IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Mean Average Accuracy for Text And Non-Text Images Using SVM Classifier

Chowdhury Md. Mizan[1], Pradipta Karmakar[2*], Sayak Dasgupta[3], Tridib Chakraborty[4]

[1,2,3,4]Department of Information Technology, Guru Nanak Institute of technology, Kolkata, India

[*]*Corresponding Author: pradipta23karmakar@gmail.com, Tel.: 8902781676*

DOI: https://doi.org/10.26438/ijcse/v7i5.16481651 | Available online at: www.ijcseonline.org

Accepted: 26/May/2019, Published: 31/May/2019

Abstract— Recognition of text and non-text images is a major challenge in the field of computer vision so as to efficiently extract the text from that image. The algorithm used for the extraction of the text from the images would have a higher efficiency if it is known beforehand that the image is a text image or a non-text image. However, there are many images such as old manuscripts where the extraction of the text becomes very difficult. In that case, the algorithm for the distinction between the text and non-text becomes very easy for text detection and have high accuracy and fast in detecting the text from the image. This method can also be applied to detect and extract the text from the signboards also. In our approach, we had built a system that takes any sort of image as an input. After the input of the image, it is then processed and converted into a binary image. Distance transform method is then applied and the measure of the distance between the various points in the image are then calculated. From the calculated points, duplicate points are merged into one point and are sorted in ascending order. The total area of the binary image is then calculated and also the image corresponding to each of the distance transform points are then calculated. The total area of the binary image is then divided by each of the area value of the corresponding distance transform points are the value extracted is known as the feature values. After getting all the feature values the whole value is then divided into small intervals and is then processed through the classifier. The accuracy of the classifier is then calculated and evaluated. This method is a very simple and accurate method for the calculation of the average accuracy of purely text and purely non-text images which can be further used to distinguish between text and non-text images. Experiment have been done with simple text and non-text image dataset and the efficiency of the proposed method is then demonstrated.

Keywords— text and non-text, distance transform, SVM classifier

I. INTRODUCTION

The text is an important part of the image. In some cases, it is clearly understandable, while in others it becomes very difficult to extract the text from that image. The work for the extraction of the text becomes more difficult if it is the case with the video. A video is nothing but a collection of image frames together which are displayed at a regular interval of time so as to create the persistence of vision in the human mind. As a result, the work becomes tedious for extraction of the text from the video frames. So it would much easier for one if they are able to distinguish between an image containing text and an image that does not contain any text before the extraction of the text from that image. This process makes it less complex as it knows beforehand that the image contains any sort of text or not.

We know that there are many functions and algorithms which can detect a text from an image. But sometimes it happens that there efficiency is not as good as it is expected to be. This is because some structures in natural images seem to represent a text. The algorithm tries to extract that text from that image. But unfortunately, it is unable to find any sort of text and at last, after a long time, it fails to extract any text. As a result, it takes a lot amount of time for the extraction of the text from that image which actually does not contain any sort of text. Also in some old manuscripts of ancient times the ancient symbols sometimes represent a text but later it came to be known that it does not represent any text.

Fig 1:- Purely text image

Fig 2:- Purely non-text image

In the recent era of the technology, it also has become an important factor for calculation of the average accuracy of the both text and no-text images. This is because it has been observed that in some purely text based images, the text is not clearly visible. Even if the text is filtered and made clear, then also in some cases the text is not visible. This may be due to the poor quality of the paper which is a result of the passage of the time. Also the viewing angle of the text is not that much clear and sometimes it happens the text is not proper horizontally aligned. Due to such reasons, it becomes important to compute the mean average accuracy of the text and non-text images. This would help us in identifying the text in the image which would further help us to distinguish between a text and non-text based image. However, the mean average accuracy can only be calculated with the help of the confusion matrix. So in our case we have chosen the SVM classifier for the computation of the average accuracy of both the text and the non-text based images.

In our proposed method, in determining the average accuracy of image i.e., it is a text and non-text, we have obtained a graph for both text-based and non-text based image and it has been observed that both the image represented a different graph. However, the data was trained with SVM classifier and the confusion matrix was obtained. From the confusion matrix the average accuracy percentage of both the text and the non-text images were calculated. In the following sections, we have reviewed the old work, propose the new classification scheme and discuss the results that are obtained.

