International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Approaches for Offline Cursive Handwritten Character Recognition

Varsha Vishwakarma, Hylish James

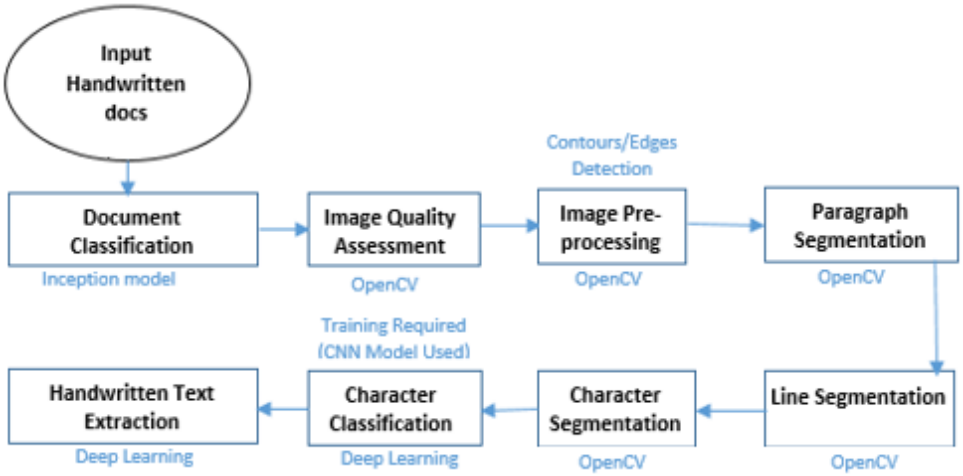
Apar Technologies, Bengaluru, India

Abstract: ***Handwritten Character Recognition (HCR) plays an important role in the retrieval of information from pixel-based images to searchable text formats. For instance (HCR) nonlinear normalization of character size, and feature compression of higher dimensional original features are studied as pre-processing and feature extraction techniques for a statistical character classifier to improve the recognition accuracy of handwritten character recognition. A high speed pre-classification technique using a linear discriminant function is employed to improve the recognition speed. The nonlinear normalization is also utilized to supply the lack of the training samples by artificially generating character samples. The Inception V3 network is trained with character images consisting of noises which are collected from receipts and newspapers. Analysis and discussion were also made on how the different layer's properties of neural network affects the HCR's performance and training time. The proposed deep learning based HCR has shown better accuracy than conventional methods of HCR and has the potential to overcome in the text.***

Keywords: Segmentation, Handwritten Characters,Classification

1. Introduction

The ability to perform human functions such as reading machines is an ancient dream. However, over the last few years, reading by a machine is no longer a dream and has grown to become a truth. Text character recognition commonly deals with the recognition of handwritten processed characters which is also called as Handwritten character recognition (HCR). The basic idea of HCR is to convert any hand written text into data files that are able to be edited and read by machine. With HCR, any handwritten text can be scanned directly and the editable text format can then be easily converted from a computer. The HCR system has two major advantages which are the ability to increase productivity by reducing staff involvement and storing text efficiently. More generally, the areas where this system can be applied are postal departments, banks, publication industry, government agencies, education, finance, health care. The universal HCR system consists of three main steps which are image acquisition and preprocessing, feature extraction and classification. Image preprocessing phase cleans up and enhances the image by noise removal, correction, binarization, dilation, color adjustment and text segmentation etc. Feature extraction is a technique for extracting and capturing certain pieces of information from data. In the classification phase, the portion of the divided text in the document image will be mapped to the equivalent textual representation. Nowadays, there are several existing HCR solutions which are commonly used in machine learning research and pattern recognition. Unfortunately, there is still a challenging problem for recognizing the cursive font or faded English characters. The performance of HCR directly depends on the quality of input image or thus making the character recognition in scene images is potentially far more complicated. In addition, English characterswith poor quality are typically obtained from old printed documents that are usually caused by damaged print cartridges. Unfortunately, these training samples are yet to be found in the existing solution. In order to recognize poor quality English characters, an improved HCR with sufficient training data distribution is needed. In traditional machine learning research, many people think that the feature vectors of test and training data are provided from the same source. However, this may not be truth in some of the HCR research cases. In the concept of transfer learning, training samples can be used to pre-train a network in the source domain, and these well-trained learning characteristics can be delivered and benefit from the training



