



NATURAL SCENE TEXT DETECTION AND RECOGNITION USING NEURAL NETWORK APPROACHES

Vandana Phatak

Research Scholar, Computer Science & Engineering, MGCGV, Chitrakoot, Satna, M.P.

Prof. Bharat Mishra,

Professor, Department of Physical Science, Faculty of Science & Environment,
MGCGV, Chitrakoot, Satna, M.P.

Dr. Mamta Lambert,

Associate Professor, Department of Computer Application, Jabalpur Engineering College, Jabalpur,
MP, India

Abstract –

It is pertinent to maintain that in the present framework we shall be using following series of processes which are very significant in their own rights- character scanning, pre-processing, segmentation, feature extraction, classification. Consequently these process may be combined together to arrive at final conclusion, which is our final destination. In other words we may say that a final process will essentially combine the results from all the sub processs to produce a final result. One more important point which is of great relevance and that is as follows: initially we shall be providing sufficient input to the system but in turn that output itself will be playing the role of input.

Keywords: Image Enhancement, NN, Text Detection.

DOI Number: 10.14704/nq.2022.20.8.NQ22988

NeuroQuantology 2022; 20(8): 9703-9720

1 INTRODUCTION

Now a day's natural scene text recognition has gained enormous importance and significant in the field of computer science and information technology. Text recognition in natural images, i.e. called *scene text recognition*, is a challenging but very useful task. Text is one of the basic tools for preserving and communicating information, and a large part of the modern world is designed to be interpreted through the use of labels and other textual cues. The fertile and precise information incorporated in text is very useful in a wide range of applications that are computer-vision based, and hence text detection and recognition in natural scenes (e.g.: traffic sign boards, license plate, Hoardings and videos etc.) have become important and active research topics in computer vision and document analysis. Past researchers clearly show that various aspect of scene text recognition have been discussed in variety of ways.

Computer scientists have always been facing a very important problem of scene text recognition. When applied to natural scene images, traditional character techniques results in failure as they are tuned to the largely black-and-white, line-based environment of printed documents, while text occurring in natural scene images suffers from inconsistent lighting conditions, variable fonts orientations, background noise and imaging distortions. There are generally two stages for recognise scene text: word detection and word recognition. The detection stage generates a large set of word bounding box candidates, and is tuned for speed and high recall. Recognition is therefore a far more challenging problem and it is the focus of this research.

To a certain extent they have succeeded to recognize these data but there were so many things which were effecting and distorting these important data. It is a matter of great pride and pleasure that in the

same sequence we have also studied the scene text recognition in slightly different manner. In fact we have developed an algorithm to elucidate these aspects in the modern technological scenario. We may pointed be claimed that we have thoroughly used neural network algorithms for scene text recognition. As a matter of course these algorithm have for the first time had been example:

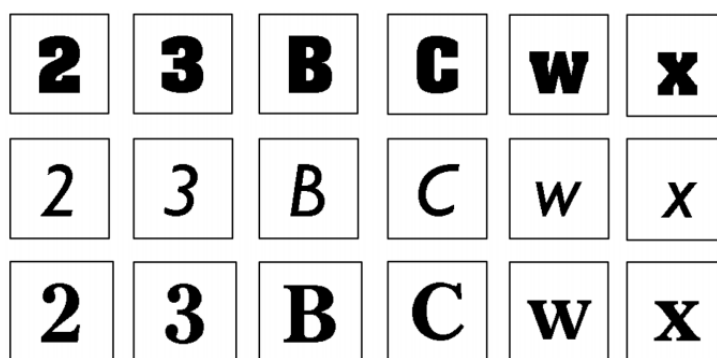
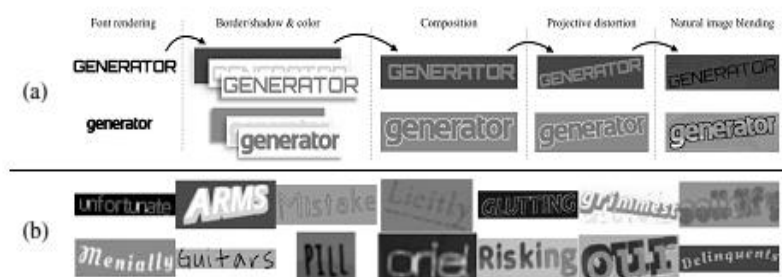


Figure 1 An example of Scene text recognitions

These ideas are basically directly or indirectly related to AI concepts like neural network etc. in fact in biological science these networks are being used to read human mind. These ideas eventually explore the ideas of how human recognize scene text in general

used by the researcher to deal with such problem.

Here it is very contextual to mention that the algorithm established by the researcher is wider and larger in its nature. By virtue of fact that it not only extends but sharpens and refines the previous working system. This may be clarified and justified by means of following

and they are also used to develop machines that simulate this process.

In present work, I describe an alternative approach to this problem of text recognition based upon recent advances in machine learning, and more precisely,

unsupervised feature learning. By leveraging these feature-learning algorithms, we were able to derive a set of specialized features tuned particularly for the text recognition problem. These learned features were then integrated into a larger, discriminatively-trained Self organizing feature map neural network (SOM). SOMs are hierarchical neural networks that have immense representational capacity and have been successfully applied to many problems such as handwriting recognition, visual object recognition and character recognition.

In spite of the simplicity, however, our system will be work on state-of-the art performance on the standard ICDAR 2003, ICDAR 2011, ICDAR 2013 and Street View Text (SVT), VID1 benchmarks. Our results thus demonstrate the viability of our alternative construction of a complete end-to-end text recognition system that does not rely extensively on hand-engineered features or hard-coded prior knowledge.

1.1 Objective

The main objective of proposed work is the assignment of some sort of specific output value to a given basic input value, which could be of different patterns, according to some specific algorithm. Aim of scene text recognition algorithms is to provide a reasonable answer for all possible inputs and to do fuzzy matching of inputs.

- Extraction of text from a scene.
- Performing word recognition on whole image as an extension of character recognition system.
- Make use of synthetic data to train the network for the recognition process.
- Introduce up-to-date works and summarize recent advances,
- Compare different methods and highlight state-of-the-art algorithms.
- Analyse different neural network approaches for scene text

recognition. Develop unsupervised learning algorithm SOM KOHONEN for scene text detection and recognition.