II. REVIEW WORK

There are many methods which have been proposed both for the purpose of extraction of the text from images and distinction between the text-based image and non-text based image. However, among them we present a few studies as follows:-

- In the year 2016, authors Najwa-Maria Chidiac, Pascal Damien and Charles Yaacoub presented a paper in which they have applied the technique of applying both the method of MSER and OCR for the extraction of the text from the image.
- In the year 2015, authors Radhika Patel and Suman K. Mitra presented a paper in which they have used the intensity values for separation of the text and the non-text area and then recognize the text from the image. The algorithm produces the best results on the Gujarati degraded document images.
- In the year 1999, authors R. Malik and SeongAh Chin presented a paper in which they they have used an algorithm which is used to extract the text blocks from the whole image and then extract the text from that image. Besides this the stroke width factor is also used for extraction of the text from the image.
- In the year 2017, authors Sezer Karaoglu, Ran Tao, Theo Gevers and Arnold W M Smoulders, presented a paper in which they assigned the images to different classes and then extract the text according to the scene and is the most successful method in both the fields of classification of the image and text retrieval.
- In the year 2015, authors Chengquan Zhang, Cong Yao, Baoguang Shi, and Xiang Bai presented a paper in which they have combined the three techniques MSER, CNN and BOW and the experiment have been performed in natural images in many varieties of scenario.
- In the year 2018, authors Vishal Chowrasia, Sanjay Shilakari and Rajeev Pandey presented a paper on "Implementation of Optical Character Recognition Using Machine Learning" in which they have applied the OCR technology after the feature extraction from the images.

III. PROPOSED METHOD

In our approach, we have built a system which can take the input of both images and video frames of many different varieties. The output of the system is that it is able to distinguish between a text-based image and a non-text based image. The steps of our proposed method are as follows:-

- At first, the image or the video frame which is intended for the distinction between text and non-text is taken as input into the system.
- The image is then converted into a black and white image which makes it easier for the extraction of the feature values from the image as the intensity levels become very much clear.
- After the conversion into the black and white image, the distance of the separation of the points in the image is then calculated which is also known as the Euclidean Distance Transform.
- After the calculation of the distance transform values, the values without repetition are selected and are sorted in ascending order.
- In the next step, two image areas are calculated. At first, the total area of the binary image was calculated and then the area formed by each of the corresponding distance transform points was calculated one by one. After the computation of all the area, the area constituted by the distance transform points were

divided by the total area of the binary image and were termed as feature values.

- After the extraction of the feature values, the values are then divided into various intervals such as 0.0000-0.0999, 0.1000-0.1999 and so on. And a bar chart has been plotted for the maximum values obtained from the specified intervals. The bar chart exhibits that the text-based image has a parabolic curve while the non-text based image has a non-parabolic curve. From here only we can somewhat infer about the type of image. For more interpretation, the analysis was further conducted.
- After the wrenching of the highest values among the feature values that fit into the corresponding periods, the value is then saved in an excel sheet in order to form a dataset. There were two datasets formed one for an image wholly containing text and another which is wholly a non-text based image.
- After the formation of the dataset, the whole dataset was fetched into a classifier and trained and then the execution of the classifier was then judged and the consequences were explained.

