Handwriting

by B3 Secondreport

Submission date: 26-Dec-2020 05:04PM (UTC+0530)
Submission ID: 1481321430
File name: B3_SecondReportToCheck.doc (54.5K)
Word count: 2947
Character count: 16377

HANDWRITING RECOGNITION AND WRITER IDENTIFICATION ON TEXT INDEPENDENT DATA

Identification of individualsusing theirhandwriting is one of the stimulating problems. It can be used in various number of fields like security and criminological analysis, antique documents and literature analysis, forgery and signature analysis etc. Handwriting has a crucialcharacter in demonstration of identity and culturedtraits of anindividual. It is the keydistinctiveness of an individual. Methodologies which are based on deep learningareverified as the versatiletechniques for extracting features from hugevolumes of diverse data and gives solid andaccurate predictions of the underlying patterns than traditional methodologies. Automatic handwritingrecognition system aids in identifying whether the stated handwriting is accurately matched and allocated to the original writer of that handwriting.

There are two modes of data capturing in any kind of writer identification, they are online and offline. Spatial coordinate values are considered mainly in the former case whereas the temporal details are considered in latter case. With respect to the textual content, offline writer identification can be categorized as text dependent and text independent. Text-dependent procedures highly emphasize on context and semantics in the handwritten image, so it needs an image having fixed text content and evaluates the resemblance of the input image with alreadylistedprototypes for this purpose. Contrarily, text independent handwritingrecognitionemphasizes on the images that do not hang on the text content which is fixed.

Since the convolutional neural networks and their state of art architectures are very powerful in extracting deep features and their huge success in the fields of computer vision made us to implementthe proposed model of writer identification system. It is done in three stages. The first stage is Image pre-processing where various filters are applied on images to remove noise. The second stage is feature extraction using AlexNet based CNN architecture where the patterns are learnt and deep features are extracted in order to accurately classify different handwriting styles and the final stage is writer classification using Artificial Neural Networks (ANN) followed by evaluation of the model.

CHAPTER 1

SUMMARY OF BASE PAPER

Automatic handwriting recognition and writer identification system helps in determining the valid writer of a particular handwritten text image among a large number of writers whose handwriting has been already trained. Many research scholars had shown interest in this domain because of its large number of applications like security and criminological analysis, antique documents and literature analysis, forgery and signature analysis etc. The traditional approaches extract features from handwritings using hand designed and manual calculations. They used techniques like scale invariant feature transforms, local binary patterns, wavelet transforms etc, where the final feature vectors are extracted from local patches of handwriting images and constructed models like bag of words etc. Before the deep learning era, researchers worked on this using hand designed features. With the advent of deep learning algorithms like CNNs, ANNs etc, the feature extraction work has become much easier. Since the convolutional neural networks and their state of art architectures like AlexNet, VGG, GoogleNet etc are very powerful in extracting deep features and their huge success in the fields of computer vision made several research scholars to employ CNN techniques for writer identification.

This paper presents a deep learning based automatic features extraction model using AlexNet CNN architecture. AlexNet is a state of art CNN architecture model trained on ImageNet
Database. In this project the pre-processed handwriting images of 10 different writers are trained on CNN architecture inspired from AlexNet and classification is done using Artificial Neural Networks.

A Convolutional Neural Network is a populardeep learning algorithm that takes an input image, process pixel data, undergoes learning by assigning weights and biases through optimizers and large number of back propagations, to various features in the image and becomes capable of distinguishing one another. The convolutional neural network architecture is very similar to the arrangement pattern of neurons present in the brains of human beings and got inspired from organization of the visual cortex. Receptive field is the visual area where from discrete neurons are restricted to receive stimuli and responds. The visual area is a collection of such fields that gets overlapped.