This chapter discusses mainly about natural scene image. Segmentation and recognition of text from the scene image is a challenging task due to blurred, low-resolution and small sized images. Innovative methods have been proposed to address this problem and to recognize the text from natural scene images. The acquired image may be color or gray scale, any file format (jpg, png, bmp, and tiff) and compressed or uncompressed formats. Then the image is converted into a gray scale image if the input image is color. To segment the text from the scene image, two types of segmentation algorithm will be carried out. From the segmented region, extract two types of features. The extracted features are used to recognize the segmented region whether it is a text region or non-text region using the SVM classifier. Template matching algorithm is applied only on the bounding box image to recognize the text. Trained SVM classifier is used to classify the image containing text and non- text region. Performances are analyzed based on the recall rate, precision, accuracy and F-measure.

2 METHODOLOGY

The input image is converted into a gray scale image and segmented to the text from the scene image. From the segmented region, extract two types of features. The extracted features are used to recognize the segmented region whether it is a text region or non-text region using the SVM classifier. Template matching algorithm is applied only on the bounding box image to recognize the text.

Natural scene image. Segmentation and recognition of text from the scene image is a challenging task due to blurred, low-resolution and small sized images. Innovative methods have been proposed to address this

problem and to recognize the text from natural scene images.

The acquired image may be color or gray scale, any file format (jpg, png, bmp, and tiff) and compressed or uncompressed formats. Then the image is converted into a gray scale image if the input image is color. To segment the text from the scene image, two types of segmentation algorithm will be carried out. From the segmented region, extract two types of features.

The extracted features are used to recognize the segmented region whether it is a text region or non-text region using the SVM classifier. Template matching algorithm is applied only on the bounding box image to recognize the text. Trained SVM classifier is used to classify the image containing text and non-text region. Performances are analyzed based on the recall rate, precision, accuracy and F-measure.

2.1 The broad outline of the proposed objective is as follows:

- To develop an algorithm, so that the OCR should be able to recognize highly degraded offline Devanagari text.
- To achieve the above objectives, the following components of the OCR are studied.
- To use some of the existing methods for preprocessing phase such as noise removal,smoothing, document skew detection/correction, thinning, binarization, digitization.
- To explore existing methods of restoration for highly degraded documents especially for the broken characters such as filtering, for restoration, joining broken characters using connected components, filter and active contour model.
- To propose new algorithms for restoration of highly broken text.

- To explore existing features (structural and statistical) and to propose new features for degraded printed Devanagari text.
- To use Support Vector Machine [SVM] classifiers and Neural Network.

2.2 Proposed Research Methodology.

Following is the assumption to carry out the proposed research.

- The scanned documents is in Binarization (digitization) form.
- Printed Natural text is considered.
- Text and non-text separation are primarily defined.
- Proposed new algorithms are tested on 500 data. It is scanned from Natural Image.
- Textual Degradation (Broken characters, Touching characters: Line touching overlapping and Inter - Character touching, Bleed through text) and Non Textual degradation Underline or other lines marked on the document and Back side view.

The work will further be carried out as described below:

- The proposed method is able to restore the highly broken characters.
- The restoration is done in two phases one before segmentation and second after character segmentations.
- The proposed method will use a hybrid approach to handle broken characters and other degradation. Some of the existing methods such as gabor filter and active contour the model along with structural properties is used.
- The proposed method will solve the line overlapping segmentation problem. This method makes use of the projection profiles and CC with some heuristics to segment the

overlapped text documents in a robust way.

- The segmentation phase shall be implemented. Its aim is to separate lines, words and characters of degraded printed text. In this phase, we will review the Run-Length Smoothing Algorithm[RLSA], projection profile, connected components and contour analysis. New algorithm is designed for segmenting the lines, words and characters from degraded printed text using structural features of the language.
- In the feature extraction phase, we will review existing algorithms and techniques like foreground skeleton, background skeleton, stroke detection and elastic matching. We will also review the structural and statistical features for degraded printed devanagari text. Based on these the features shall be detected for degraded printed i text.
- In the classification phase SVM is used.
- Post processing also is implemented.

There are many OCR systems available for handling the clean printed documents in English. But till now there is no system available for recognition of printed degraded natural script. As such, there is a need for degraded OCR for scripts that can help people convert the text to computer processable format. The system developed in the present work can be used for recognition of clean, touched documents of script. There is a need to develop OCR, which can handle broken characters and half characters in natural Image text. .

3 METHODOLOGY

- The input natural image is converted into a grayscale image and segmented to the text from the scene image.

From the segmented region, extract two types of features.

- The extracted features are used to recognize the segmented region whether it is a text region or non-text region using the SVM classifier.
- Template matching algorithm is applied only on the bounding box image to recognize the text.

3.1 Text Recognition

Text can be known by the optical character recognition (OCR) algorithm. OCR is the authorization of printed text or written text characters by a computer. In OCR processing, the scanned-in image or bitmap is analyzed for light and dark areas in order to identify each alphabetic letter or numeric digit. After splitting the text regions, use the OCR function to recognize the text within each bounding box.

The output of the OCR will be noisy when the text region is not found. There are two types of OCR algorithms used here. They are template matching algorithms and SVM classifiers.

3.2 Optical Character Recognition Algorithm

Optical character recognition (OCR) is the process of classification of optical patterns contained in the digital image. Optical Character recognition can be described as mechanical or electronic conversion of scanned images where images can be handwritten, typewritten or printed text. Optical character recognition (OCR) is classifying the optical patterns corresponding to the other characters. The process of OCR involves several steps: segmentation, feature extraction and classification. OCR helps to modify various types of documents such as scanned documents, images.

In OCR can be used to recognize the text in the segmented frames. An OCR included machine recognition techniques performing automatic identification. Automatic identification is the process of a

recognition system that identifies the object, automatically collects the data and enters data directly into computer systems.

In OCR form include some special symbols such as letters, numbers, commas, question marks as well as different characters. A typical OCR system has many components as: The first step is to digitize analog documents using an optical scanner.

When regions containing text are located each symbol is extracted through a segmentation process. The extracted symbols are pre-processed and eliminate the noise to facilitate feature extraction. Each symbol is found by comparing extracted features with descriptions of symbols. Finally contextual information is used to reconstruct words and numbers of the original text.

3.3 Flowchart

A OCR system is improved from primitive schemes suitable only for reading stylized printed numerals to more complex and sophisticated techniques for the recognition and printed characters. The character recognition continues to appear with development of computer technology and decrease in computational restrictions. Some techniques recognize the whole word instead of individual characters A typical OCR system consists of steps as:

- To scan the textual image.
- To convert the color image into gray image and then binary image.
- When the gray images are preprocessed like noise removal and skew correction etc.
- To segmented by the textual image from separating

3.4 Template Matching Algorithm

Template matching is a system model which compares the pair of images to recognize the character or alphabet. It found the location of the sub image which is called a template. Centre of the template is taken as a

corresponding point to calculate the registration parameters.