process in the target domain of the second network. In recent years, traditional methods in the field of HCR research have been almost substituted by deep learning methods such as Convolutional Neural Networks (CNN). The CNN to learn image representations on a large annotation dataset can adequately transfer this information to othervisual recognition tasks with a limited amount of training data. To add an adaptation layer in CNN using transfer learning, which achieves performance improvement in historical English character recognition tasks. The existing conventional HCR with machine learning is trained based on hand-written text. There is still a challenge for poor quality (broken, blurred and incomplete) English text character recognition. In addition, due to insufficient labeled training samples of poor quality English character, neural network used for HCR will suffer from imbalanced training data distribution issue. However, the data labeling process requires new train data which is very costly to train up a new network to recognize poor quality text characters. The process will also consume a huge amount of training time as well. Furthermore, there is also another challenge where the performance of deep neural network will potentially be affected by the new training data distributions. For example, a neural network is pre-trained to recognize good quality of "O" character, and then if the network trained again with different "broken" pattern of poor quality of "O" character, the weights adjusted in the network will actually negatively be affected by the new training data. If training with overrepresented class distributions, this will cause the performance of neural network to degrade. From the challenges stated above, the classification and training data distribution is the most crucial stage and challenge. The objective is to collect training materials with a set of blur, incomplete English text characters in images. Next is to develop an HCR method by using deep learning neural network approach. Moreover, investigate the method that will achieve high accuracy while reducing the training time. However, some existing solutions such as OpenCV will be used and applied for the image processing part (segmentation and filtering). While going through preprocessing, each image will segment out the text character instead of word, so as similar to the input for classification. The text language is English only, where the font size and type of the text dataset are typically depend on the data-set's resources. In addition, the dataset will be collected from check books, documents and receipts.

Image Pre-processing

The aim of pre-processing is to eliminate the inconsistency that is inherent in cursive handwritten words. The handwriting samples may be written on a noisy or colored background and also the quality of the images may be degraded due to the bad resolution that is introduced in the process of scanning or capturing the images. It is necessary to improve the quality of the images by applying pre-processing rules.The outcome of the pre-processing techniques, which has been employed in an attempt to increase the performance of the segmentation process.

Image Scaling

This helps to reduce the time of training of a neural network as more is the number of pixels in an image more is the number of input nodes that in turn increases the complexity of the model. Other advantage of image scaling is that it also helps in zooming in images for better understanding. Opencv which provides some interpolation method for resizing like "cv2.Inter_Area" which helps to reduce an image, "cv2.Inter_Cubic"which helps to zoom an image where slowly and more efficient way."cv1 InterLinear "also helps to zoom an image faster.

DE noising

The main aim of an image denoising algorithm is to achieve both noise reduction and feature preservation. Likewise "cv2.fastNIMeansDenoising()" works with the single grayscale image and "cv2.fastNIMeansDenoisingColored()" works with the color image. Also Mean and Median filters helps for denoising in image,where Mean filter will be easy to implement the method of smoothing the images for salt and pepper noisereducing. The amount of intensity variation between one pixel and the next which causes blurred image.Median filter used to reduce noise in an preserving edges in the image.The median is calculated by sorting all the pixel values from neighbourhood into numerical order and then replacing the centre pixel being considered with the median pixel value.

Perspective Transformation

Detecting the corner points based on intensity variation by drifting the topmost row towards down, until the first high intensity pixel variation has detected and store the point and then repeat the process by drifting from bottom row towards upward . Similarly the extreme left and right columns are drifted towards right and left side and stored the other three points eventually.