This paper proposed a model for writer identification, using Convolutional Neural Networks. CNNs are mainly used for their abilities of features extraction. The main feature that the CNN algorithm uses in this writer identification system is 'Raw Pixel'. Apart from the pixel value, the features like i) Spacing between letters, words, sentences and lines. For ex: cursive handwriting letter space is zero ii) Thickness of letter edges, since pressure applied by every writer while writing will be different, this leads to different thickness in letter edges. iii) Curvature along with depth and height of the letters etc

The extracted feature vector is sent to Artificial Neural Networks or Fully Connected layers for performing handwriting classification. The obtained training and testing accuracy are considered as the evaluation parameter for the model. The entire workflow can be briefly divided into 3 stages, each stage has 3-4 sub stages called as pipelines.

Stage 1: Development of the Classification Model

This stage is mainly for building the classification model and its evaluation. It is broadly divided into 4 pipelines. Each pipeline has its own well-defined task.

Pipeline 1: In this stage the main objective is to remove the noise and variations in background lighting, since the images are differently captured by every writer. This is done using Gaussian blur filter and finally extracting the edges alone from the handwriting image. This is done using Canny Edge Detector.

Gaussian Blur filtering is an operation in which a Gaussian filter instead of a box filter is convolved over an image. It is a low-pass filter which eliminates the high-frequency components and reduces the unwanted noise in the image.

The Canny edge detector is an edge detection operation that uses multi-stage algorithm to detect a extensive range of edges in images.

Pipeline 2: The main objective of this stage is Image Augmentation. Image augmentation is a procedurewhichis used to theatrically enlarge the size of a training and testing dataset by creating altered variants of images in the dataset. There are various augmentation techniques like i) Rotation ii) Scaling iii) Random distortion iv) Flipping v) Zooming vi) Horizontal and Vertical shifting, etc

To increase the image samples, for training the model efficiently, the paper proposed two augmentation techniques, rotation and random distortion followed by random patching of images to obtain a size of 224*224. The following are the operations performed on data.

1)50 images by rotation with 12 degrees clockwise and 14 degrees anti clockwise.
2)50 images by rotation with 10 degrees clockwise and 8 degrees anti clockwise.
3)50 images by distortion with grid width 10 units, grid height 10 units and magnitude 12 units.
4)50 images by distortion with grid width 4 units, grid height 4 units and magnitude 8 units.
5)random slices of 224*224 are extracted for every image.

Pipeline 3**:** Creation of appropriate training dataset and testing datasets is the main objective.

In this project, we used 10 writers handwriting, i.e. a total of 10000 images, 1000 images for every writer, of size 224*224

For every writer 800 images are used for training the model and 200 images are used for testing or validating. So, a total of 8000 images are used for training and a total of 2000 images are used for testing. The created datasets are uploaded in Google Drive, which is considered as a database for this project. Uploading in google drive is efficient for accessing data while working in google colab

Pipeline 4**:** Constructing a Convolutional Neural Network model based on AlexNet architecture is essential for training the handwriting images for classification. The entire neural network is designed using TensorFlow and keras library. The network architecture is shown in the figure 1.2 The network is constructed by referencing AlexNet State of Art CNN architecture. It has 4 convolution layers followed by 4 pooling layers, each after every convolution layer. Feature vector is obtained from flatten layer and sent to two fully connected layers. Other configurations of the network are described below.

Activation Functions used: -

In convolution layer:- Relu

In Output layer:- Softmax

Optimizer used: - Stochastic Gradient Descent (SGD)

Loss Function used: - categorical_crossentropy

No of Epochs: - 20

Evaluation Metric used: - accuracy

To reduce overfitting, Batch Normalization and Dropout of 40% are used in CNN and ANN layers respectively.

Stage 2: Creating backend REST API of web application

The main objective of this stage is to build a REST API for the developed CNN model using flask. It consists of 3 pipelines.

Pipeline 1: - The trained AlexNet CNN model is stored in disk and is loaded whenever necessary. So, the main objective of this stage is to extract the Handwriting classification model from the disk and gets loaded with all the pretrained weights and can be used for further testing.

Pipeline 2: - Designing an Image Output Generator (IOG) method, which gives the classification results in the form of 20 votes, which were distributed among the possible writers, for an uploaded handwriting image of any writer. For ex: for any image if the model is 100% confident on a particular writer's handwriting, all the 20 votes are given to that writer, else it gets divided.