Template matching is a system model which compares the pair of images to recognize the character or alphabet. It found the location of the sub image which is called a template. Centre of the template is taken as a corresponding point to calculate the registration parameters. When the image is compared to the template image. To Detect the techniques for the highest matching area. Template matching is one of the areas of profound interests.

Template matching provides a new dimension into the image processing capabilities, although there are many attempts to resolve different issues in the field. Template matching is a high-level machine vision technique that is allowed to identify the image pattern. Template matching algorithm is allowed to find the pattern occurrences regardless of the orientation and local brightness.

Template matching technique can be applied only for template images. It is used to find the pixel level and match the character boundaries in the template image. Template matching techniques are used for image processing, which is used to find the location of the scene image. Template matching is classified into two approaches: Feature-based approaches and Area- based approaches.

A feature-based approach is used to find the relationship between the scene image and template image using the feature descriptors. Feature-based approaches include curves, points and surface. Area-based approaches for the templates have no strong feature with image. Matches are based on the intensity values of both input image and template image.

The Template Matching algorithm implements the following steps:

- Firstly, the character image from the detected string is selected.

- After that, the image to the size of the first template is rescaled.
- After rescale the image to the size of the original image, the matching metric is computed.
- Then the highest match found is stored. If the image does not match repeat again the third step.
- Good competition code is stored as the recognized character.

3.5 Advantages of SVM Classifier

- SVMs are effective when the number of features is quite large.
- It works effectively even if the number of features is greater than the number of samples.
- Non-Linear data can also be classified using customized hyper planes built by using kernel trick.
- It is a robust model to solve prediction problems since it maximizes margin.

Next this research work explained about the various types of classifiers for comparison. The different types of classifiers are KNN, DNN and ANN. The proposed SVM classifier is compared with those classifiers. These outputs are explained in this chapter. The input image is pre-processed, segmented and features extracted using the same method explained in chapter 3, chapter 4 and chapter 5. Then those features are given to the different type of classifiers and these are also explained in this section (Andrew Zisserman et al. 2015).

3.6 KNN (k Nearest Neighbour)

KNN (k Nearest Neighbour) denotes the number of nearest neighbors that are helpful

to predict the class of the test sample (102). It is a nonparametric and learning algorithm because it does not build the classifier at training time; rather it delays until classification time. KNN uses standard Euclidean distance to measure the variation between the training instance and test instance.

One of the major difficulties faced when the KNN classifier is employed is that each of the sample data is given equal importance when the class label is to be assigned to the input data. This causes problems where the samples overlap. Another difficulty is that once an input sample is assigned a class label, there is no information about its strength of membership in that class.

3.7 ANN (Artificial Neural Network)

An artificial neural network is an information processing system which has certain performance characteristics similar to biological neural networks. Neural network consists of a large number of simple processing elements called neurons. Each neuron connects to other neurons by means of directed communication links, each with an associated weight. The weights represent information used by the network to solve a problem. Figure 2 shows that, there are 9 features, the classifier network consists of 9 inputs. Number of hidden neurons taken is 4 and one output neuron. used to minimize the number of input features. Finally the input feature taken for this classifier is 9. The activation function used is tan sigmoid function. The output of the network is 0 or 1. Zero indicates a text condition and one indicates non-text condition.

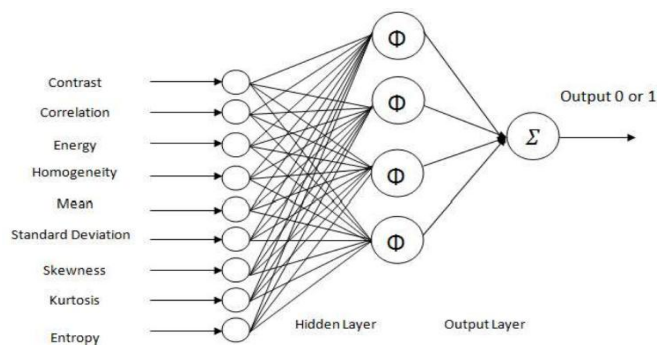


Figure 2 9 Features, the classifier network consists of 9 inputs.

3.8 DNN (Deep Neural Networks)

In a deep neural network (DNN), several neurons are stacked in layers and are connected such that the outputs from the previous layer are inputs to the next layer. The first layer in a neural network is called the input layer, the final layer the output layer, and the intermediate layers the hidden layers (90). A neural network is called deep when it consists of one or more hidden layers and a fully connected neural network is when each neuron in the layers is connected. Figure 3 is a simple example of a fully connected neural network, but is also a DNN that consists of two input nodes, a hidden layer with three neurons and two output classes.

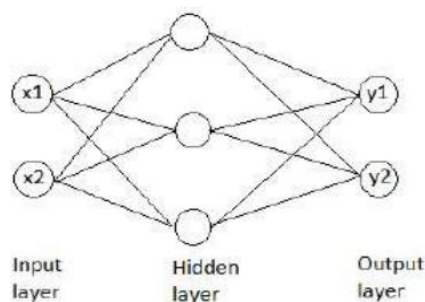


Figure 3

Table 1 Performance values of different classifier

Classifier	KNN	ANN	DNN	SVM(Proposed)
Sensitivity	89	90	92	95
Specificity	91	93	94	98
Precision	84	87	90	100
Accuracy	87	89	92	95
Recall	82	84	86	90
F-measure	87.2	89	90	94.73

3.9 Optical Character Recognition Through Neural Network

Optical Character Recognition deals in recognition and classification of characters from an image.

For accurate recognition, a character is classified and recognized based on calculated topological and geometrical properties.

3.10 Character properties'

- Features are extracted from the image by means of spatial pixel-based calculation.
- A collection of such features help in defining a character uniquely.

