1].

2.2 Modern Approach

Modern techniques focus on recognizing all the characters in a segmented line of text. Particularly they focus on machine learning techniques which are able to learn visual features, avoiding the limiting feature engineering previously used [1]. State-of-the-art methods use convolutional networks to extract visual features over several overlapping windows of a text line image [1].

Handwriting Recognition has an active community of academics studying it. The biggest conferences for handwriting recognition are the International Conference on Frontiers in Handwriting Recognition (ICFHR), held in even numbered years, and the International Conference on Document Analysis and Recognition (ICDAR), held in odd numbered years. Both of these conferences are endorsed by the IEEE and IAPR [1].

To address the challenging problems, this leverages deep CNNs (Convolutional Neural Network) as a powerful model to learn effective representations for off-line text independent writer identification [3]. Deep CNNs have demonstrated its effectiveness in various computer vision problems by improving state-of-the-art results with a large margin, including image classification, object detection, face recognition, handwriting recognition etc. [3].

III. IMPLEMENTATION

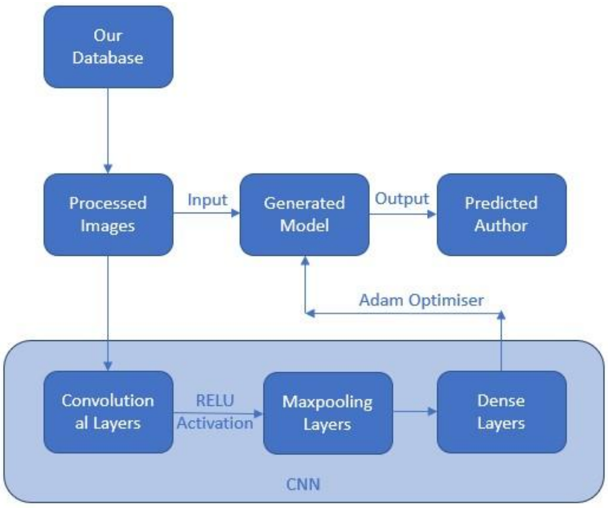

Fig. 3. Block Diagram

To start making our model we first needed a good database. The most popular database available right now is the IAM handwriting database which is free but has a lot of issues. We first developed our model on this database and found out that due to several factors like very low datapoints per author, inconsistent image quality, uneven cropping and several one lettered image, the accuracy of the model suffered a lot. There were a lot of images but the total images per author were only 14. To overcome this, we decided to build our own database with the help of our friends and colleagues. The database though small is sufficient to test our model and act as a proof of concept. We gathered handwriting samples from 10 authors and after processing the data we had about 207 images per author. All authors were asked to write down the same paragraph giving us about 2070 images containing consistent uniform data that then we could use as our database.

Next step was to design a convolution network. It consists of basically three layers repeated according to need. These are the convolution layer, the activation layer and pooling layer. In the convolution layer the data is processed in 4 steps.
Step1: The model first extracts feature from the sample image lines up the features and the image patch which by default is about 9 pixels.
Step 2: Multiply each image by the corresponding feature pixel.
Step 3: Add them up.
Step 4: Divide by total number of pixels in feature.

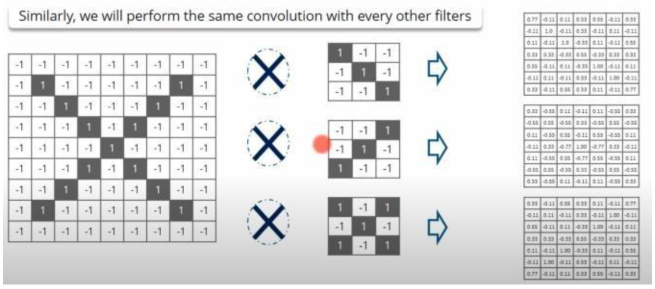

Fig. 4. Process in Convolution Layer. [5]

Next layer is the Activation layer. The activation function is a mathematical "gate" in between the input feeding the current neuron and its output going to the next layer. It can be as simple as a step function that turns the neuron output on and off, depending on a rule or threshold. Or it can be a transformation that maps the input signals into output signals that are needed for the neural network to function. [6].

Our model is a multiclass classification model so we cannot use Binary step function which only gives values in yes or no, so instead we use Rectilinear Activation Function or the RELU function which provides a range of values for all classes.

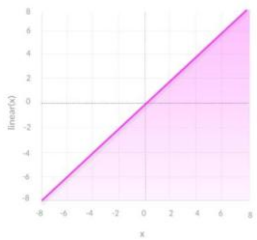

Fig. 5. Rectilinear Activation Function Graph [6]

After passing the data through RELU layer it removes all the negative values and produces an output like this for each feature in the handwriting sample.