IV. FLOW CHART OF THE PROPOSED ALGORITHM

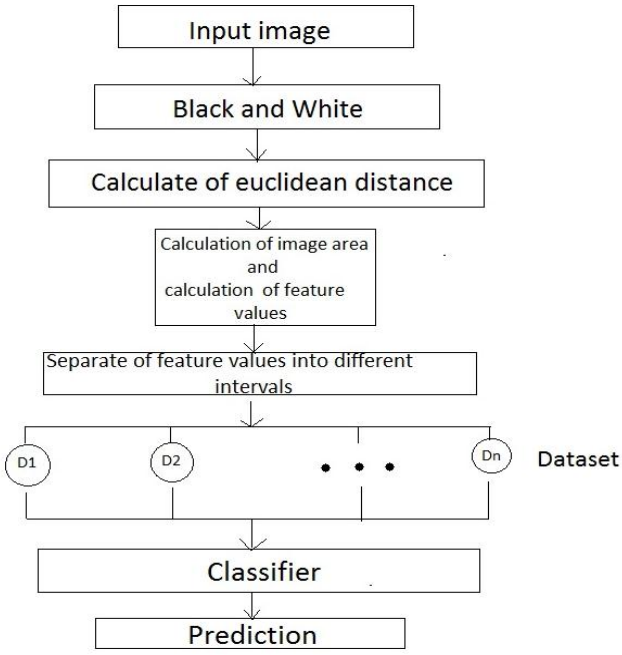

Fig.3

V. RESULTS AND DISCUSSION

The analysis was implemented using purely text and non-text based image set. A set of images was fed into the system and the bar chart, which was obtained from the extracted feature values of the images were examined and analyzed. The analysis was further extended to study the average accuracy of both the images for the distinction between the text and non-text based images. The classifier SVM was employed to study the matter of the subject so as to draw a clear understanding of the matter. The result from the images obtained are as follows:-

(i). Text- based image

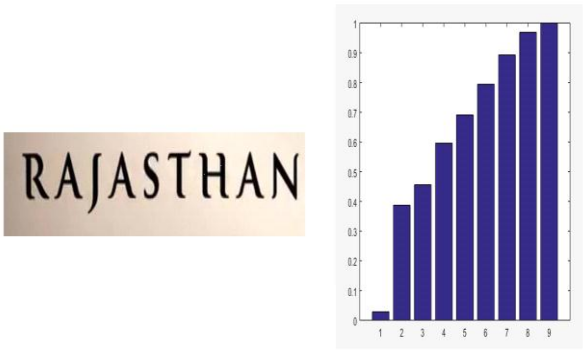

Fig:4 Bar chart of text based image

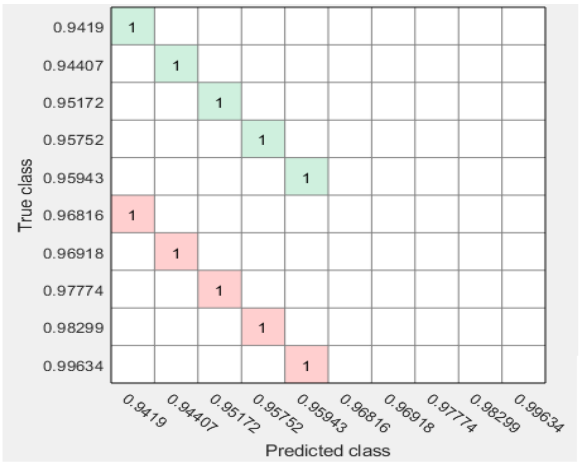

Fig:5 Confusion matrix of text image in ensemble classifier

(ii). Non-text based image

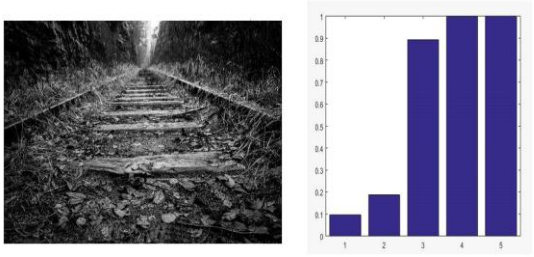

Fig:6 Bar chart of Non-text image

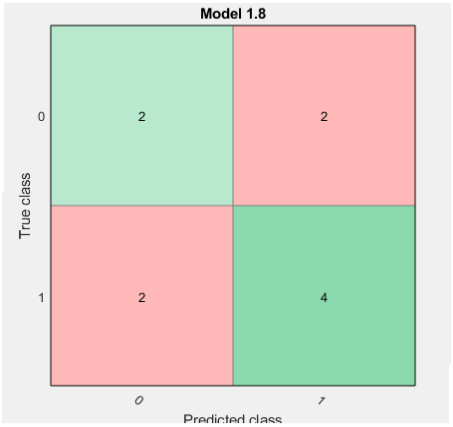

Fig:7 Ensemble classifier of a non-text image

From the results obtained from the text based image, we can infer that the pattern of the curve obtained by joining the top points of a bar chart is a parabolic curve. It gradually increases from zero to the highest level. When this data is fed into a classifier, it has been observed that and for SVM classifier the mean average accuracy is about 50%.