Preprocessing



Figure 4 Preprocessing

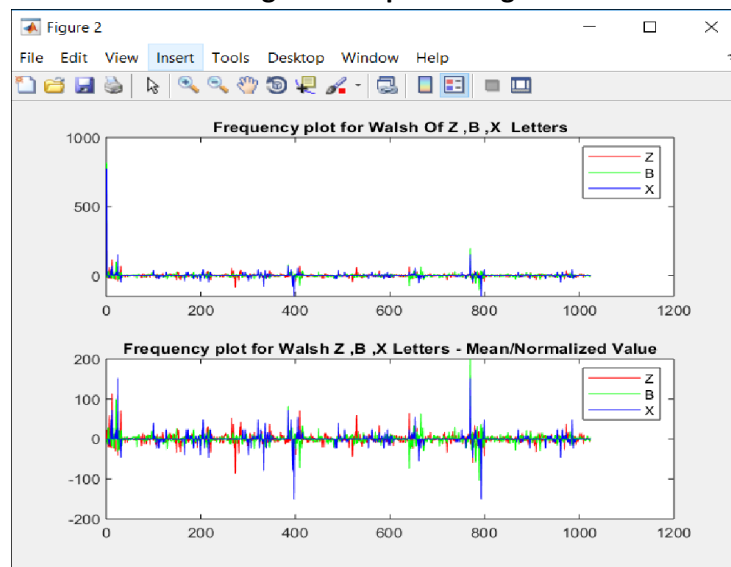


Figure 5 Walsh Comparison

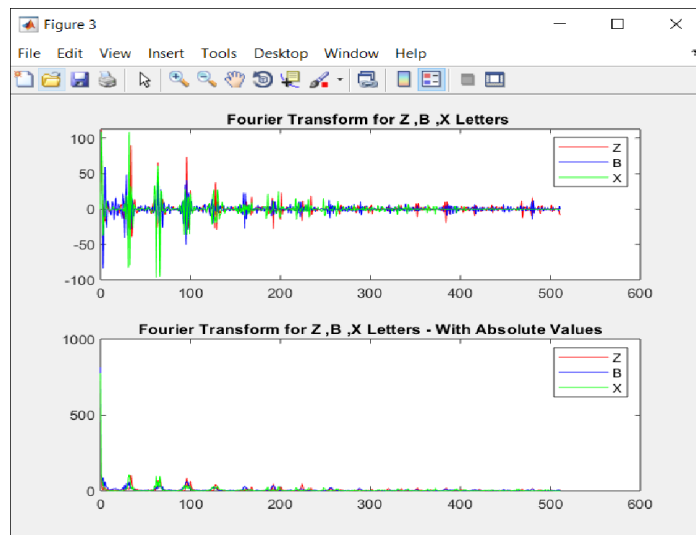


Figure 6 Fourier Comparison

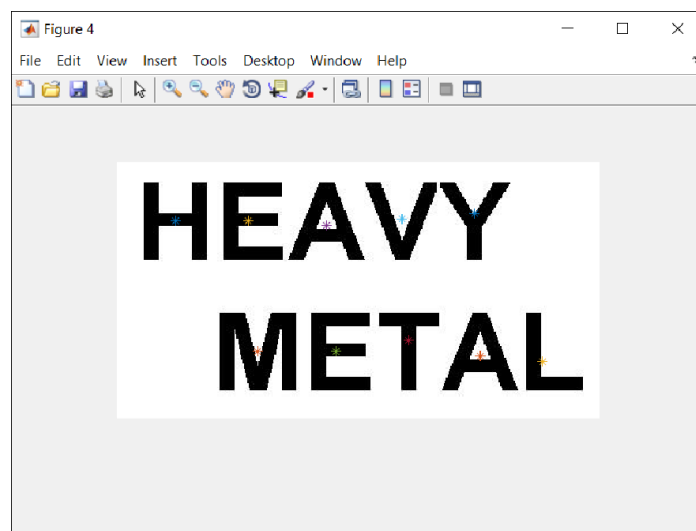


Figure 7 Centroid

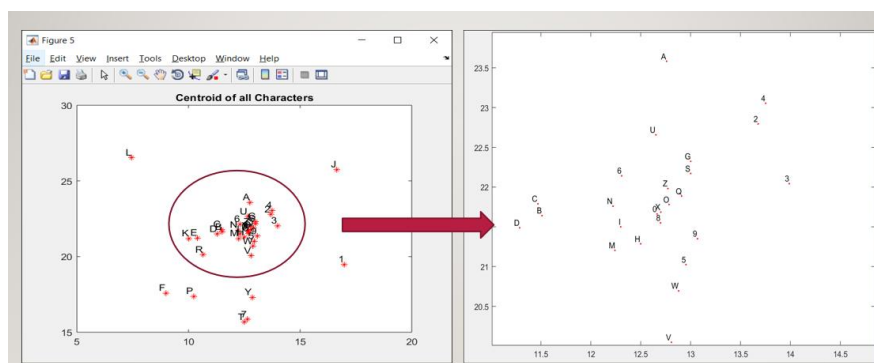


Figure 8 Centroid and Center of Mass

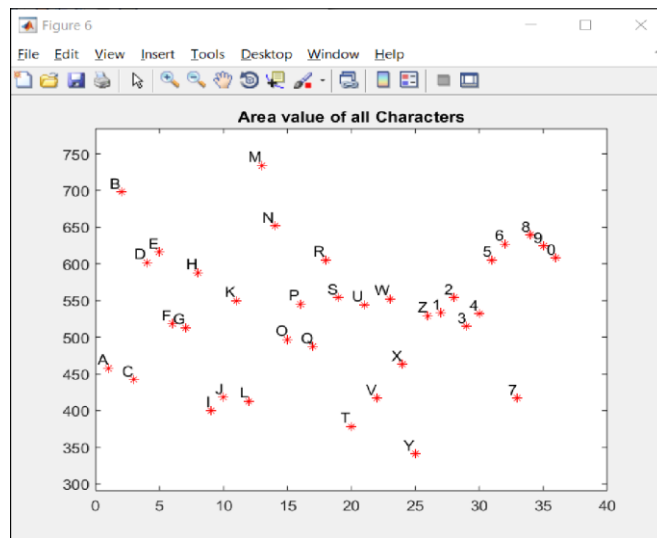
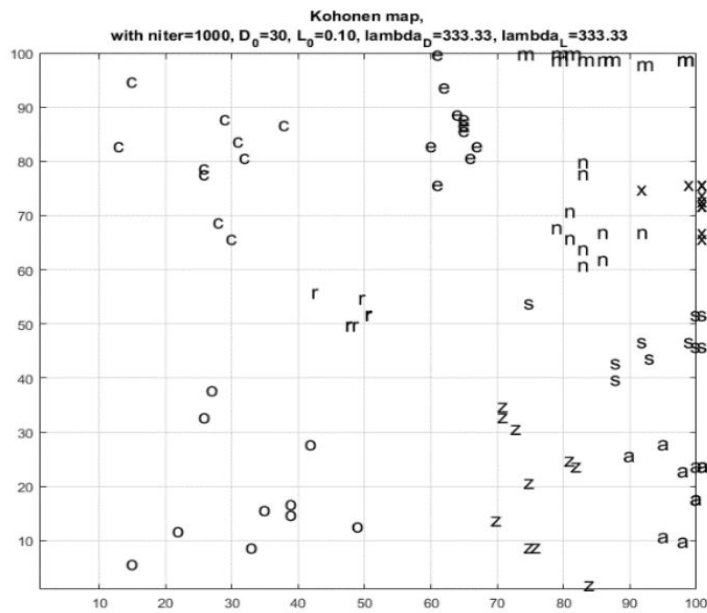