Text skew correction

The deskew can be applied to detecting the block of text in the image and computing the angle of rotated text and also for rotating the image to correct for the screw.It increases the rate of character recognition accuracy because aligned text is much closer to encounter when performing the image analysis. Several approaches have been proposed for skew angle detection of document images which is Hough transform, projection profile, nearest neighbour.

Normalization

The process which helps to change the intensity values of the pixel to the range [0,1].where the complication occurs during the feature extraction unless the normalization has covered in earlier stages.Normalization also used to generalize the preprocessing algorithms suitable for different images.For example, the intensity range of the image is 30 to

150 pixel and the desired range is 0 to 255 pixel, the process entails subtracting 30 from each of pixel intensity, making the range 0 to 120 pixel. Each pixel intensity is multiplied by 255/120, making the range 0 to 255 pixel.
Inew=(I-I.min)*(newmax-newmin)/(I.max-I.min)+newmin

Morphological Transformations
Thinning is the process to eliminate the selected foreground pixels from binary image format(where always try to keep foreground in white). It erodes away the boundaries of foreground by convolving the kernel filter, through the image and will be considered as 1 alone. If all the pixels which is under kernel 1 or it is considered as 0. This will help to reduce the complexity processing time and eliminate noise from an image. Once the white noise removed from an image, it will increase the object area by convolving the kernel through the image and will be considered as 1.

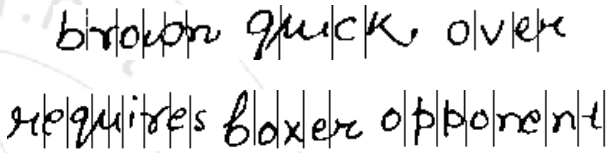
Segmentation
This applies the canny edge detector to the image which produces white pixels, which has an edge in the original image, and removes most of the background noise from the image and turns the text regions into bright clumps of edges which contain both the image borders and the text. vertical and horizontal rank filter helps to eliminate the borders.The text areas represent many white pixels, where the borders contain only the thin 1-pixel line.The problem with rank filter is that can remove only the borders, where a small piece of text outside the borders exist.To overcome from this problem, will apply contours tracing at the edge of image where pixels are connected to one another and marked out everything, otherwise apply minAreaRect to find all white points and extract the minimum area region from an image.The point which represent for a crop to extract the text region from an image which apply the binary dilation to the de-bordered edge image. Binary dilation process bleeds the white pixels into one another. This has repeated until few connected components represented. The text areas has all the bled into a few components. This components are ordered, as which contains the most white pixels.

Precision/Recall Trade-off
Recall is the fraction of white pixels inside the cropping rectangle and precision is the fraction of the image which is outside of the cropping rectangle. A standard way to solve precision/recall problems is to optimize the F1 score.While cropping the de-bordered edge image, keep adding the components which increases the F1 score and help to focus on important content rather than the noise.

Line segmentation
To find upper and lower bounds to extract the lines from paragraph to compute the line segmentation algorithm is to find points of lines separating upper and lower text fields. Most often those steps uses horizontal projection profiles in where the amount of black ink is summed over the x-axis to obtain a text areas has more or lesser black ink not presented.

Word segmentation
After getting the text lines, words and sub-words are separated by spaces, when word boundaries are always represented by a space. According to this, the distance between each pair of consecutive sub-words are obtained. Normally the distance between words are larger than the distance between sub-words, thus words can be segmented by comparing this distance against a suitable threshold. To determine such a threshold, Bayesian criteria of minimum classification error is employed a given distance of d, the probability that represents a separation of word or sub-word is denoted as pw(d) and ps-w(d) respectively.

Character Segmentation
The Off-line handwritten word segmentation is a subject of much attention due to the presence of many difficulties. There can be variation in shapes and writing styles of different writers and cursive nature of handwriting i.e. two or more characters in a word can be written connected to each other. characters can have more than one shape according to their position inside the word image. words may be written by a pen having ink of different colours and some characters in a handwritten word image can have similar contours (‘u’ and ‘v’).There are two types of characters in English language.