Pipeline 3: - Designing a Flask based API having 2 routes

i).‘/’which is a GETroute for displaying home page of the app
ii).‘/predict’which is a POSTroute, that calls predict function and sends the result string to the frontend
Stage 3: Creating frontend user interface of web application

The main objective of this stage is to design a simple user interface for testing the model by uploading handwriting images. It consists of three pipelines.

Pipeline 1: - Creating UI elements for uploading handwriting and making AJAX calls to the REST API to obtain the result for the uploaded handwriting.

Pipeline 2: - Formatting of the result string is done in this pipeline i.e. the writer name which got a greater number of votes out of 20 are considered as 1st priority and displayed appropriately.

Pipeline 3**: -** Designing other UI elements of the front end and rendering them appropriately for a better user experience.

To prove correctness of any deep learning algorithm, there are certain metrics like accuracy, precision, recall, loss, validation loss, validation accuracy etc to evaluate the model and which in turn prove its correctness. Various metrics should be considered in various situations. In image classification using CNN mainly in every epoch we will be considering parameters like accuracy, validation accuracy, loss and validation loss.

Accuracy: - Accuracy is considered as a significantmetricin any classification project. It compares the classified image with another data source that is considered to be accurate or ground truth data. Through field work and experimentation ground truth can be collected; however, this is highly expensive and more time taking. Briefly it can be considered as how much correct a classified result is, with respect to correct result. Higher the accuracy score, better the learning and performance of the model

Validation Accuracy: - It is nothing but accuracy score obtained on validation dataset. It shows how well the model is behaving on the unknown data after getting trained. Higher the validation accuracy, better the classification on unknown images.

Loss:- In order tooptimize any algorithm, the function that is used to figure out a solution, is often considered as an objective function. The main aim is to maximize or minimize this objective function, which means that the solution which we are finding should haveeither highest score or lowest score respectively. Naturally, in neural networks, the error must be as minimum as possible.This objective function is consideredasloss function or cost function and the value obtained by this loss function is simply called as loss.

Validation Loss: - Loss obtained on validation data is considered as validation loss. For a better generalized model, the loss and validation loss should be as minimum as possible. Higher validation loss and lower validation accuracy is a result of overfitting.

While training the CNN model on certain epochs, the following trends can be seen on these values.

1. Validation loss slowly starts to grow and validation accuracy slowly starts falling. This means model is stuffingvalues and not learning

2. Validation loss slowly starts to grow and validation accuracy also slowly raises. This must be a situation of overfitting or miscellaneous probability values in scenarios where SoftMax activation function is being used in the output layer

3. Validation loss starts decreasing and validation accuracy starts increasing. This is also acceptablewhich means that the model built is learning well and working fine.

While training our model we observed the 3rd trend, i.e. validation loss is decreasing and validation accuracy is increasing. This shows that the CNN model is generalized and not overfitting. In the 1st epoch, the accuracy is 0.3341, validation accuracy is 0.1870, loss is 2.0037 and validation loss is 3.4596, while in the final epoch i.e. 20th epoch, the accuracy is 0.9919, validation accuracy is 0.9900, loss is 0.0487 and validation loss is 0.0421. With these metrics we can say that the model is not overfitting and it is a highly generalized model. This proves the correctness of our CNN based Handwriting classification model.

CHAPTER 2

MERITS AND DEMERITS OF THE BASE PAPER

The traditionalmethodologies of offline text-independent handwriting identification have been classified into two types based on the extracted features. The first approach is based onmanually designed features and the second approach is based on deeplearning-basedfeature extraction approaches. Some hand-designed features such as Scale-Invariant-Feature-Transforms, Local Binary Patterns, etc were extracted from the local patches of the handwriting images and models like bag of words has been constructed to represent the obtained final feature vectors.To extract features from handwriting images for writer identification, filters like Gabor filters and local binary patterns were used. Other than this Most of the works are based on traditional methods like finding chi^2 distance between letters, using wavelet transformation which is computationally intensive, and other methods like grapheme generation which is time consuming. Many other researchers like Christlein et al. used CNN algorithms but followed a technique of extracting features from local structures which resulted in high feature loss and low accurate models. Tang and Wu used Bayesian networks and CNN algorithm on ICDAR 2013 dataset.