Figure 9

From figure 10 we can see that L_0 will affect the density of each cluster, i.e. the average distance between elements in each cluster. Smaller D_0 will lead to higher density of the clusters, making the same characters more condensed.



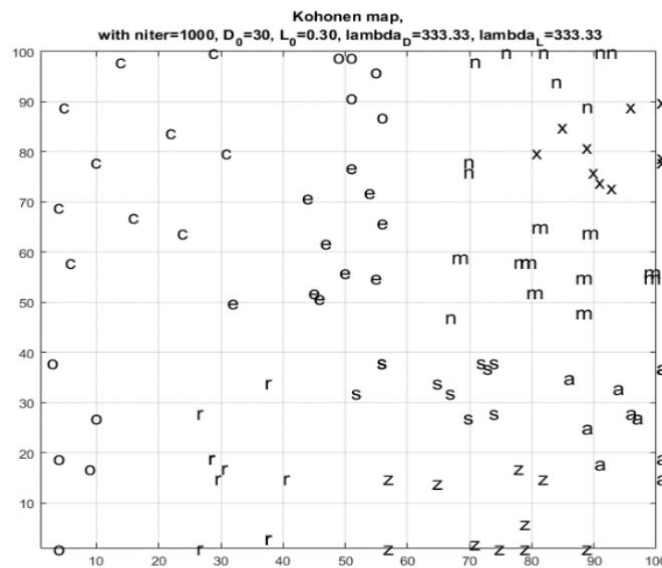


Figure 10 Kohonen map

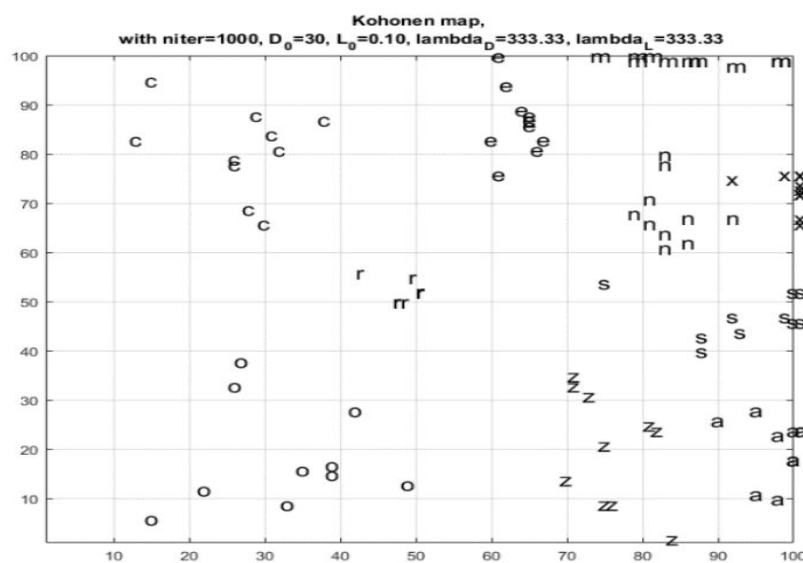
4 CLASSIFICATION RESULTS

According to the optimized parameters obtained in last subsection, we fix the parameters for the final classification as follows:

Table 2 Optimal parameters for classification

N	1000
s	100
D_0	30
L_0	0.1

In figure 11, the samples in training and test set are located in the fitted Kohonen map according to the best matching principle stated in the previous section.



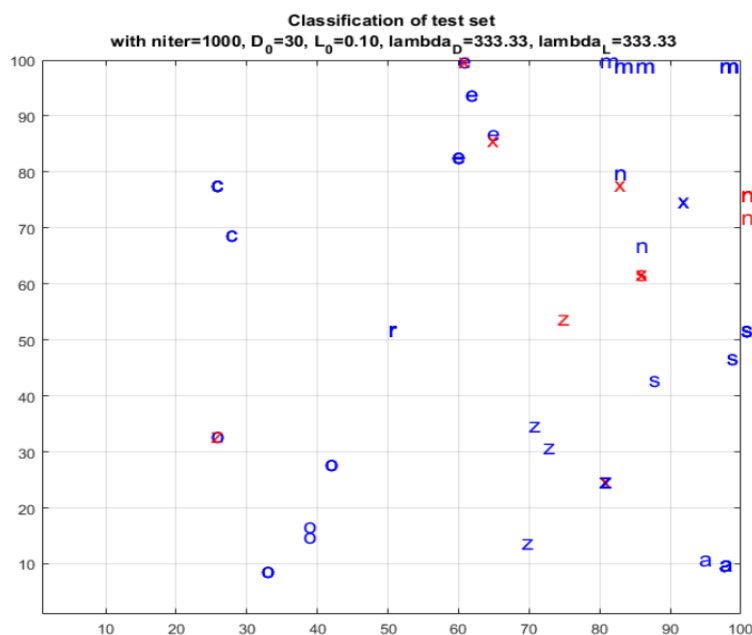


Figure 11 Kononen map with location of training samples and classification results for test samples

We can see from figure 11 that the same characters are located closed to each other in the map, which indicates that the sample in the same class are well cluster by their location on the map. The whole map can be segmented into 10 sections corresponding to 10 different classes. If a sample in the test set is located in a certain section, we will determine its label as the label of this section.

Table 3 Prediction results of test set

	a	c	e	m	n	o	r	s	x	z	accuracy
a	10										1.0
c		10									1.0
e			10								1.0
m				10							1.0
n					3				7		0.3
o						10					1.0
r							10				1.0
s					1			9			1.9
x			2		2				2	4	0.2
z					1	1				8	0.8
Average											0.82

The prediction results of the test set are shown in table 3 where we can see that most prediction errors occur for the character 'x' and 'n'. If we verify the image for x in the data set, we will find the character x in the training set is seriously distorted, which makes it hard to predicted for the test set.