On the other hand, the results obtained from a non-text image, we can infer that the pattern of the curve obtained by joining the top points of a bar chart is a non-parabolic curve. It starts from zero and rises irregularly to the highest point. When this data is fed into a classifier, it has been observed that for SVM classifier the mean average accuracy is about 60%.

In this study, we have presented the approach of distinguishing between the text and the non-text image using the both the ensemble and SVM classifier. This proposal is totally based on extraction of the feature value from the images and then fetching this value into a classifier. After the training of the classifier, the accuracy of the classifier is then observed and studied. MATLAB has been used to develop the software to improve the efficiency of the method.

VI. CONCLUSION AND FUTURE SCOPE

From the whole experiment we conclude that SVM (Support Vector Machine) is capable of distinguishing between a purelty text and a purely non-text images and also with an accuracy of about 50% for the text image and 60% for a non-text image. Though, there are many cases in which the experiment fails to perform but in 80% of the cases the experiment has succeeded in classifying between the text and the non-text images. The work can be more improved by optimizing the whole code and also the for loops which may lead to the better performance of the system. In future we intend to apply this method for the extraction of the text from an image as locating and detecting the text is prior to the extraction of the text. This method will help to classify between the text and the non-text image which in turn will help in optimizing of the text extraction algorithms and thus will have a better performance in extraction of the text from the image.

REFERENCES

[1]. Najwa Maria Chidiac, Pascal Damein and Charles Yacoub, "A robust algorithm for text extraction from images", 39th International conference on Telecommunication and Signal Processing, 2016.
[2]. Radhika Patel and Suman K Mitra, "Extracting text from degraded documents", 5th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, 2015.
[3]. R. Malik and SeongAh chin, "Extraction of text in images", Proceedings of International Conference on Information Intelligence and Systems, 1999.
[4]. Sezer Karaoglu, Ran Tao, Theo Gevers and Arnold W. M. Smeulders, "Words matter: Scene Text for Image Clssification and Retrieval", IEEE transactions on multimedia, vol. 19, no. 5, may 2017.
[5]. Chengquan Zhang, Cong Yao, Baoguang Shi and Xiang Bai, "Automatic discrimination of text and non-text natural images", 13th International Conference on Document Analysis and Recognition, 2015.
[6]. Vishal Chowrasia, Sanjay Shilakari and Rajeev Pandey, "Implementation of Optical Character Recognition Using Machine Learning",International Journal of Computer Science and Engineering, Vol.-6, Issue-6, jun 2018.

Authors Profile

Mr. Mizan chowdhury pursued Bachelor of Technology from Narula Institute of Technology in Computer Science and engineering. He has also pursued his Masters of Technology from Calcutta University in Information technology. Currently he is pursuing his PhD in Computer Science and Engineering from Aliah University and working as a Assistant Professor in Guru Nanak Institute of Technology.

Mr. Pradipta Karmakar is a Final Year student of Information Technology and is persuing his Bachelor of Technology from Guru Nanak Institute of Technlogy. He has completed his Higher Secondary exam from North Point Senior Secondary Boarding School. He is currently the member of Institution of Engineers.

Mr. Sayak Dasgupta is a Final Year student of Information Technology and is persuing his Bachelor of Technology from Guru Nanak Institute of Technlogy. He has completed his Higher Secondary exam from Ichapur Anandamath Vidyapith. He is currently the member of Institution of Engineers.

Mr. Tridip Chakraborty pursued Bachelor of Technology from St. Thomas College of Engineering and Technology in Computer Science and engineering. He has also pursued his Masters of Technology from Jadavpur University in Information technology. Currently he is pursuing his PhD in Artificial Intelligence and Neural Network and working as a Assistant Professor in Guru Nanak Institute of Technology.