First type of characters are called closed characters and contain a loop or semi-loop such as ‘a’, ‘b’, ‘c’, ‘d’, ‘e’, ‘g’ etc. Second type of characters are termed as open characters and are without a loop or semi-loop e.g. ‘u’, ‘v’, ‘w’, ‘m’ etc. In case of cursive handwritten words, a ligature is a link (small foreground component) which is present between two successive characters to join them.To overcome such type of challenges, a new segmentation approach is developed which is based on the analysis of the character’s geometric features, such as the shape of the character to identify the characters and the ligatures. Height and width of the word image is calculated for the analysis of ligatures.When the two consecutive characters in the word image are not touching each other and the sum of the foreground pixels of the columns in this area are 0.When there is a ligature between two consecutive characters or there is a ligature-within-character over-segmentation is eliminated to a great extent by taking average of those segmentation columns which are at a distance less than a particular value (threshold) and by merging them into a single segmentation column. The threshold value is the minimum distance along the width of the word image and its value must be less than the width of the thinnest character possible.

2. Conclusions and Future Scope

The proposed segmentation approach guaranteed correct segmentation when characters in a word image were not touching each other. In future, better pre-processing techniques will be used to enhance the quality of the images specially the character segmentation technique to validate the correct segmentation points. Also will be used CNN model for word recognition instead of character recognition to improve the handwritten text accuracy in future work.

References

[1] Peng Y., Kou G., Shi Y. and Chen ZA. Descriptive Framework for the Field of Data Mining and Knowledge Discovery , InternationalJournal of Information Technology & Decision Making, 2008, Vol. 7, Issue: 4, Page 639 682.
[2] Rehman A, Mohamad D, Sulong G. Implicit Vs Explicit based Script Segmentation and Recognition : A Performance Comparison onBenchmark Database, Int. J. Open Problems Compt. Math., 2009, Vol. 2, No. 3, 352-364.
[3] Rahul Rishi, Amit Choudhary, SavithaAhlawat, 2013, A New Character Segmentation Approach for Off-Line Cursive Handwritten Words.
[4] Zheheng Rao, Chunyan Zeng, Minghu Wu,2018,Research on a handwritten character recognition algorithm based on an extended nonlinear kernel residual network.
[5] Al Hamad HA, Zitar RA. Development of an efficient neural-based technique for Arabic handwriting recognition, Pattern Recognition,2010, (43) 2773-2798.
[6] Cheriet M. Reading cursive script by parts. In: Proceedings of the 3rd international workshop on frontiers in handwriting recognition,Buffalo, 1993, 25 27 May, pp 403 408.
[7] Han K, Sethi IK. Off-line cursive handwriting segmentation. In: Proceedings of the 3rd international conference on documents analysisand recognition, 1995, pp 894 897.
[8] LeCun Y, Bengio Y, and Hinton G., "Deep learning," Nature, 521(7553), 436-444, 2015.
[9] Veloso LR, Sousa RP, De, Carvalho JM. Morphological cursive word segmentation. In: Symposium on computer graphics and imageprocessing, 2000, XIII, Brazilian, pp 337 342.
[10] Zhang X Y, Bengio Y, Liu C L., "Online and offline handwritten Chinese character recognition: A comprehensive study and new benchmark," Pattern Recognition, 61, 348-360, 2017.
[11] Yamada H, Nakano Y. Cursive handwritten word recognition using multiple segmentation determined by contour analysis. IEICE TransInfSyst, 1996, E79-D:464 470.
[12] Marti U, Bunke H. The IAM database: An English sentence database for off-line handwriting recognition, Int. J. Doc. Anal. Recognit,2002, (15): 65 90.
[13] Hull JJ. A database for handwritten text recognition, IEEE Trans. Pattern Anal. Mach. Intell.; 1994, (16):550 554.
[14] Cheriet M. Reading cursive script by parts. In: Proceedings of the 3rd international workshop on frontiers in handwriting recognition,Buffalo, 1993, 25 27 May, pp 403 408.