5 CONCLUSION

Kohonen algorithm provides a practical tool for visualizing the high dimensional data. If all types of data can be viewed in 2D or 3D space, it will be more intuitional for us to observe the characteristic, the similarity and the diversion of the data, and then classify them.

In our work this approach is applied to the optical character recognition, and we have obtained 82% accuracy for 10 classes of scanned characters. Further work including comparing the performance of different classifiers, implementing for handwritten characters recognition will be done.

REFERENCES

1. Abdelkarim Ben Ayed, Mohamed Ben Halima and Adel M. Alimi, "MapReduce Based Text Detection in Big Data Natural Scene Videos", *Procedia Computer Science*, Volume 53, 2015, Pages 216–223.
2. Baoguang Shi, Xiang Bai and Cong Yao, "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition", arXiv:1507.05717v1 [cs.CV] 21 Jul 2015.
3. Tong He, Weilin Huang, "Text-Attentional Convolutional Neural Network for Scene Text Detection", arXiv:1510.03283v2 [cs.CV] 24 Mar 2016.
4. Cong Yao, Jianan Wu, Xinyu Zhou, Chi Zhang, Shuchang Zhou, Zhimin Cao, Qi Yin Megvii Inc., "Incidental Scene Text Understanding: Recent Progresses on

- ICDAR 2015 Robust Reading Competition Challenge 4", arXiv:1511.09207v2 [cs.CV] 3 Feb 2016.
5. Andreas Veit, Tomá's Madera, Luk'a's Neumann, Ji'r'i Matas, Serge Belongie, "COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images", arXiv:1601.07140v2 [cs.CV] 19 Jun 2016.
6. Llu'is G'omez and Dimosthenis Karatzas, "A fine-grained approach to scene text script identification", arXiv:1602.07475v1 [cs.CV] 24 Feb 2016.
7. Zhuoyao Zhong, Lianwen Jin, Shuye Zhang, Ziyong Feng, "Deep Text: A Unified Framework for Text Proposal Generation and Text Detection in Natural Images", arXiv:1605.07314v1[cs.CV] 24 May 2016.
8. Cong Yao, Xiang Bai¹, Nong Sang, Xinyu Zhou, Shuchang Zhou, Zhimin Cao, "Scene Text Detection via Holistic, Multi-Channel Prediction", arXiv:1606.09002v2 [cs.CV] 5 Jul 2016.
9. Jianqi Ma, Weiyuan Shao, Hao Ye, Li Wang, Hong Wang, Yingbin Zheng, Xiangyang Xue, "Arbitrary-Oriented Scene Text Detection via Rotation Proposals", arXiv:1703.01086v3 [cs.CV] 15 Mar 2018.
10. Yingying Jiang, Xiangyu Zhu, Xiaobing Wang, Shuli Yang, Wei Li, Hua Wang, Pei Fu and Zhenbo Luo, "R²CNN: Rotational Region CNN for Orientation Robust Scene Text Detection", arXiv:1706.09579 v2 [cs.CV] 30 June 2017.
11. Christian Bartz Haojin Yang Christoph Meinel, "STN-OCR: A single Neural Network for Text Detection and Text Recognition", arXiv:1707.08831v1 [cs.CV] 27 Jul 2017.
12. Minghui Liao, Baoguang Shi, Xiang Bai, "TextBoxes++: A Single-Shot Oriented

- Scene Text Detector”, arXiv:1801.02765v3 [cs.CV] 27 Apr 2018.
13. Zichuan Liu, Guosheng Lin, Sheng Yang, Jiashi Feng, Weisi Lin and Wang Ling Goh, “Learning Markov Clustering Networks for Scene Text Detection”, arXiv:1805.08365v1 [cs.CV] 22 May 2018.
 14. Zhida Huang, Zhuoyao Zhong, Lei Sun, Qiang Huo, “Mask R-CNN with Pyramid Attention Network for Scene Text Detection”, ICDARarXiv:1811.09058v1 [cs.CV] 22 Nov 2018.
 15. Yixing Zhu, Jun Du, “TextMountain: Accurate Scene Text Detection via Instance Segmentation”, arXiv:1811.12786v1 [cs.CV] 30 Nov 2018.
 16. Yongchao Xu, Yukang Wang, Wei Zhou, Yongpan Wang, Zhibo Yang, Xiang Bai, “TextField: Learning A Deep Direction Field for Irregular Scene Text Detection”, arXiv:1812.01393v2 [cs.CV] 29 Jul 2019.
 17. Ning Lu, Wenwen Yu, Xianbiao Qi, Yihao Chen, Ping Gong, Rong Xiao, Xiang Bai, “MASTER: Multi-Aspect Non-local Network for Scene Text Recognition?”, arXiv:1910.02562v3 [cs.CV] 11 Apr 2021.
 18. Yuliang Liu, Tong He, Hao Chen, Xinyu Wang, Canjie Luo, Shuaitao Zhang, Chunhua Shen, Lianwen Jin, “Exploring the Capacity of an Orderless Box Discretization Network for Multi-orientation Scene Text Detection”, arXiv:1912.09629v3 [cs.CV] 16 Jan 2021.
 19. Chixiang Maa, Lei Sun, Zhuoyao Zhong, Qiang Huo, “ReLaText: Exploiting Visual Relationships for Arbitrary-Shaped Scene Text Detection with Graph Convolutional Networks”, arXiv:2003.06999v1 [cs.CV] 16 Mar 2020.
 20. Wenqing Zhang, Yang Qiu, Minghui Liao, Rui Zhang, Xiaolin Wei, Xiang Bai, “Scene Text Detection with Scribble Lines”, arXiv:2012.05030v2 [cs.CV] 10 Dec 2020.
 21. Youjiang Xu, Jiaqi Duan, Zhanghui Kuang, Xiaoyu Yue, Hongbin Sun, Yue Guan, Wayne Zhang, “Geometry Normalization Networks for Accurate Scene Text Detection”, IEEE Xplore.
 22. Manuel A. Córdova, Luis G. L. Decker, Jose L. Flores-Campana, Andreza A. dos Santos, Jhonatas S. Conceic, ao Allan Pinto, Helio Pedrini and Ricardo da S. Torres, “Pelee-Text: A Tiny Convolutional Neural Network for Multi-Oriented Scene Text Detection”, IEEE 2019.
 23. Youhui Guo, Yu Zhou, Xugong Qin, and Weiping Wang, “Which and Where to Focus: A Simple yet Accurate Framework for Arbitrary-Shaped Nearby Text Detection in Scene Images?”, arXiv:2109.03451v1 [cs.CV] 8 Sep 2021.
 24. Enze Xie, Yuhang Zang, Shuai Shao, Gang Yu, Cong Yao, Guangyao Li, “Scene Text Detection with Supervised Pyramid Context Network”, Association for the Advancement of Artificial Intelligence, 2019.
 25. Minghui Liao, Baoguang Shi, Xiang Bai, Xinggang Wang, Wenyu Liu, “TextBoxes: A Fast Text Detector with a Single Deep Neural Network”, Association for the Advancement of Artificial Intelligence, 2017.
 26. Han Lin, Peng Yang, Fanlong Zhang, “Review of Scene Text Detection and Recognition”, Archives of Computational Methods in Engineering, January 2019.
 27. Youngmin Baek, Bado Lee, Dongyoon Han, Sangdoo Yun, and Hwalsuk Lee, “Character Region Awareness for Text Detection”, IEEE 2019.