These methodologies didn't consider, custom (real time) writers' handwritings rather used benchmark datasets handwritings as shown above. They also didn't used any modern State of art CNN architecture models like ResNet, AlexNet, GoogleNet etc which are highly accurate and powerful.

In this paper, we follow a technique of extracting features from Global structuresthan local structures, due to which the feature loss is minimized and performance gets improved. Apart from it we used a CNN architecture inspired from the famous State of Art CNN architecture, AlexNet.
The entire project is carried out on custom writers than using the existing old benchmark datasets like QUWI, ICDAR-13 etc., on which more research has already been done. The proposed model performs better than existing models in the areas of Image Augmentation, efficient ways of storing the images, better CNN architecture with a very high accuracy etc.

Analysis on Image Augmentation techniques:

The proposed scheme used basic as well as latest image augmentation techniques like rotation and random distortions respectively. With this, the final datasets have high variant images which is very useful for avoiding overfitting and feature loss. With the technique of random distortion in the line alignments, we can easily relate it with the daily life scenario where writers always won't write uniformly. We followed random cropping technique over sliding window technique in order to ensure the capturing of as many features as possible. With sliding window technique there is loss in features like word gap, curvature and line height etc. It also avoids overfitting of the model.

Architecture Analysis**:**

The architecture is built by taking reference of AlexNet CNN architecture. Since the handwriting images that we are using in this paper are not a part of the ImageNet dataset, the actual weights used in AlexNet architecture didn't gave a valid accuracy. Hence Transfer learning techniques cannot be applied directly. Instead, tweaking the configurations like number of filters, strides etc in each layer by keeping the same architecture worked well for us. Usage of Stochastic Gradient Descent optimizer over other optimizers like Adam which was used in other existing works gave a relatively less loss. The number of convolution layers and pooling layers can still be optimized.The total trainable parameters can still be reduced.

CHAPTER 5

CONCLUSION AND FUTURE PLANS

Automatic HandwritingIdentification is one of thefascinating research problems in the fields of document analysis and criminological analysis etc. The effective execution of handwriting identification systems can be applicable in banks, check processing, historical and forensic analysis,signature identification, graphology, legal documents and ancient manuscripts etc. In this paper we have implemented a classification model based on AlexNet state of art CNN architecture model on ten different writers and got a decent accuracy and validation accuracy without any overfitting issues. These results and analyses show the correctness of the implemented model and shows the productiveness of the proposed implementation for writer identification.

Apart from AlexNet, there are many state of art CNN architecture models like VGG, GoogleNet, ResNet etc. Implementing the same model using any other such architecture can be worked out. Emphasis can be further increased in optimization of architecture by using techniques like Hyper Parameter Tuning, Normalization, Effective dropout ratios etc. Constrains on uploading user input images can be improved, so that unnecessary variations that occur in line spacing while capturing the photo can be removed. Dataset size and number of writers can be further increased. The performance of the model can also be tweaked using other optimizers and loss functions with different number of epochs and can be compared with the implemented model's configurations.

Handwriting

ORIGINALITY REPORT

3%	2%	2%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	desktop.arcgis.com Internet Source	1%
2	ijircce.com Internet Source	1%
3	Li Li, Dawei Qi, Jingwei Song, Hongbo Mu. "A Method for Nondestructive Testing of Wood Defects Based on Fractional Brownian Motion", 2007 IEEE International Conference on Control and Automation, 2007 Publication	<1%
4	mafiadoc.com Internet Source	<1%
5	"Advances in Signal Processing and Intelligent Recognition Systems", Springer Science and Business Media LLC, 2014 Publication	<1%

Exclude quotes On

Exclude matches Off

Exclude bibliography On