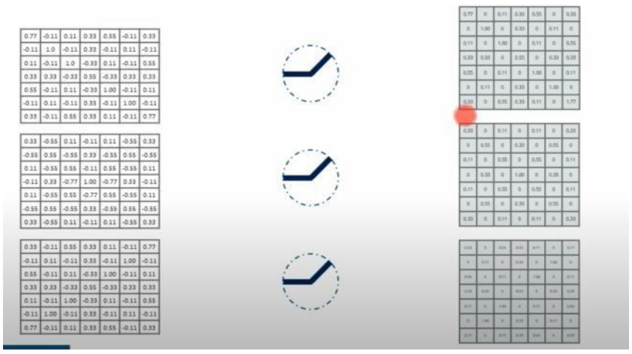

Fig. 6. Output Produced by RELU function [5]

Next comes the pooling layer. The main idea behind a pooling layer is to “accumulate” features from maps generated by convolving a filter over an image. Formally, its function is to progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network. The most common form of pooling is max pooling. Max pooling is done to in part to help over-fitting by providing an abstracted form of the representation. As well, it reduces the computational cost by reducing the number of parameters to learn and provides basic translation invariance to the internal representation. Max pooling is done by applying a max filter to (usually) non-overlapping subregions of the initial representation. The other forms of pooling are: average, general. [7]

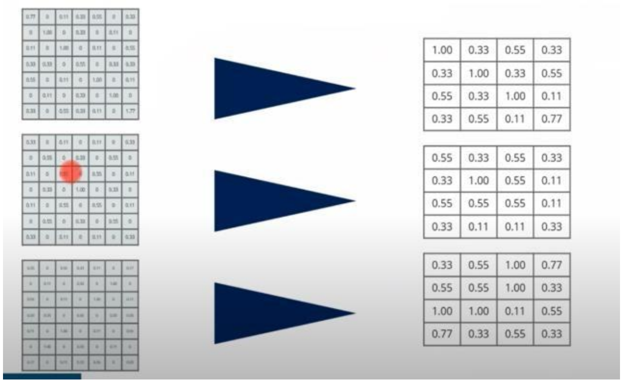

Fig. 7. Data after Max Pooling Layer

Lastly, we repeat the layers again and again. How many times the layers need to be repeated is a relatively trial and error process. We repeated the 3 layers 6 times in our model.

After all these layers the data matrix obtained is passed through dense layers which reduce down the prediction probabilities to our number of classes which in this case would be the number of authors. This gives us a vector with probabilities of the sample belonging to each of the author. If the model has been trained properly more often than not it the value in vector with highest probability will correspond to the author to which it belongs.

Building our model on these lines we tried to make several models with a varying combination of layers and epochs and the one which succeeded in giving most accuracy was when we used 20 layers and trained the model for 25 epochs. This yielded a model with following parameters:

Loss: 5.04e-05

Training Accuracy: 100%

Validation Loss: 0.1494

Validation Accuracy: 96.62%

These parameters were obtained with a dataset of 10 authors with about 207 images per author. We can expect the accuracy to drop a little when training the model for hundreds of authors.

IV. PROBLEMS FACED

There were several issues that arose along the way while training the model. Some were resolved and some still pose a problem that needs to be solved.

1. Proper database for handwriting required to train a multiclass model like ours is not present. Although the IAM dataset is vast and provides data in a bunch of different forms, it fails to deliver quantity per author which is required for training a classification model. We overcame this by compiling our own dataset but as the task in too tedious we were only able to secure a finite amount of data. Further improvement would require a way to extract data by directly scanning pages instead of someone manually separating the data from images.
2. When we do acquire large amounts of data for hundreds of authors the next challenge would be to train the model for such vast amounts of data. Propper resources like a computer with high end GPU and powerful memory were not accessible to us. Although cloud services do offer GPU’s and other virtual machines for exactly this purpose proper documentation to use them without causing a hole in your pocket is not available.
3. Google Collab does provide free GPU and is easy to use but it provides only 12Gb ram which is really insufficient for training a model on large database. Google does provide Google Collab Pro but its available only in USA and Canada.

V. CONCLUSION AND FUTURE SCOPE

A highly accurate model like this has as many applications as one can think of.

1. It can be used in banks and other industries for more accurate signature authentication.
2. It can be trained for a class and then be used by the teachers in schools and colleges for verifying if each work being submitted is coming from the student and has not been copied or written by someone else.

An additional complimentary AI powered tool can be developed which can scan and separate data samples from scanned images automatically. This would eliminate any need for human labor for preparing database and would improve efficiency of the model as a system exponentially.

REFERENCES

[1]. Priya Dwivedi,"Handwriting Recognition Using Tensorflow and Keras" https://towardsdatascience.com/handwriting-recognition-using-tensorflow-and-keras-819b36148fe5

[2]. IAM Handwriting Database http://www.fki.inf.unibe.ch/databases/iam-handwriting-database

[3]. Bulacu M, Schomaker L. Text-independent writer identification and verification using textural and allographic features[J]. Pattern Analysis and MachineIntelligence, IEEE Transactions on, 2007, 29(4): 701-717

[4]. Linjie Xing, Yu Qiao, Shenzhen key lab of Comp. Vis. & Pat. Rec.,Shenzhen Institutes of Advanced Technology, CAS, China. DeepWriter: A Multi-Stream Deep CNN for Text-independent Writer Identification.

[5]. Edureka," Convolutional Neural Networks with Tensorflow" https://www.youtube.com/watch?v=umGJ30-15_A

[6]. DeepAI Blog https://deepai.org/machine-learning-glossary-andterms/max-pooling#:~:text=Max pooling is done to,invariance to the internal representation.

[7]. Missing Link AI Blog https://missinglink.ai/guides/neural-network-concepts/7- types-neural-network-activation-functionsright/#: ~:text=Role of the Activation Function in a Neural Network Model&text=The activation function is a,on a rule or threshold