28. Saif Hassan Katper, Abdul Rehman Gilal, Ahmad Waqas, Aeshah Alsughayyir, Abdullah Alshantiti, Jafreezal Jaafar, "Deep Neural Networks Combined with STN for Multi-Oriented Text Detection and Recognition", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 4, 2020.
29. Fangneng Zhan, Shijian Lu and Chuhui Xue, "Verisimilar Image Synthesis for Accurate Detection and Recognition of Texts in Scenes", CVF, 2018.
30. Yongcun Cao, Shuaisen Ma, and Haichuan Pan, "FDTA: Fully Convolutional Scene Text Detection with Text Attention", IEEE Access, September 4, 2020.
31. Wenhao He, Xu-Yao Zhang, Fei Yin, Cheng-Lin Liu, "Deep Direct Regression for Multi-Oriented Scene Text Detection", IEEE Xplore, 2017.
32. Muhammad Fraz, M. Saquib Sarfraz, "Exploiting colour information for better scene text detection and recognition", Springer, Jan 2015.
33. Minghui Liao, Zhen Zhu, Baoguang Shi, Gui-song Xia, Xiang Bai, "Rotation-sensitive Regression for Oriented Scene Text Detection", IEEE Xplore, 2018.
34. Xuebo Liu, Ding Liang, Shi Yan, Dagui Chen, Yu Qiao, and Junjie Yan, "FOTS: Fast Oriented Text Spotting with a Unified Network", IEEE Xplore, 2018.
35. Pengyuan Lyu, Cong Yao, Wenhao Wu, Shuicheng Yan, Xiang Bai, "Multi-Oriented Scene Text Detection via Corner Localization and Region Segmentation", IEEE Xplore, 2018.
36. Tao Sheng, Jie Chen, Zhouhui Lian, "CentripetalText: An Efficient Text Instance Representation for Scene Text Detection", Neural Information Processing Systems (NeurIPS 2021).
37. Nguyen Nguyen, Thu Nguyen, Vinh Tran, Minh-Triet Tran, Thanh Duc Ngo, Thien Huu Nguyen, Minh Hoai, "Dictionary-guided Scene Text Recognition, IEEE Xplore, 2021.
38. Manuel Córdova, Allan Pinto, Helio Pedrini, and Ricardo Da Silva Torres, "Pelee-TextCC: A Tiny Neural Network for Scene Text Detection", IEEE Access, December 2020.
39. Pengyuan Lyu, Minghui Liao, Cong Yao, Wenhao Wu and Xiang Bai, "Mask TextSpotter: An End-to-End Trainable Neural Network for Spotting Text with Arbitrary Shapes", IEEE Xplore, 2018.
40. Mukhriddin Mukhiddinov, "Scene Text Detection and Localization using Fully Convolutional Network", IEEE Xplore 2019.
41. Shuangshuang Li and Wenming Cao, "SEMPANet: A Modified Path Aggregation Network with Squeeze-Excitation for Scene Text Detection", Sensors 2021, 21, 2657. <https://doi.org/10.3390/s21082657>.
42. Baoguang Shi, Xinggang Wang, Pengyuan Lyu, Cong Yao, Xiang Bai, "Robust Scene Text Recognition with Automatic Rectification", IEEE Xplore 2016
43. Muhammad A.Panhwar, Kamran A. Memon, Adeel Abro, Deng Zhongliang, Sijjad A. Khuhro, Saleemullah Memon, "Signboard Detection and Text Recognition Using Artificial Neural Networks", IEEE 2019.
44. Shangxuan Tian, Yifeng Pan, Chang Huang, Shijian Lu, Kai Yu, and Chew Lim Tan, "Text Flow: A Unified Text Detection System in Natural Scene Images", IEEE Xplore 2015.
45. Xiaobing Wang, Yingying Jiang, Zhenbo Luo, Cheng-Lin Liu, Hyunsoo Choi, Sungjin Kim, "Arbitrary Shape Scene Text Detection with Adaptive Text Region Representation", IEEE 2019.
46. Fangfang Wang, Liming Zhao, Xi Li, Xinchao Wang, Dacheng Tao, "Geometry-Aware Scene Text Detection

- with Instance Transformation Network”, IEEE Xplore 2018.
47. Hao Wang, Xiang Bai, Mingkun Yang, Shenggao Zhu, Jing Wang, Wenyu Liu, “Scene Text Retrieval via Joint Text Detection and Similarity Learning”, IEEE Xplore 2021.
48. Fangneng Zhan, Chuhui Xue, and Shijian Lu, “GA-DAN: Geometry-Aware Domain Adaptation Network for Scene Text Detection and Recognition”, IEEE Xplore 2019.
49. Zheng Zhang, Wei Shen, Cong Yao, Xiang Bai, “Symmetry-Based Text Line Detection in Natural Scenes”, IEEE Xplore 2015.
50. Dongping Cao, Yong Zhong, Lishun Wang, Yilong He and Jiachen Dang, “Scene Text Detection in Natural Images: A Review”, *Symmetry* 2020, 12, 1956; doi:10.3390/sym12121956.
51. Yuliang Liu, Lianwen Jin, Shuaitao Zhang, Canjie Luo, Sheng Zhang, “Curved scene text detection via transverse and longitudinal sequence connection”, *Pattern Recognition* 90 (2019) 337–345.
52. Thananop Kobchaisawat and Thanarat H. Chalidabhongse, “Thai Text Localization in Natural Scene Images using Convolutional Neural Network”, APSIPA 2014.
53. Max Jaderberg, Karen Simonyan, Andrea Vedaldi, Andrew Zisserman, “Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition”, arXiv:1406.2227v4 [cs.CV] 9, Dec 2014.
54. Xiaohang Ren, Kai Chen, Jun Sun, “A Novel Scene Text Detection Algorithm Based On Convolutional Neural Network”, arXiv:1604.01894v1 [cs.CV], 7 Apr 2016.
55. Yuchen Dai, Zheng Huang, Yuting Gao, Youxuan Xuz, Kai Che, Jie Guo and Weidong Qiu, “Fused Text Segmentation Networks for Multi-oriented Scene Text Detection”, arXiv:1709.03272v4 [cs.CV] 7 May 2018.
56. Chuhui Xue, Shijian Lu1 and Wei Zhang, “MSR: Multi-Scale Shape Regression for Scene Text Detection”, arXiv:1901.02596v2 [cs.CV] 9 Jul 2019.
57. Nibal Nayef, Yash Patel, Michal Busta, Pinaki Nath Chowdhury, Dimosthenis Karatza, “ICDAR2019 Robust Reading Challenge on Multi-lingual Scene Text Detection and Recognition – RRC-MLT-2019”, arXiv:1907.00945v1 [cs.CV] 1 Jul 2019.
58. Fedor Borisyuk, Albert Gordo, Viswanath Sivakumar, “Rosetta: Large scale system for text detection and recognition in images”, arXiv:1910.05085v1 [cs.CV] 11 Oct 2019.
59. Minghui Liao, Jian Zhang, Zhaoyi Wan, Fengming Xie, Jiajun Liang, Pengyuan Lyu, Cong Yao, Xiang Bai, “Scene Text Recognition from Two-Dimensional Perspective”, *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)* 2019.
60. Lei Sun, Qiang Huo, Wei Jia, Kai Chen, “Robust Text Detection in Natural Scene Images by Generalized Color enhanced Contrasting Extremal Region and Neural Networks”, *International Conference on Pattern Recognition*, 2014.
61. Minghui Liao, Zhaoyi Wan, Cong Yao, Kai Chen, Xiang Bai, “Real-Time Scene Text Detection with Differentiable Binarization”, *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)*, 2020.
62. Jie Chen, Zhouhui Lian, Yizhi Wang, Yingmin Tang & Jianguo Xiao, “Irregular scene text detection via attention guided border labeling”, *Science China Press and Springer-Verlag GmbH Germany*, part of Springer Nature, 2019.

63. Xiaohang Ren, Yi Zhou, Zheng Huang, Jun Sun, Xiaokang Yang and Kai Chen, "A Novel Text Structure Feature Extractor for Chinese Scene Text Detection and Recognition", IEEE Access, VOLUME 5, 2017.
64. Dafang He, Xiao Yang, Wenyi Huang, Zihan Zhou, Daniel Kifer and C.Lee Giles, "Aggregating Local Context for Accurate Scene Text Detection", 2016.
65. Hongbo Qin, Haodi Zhang, Hai Wang, Yujin Yan, Min Zhang and Wei Zhao, "An Algorithm for Scene Text Detection Using Multibox and Semantic Segmentation", Appl. Sci. 2019, 9, 1054; doi:10.3390/app9061054.
66. Xiao Qin, Jianhui Jiang, Chang-An Yuan, Shaojie Qiao and Wei Fan, "Arbitrary Shape Natural Scene Text Detection Method Based on Soft Attention Mechanism and Dilated Convolution", IEEE Access, Volume 8, 2020.
67. Hojin Cho, Myungchul Sung, Bongjin Jun, "Canny Text Detector: Fast and Robust Scene Text Localization Algorithm", IEEE Xplore, 2016.
68. Chuhui Xue, Shijian Lu and Fangneng Zhan, "Accurate Scene Text Detection through Border Semantics Awareness and Bootstrapping", ECCV 2018.
69. Xiaohang Ren, Yi Zhou, Jianhua He, "A Convolutional Neural Network Based Chinese Text Detection Algorithm via Text Structure Modeling", IEEE 2016.
70. Fagui Liu, Cheng Chen, Dian Gu, And Jingzhong Zheng, "FTPN: Scene Text Detection With Feature Pyramid Based Text Proposal Network", IEEE Access, Volume 7, 2019.
71. Han Hu, Chengquan Zhang, Yuxuan Luo, Yuzhuo Wang, Junyu Han, Errui Ding, "WordSup: Exploiting Word Annotations for Character based Text Detection", IEEE Xplore, 2017.
72. Weilin Huang, Yu Qiao and Xiaoou Tang, "Robust Scene Text Detection with Convolution Neural Network Induced MSER Trees", ECCV 2014, Part IV, LNCS 8692, pp. 497–511, 2014.
73. Sunil Kumar, Krishan Kumar, Rahul Kumar Mishra, "Scene Text Recognition using Artificial Neural Network: A Survey", International Journal of Computer Applications (0975 – 8887), Volume 137 – No.6, March 2016.
74. Yuliang Liu, Hao Chen, Chunhua Shen, Tong He, Lianwen Jin, Liangwei Wang, "ABCNet: Real-time Scene Text Spotting with Adaptive Bezier-Curve Network", IEEE Xplore, 2020.
75. Yuliang Liu, Lianwen Jin, "Deep Matching Prior Network: Toward Tighter Multi-oriented Text Detection", IEEE Xplore, 2017.
76. Anna Zhua, Renwu Gaob, Seiichi Uchida, "Could Scene Context be Beneficial for Scene Text Detection?", Preprint submitted to Journal of Pattern Recognition, March 24, 2016.
77. Yingying ZHU, Cong YAO, Xiang BAI, "Scene text detection and recognition: recent advances and future trends", Front. Comput. Sci., 2016, 10(1): 19–36, DOI 10.1007/s11704-015-4488-0.
78. Jiahao Shi, Long Chen, and Feng Su, "Accurate Arbitrary-Shaped Scene Text Detection via Iterative Polynomial Parameter Regression", IEEE Xplore, 2020.
79. Jieming Wang, Wei Wu, "Research on Natural Scene Text Detection and Recognition", Journal of Physics: Conference Series 1754 (2021) 012200, doi:10.1088/1742-6596/1754/1/012200.
80. Ines Beltaief, Mohamed Ben Halima, "Deep FCN for Arabic Scene Text Detection", IEEE 2018.