EFFECTIVE SECURE TEXT EXTRACTION FOR IMAGE DATASET USING MULTI COLOR SEGMENTATION WITH STRUCTURE ANALYSIS MODEL

Student Details

Santhi. k,
Student,
Department of computer science,
Navarasam arts and science college for women,
Arachalur, Erode, Tamilnadu, India.

Guide Details

Dr. P. Rathiga
HOD, Department of computer Applications,
Navarasam arts and Science College for women,
Arachalur, Erode, Tamil Nadu – 638101.

Abstract: Image-based mostly text data are effective tags or clues for several security applications. In natural pictures, text characters and strings typically seem in close signboards and supply significant information of close settings and objects. Text characters and strings in a natural scene will offer valuable data for several applications. Extracting text directly from natural scene pictures may be a difficult task owing to various text patterns and variant background interferences. This paper proposes a way of text extraction exploitation scene image. The planned paper styles a discriminative character descriptor by combining many progressive feature detectors and descriptors. Second, it models character structure at every character category by coming up with stroke configuration maps. the look is compatible with the appliance of scene text extraction in pictures. The paper is developed to point out the effectiveness of our planned text extraction methodology on image dataset with text data extraction from close objects

Keywords— Text Detection, Scene Images Dataset, Text Extraction, Segmentation, Color Model.

I. INTRODUCTION

Image process systems are getting common because of simple availableness of powerful personnel computers, massive size memory devices, graphics softwares etc. Most of the techniques are developed for enhancing pictures obtained from remote-controlled spacecrafts, area probes and military intelligence activity flights. Image Dataset analysis is bothered with creating quantitative measurements from a picture to supply an outline. within the simplest type, this task can be reading a label on a grocery item, sorting totally different components on an production line, or activity the scale and orientation of blood cells in an exceedingly medical image. a lot of advanced image analysis systems live quantitative data and use it to create a complicated call, like dominant the arm of a automaton to maneuver an object once characteristic it or navigating an craft with the help of pictures nonheritable on its flight.

Image analysis techniques need extraction of bound options that aid within the identification of the item. Segmentation techniques are accustomed to isolate the specified object from the scene so that measurements may be created on that after. Quantitative measurements of object options permit classification and outline of the image.

Image segmentation is the method that subdivides a picture into its constituent components or objects. the extent to that this subdivision is disbursed depends on the matter being resolved, i.e., the segmentation ought to stop once the objects of interest in an application are isolated e.g., in autonomous air-to-surface target acquisition, suppose our interest lies in distinctive vehicles on a road, the primary step is to phase the road from the image so to phase the contents of the road all the way down to potential vehicles. Image thresholding techniques are used for image segmentation.

Classification is that the labeling of a component or a gaggle of pixels supported its gray worth. Classification is one of the foremost typically used ways of data extraction. In Classification, sometimes multiple options square measure used for a group of pixels i.e., several pictures of a specific object are required. In Remote Sensing space, this procedure assumes that the mental imagery of a selected {geographic square measurea|geographical area|geographical region|geographic region|region} is collected in multiple regions of the spectrum which the photographs are in smart registration. Most of the data extraction techniques consider an analysis of the spectral reflection factor properties of such mental imagery and use special algorithms designed to perform varied forms of 'spectral analysis'. the method of multispectral classification may be performed mistreatment either of the two methods: supervised or unsupervised.

In supervised classification, the identity and placement of a number of the land cowl sorts like an urban, wetland, forest, etc., are called Apriori through a mix of field works and toposheets. The analyst tries to find specific sites within the remotely perceived knowledge that represents homogenous samples of these land cowl sorts. These square measures are normally referred to as "Training Sites" as a result of the spectral characteristics of those far-famed areas that are wont to 'train' the classification rule for ultimate land cowl mapping of the reminder of the image.

Multivariate applied mathematics parameters are calculated for every coaching website. each component each inside and out of doors these coaching sites is then evaluated and appointed to a category of that it's the very best chance of being a member.

In unsupervised classification, the identities of land cowl sorts must be nominal as categories inside a scene are not usually called Apriori as a result of ground truth is lacking or surface options inside the scene don't seem to be well outlined. the pc is needed to group component knowledge into totally different spectral categories in keeping with some statistically determined criteria. The comparison in medical space is that the labeling of cells supported their form, size, color and texture, that act as options. This technique is additionally helpful for magnetic resonance imaging pictures.

The natural scene pictures dataset, text characters, and strings typically seem in close signboards and hand-held objects and supply vital data of close atmosphere and objects. Text-based tags are way more applicable than barcode or fast response code as a result of the latter techniques contain restricted data and need pre-installed marks. To extract text data by mobile devices from the natural scene, automatic and economical scene text detection and recognition algorithms are essential. However, extracting scene text could be a difficult task thanks to two main factors are littered backgrounds with noise and non-text outliers, and various text patterns like character sorts, fonts, and sizes.

The frequency of incidence of text in a natural scene is incredibly low, and a restricted range of text characters square measure embedded into advanced non-text background outliers. Background textures, like the grid, window, and brick, even gibe text characters, and strings. though these difficult factors exist in face and automobile, several progressive algorithms have incontestable effectiveness on those applications, as a result of face and automobile, have comparatively stable options.

In this planned system to unravel these difficult issues, scene text extraction is split into two processes text detection and text recognition. It aims to get rid of most non-text background outliers. Text recognition is to rework pixel-based text into clear code. It aims to accurately distinguish completely different text characters and properly compose text word. the most objective of this paper is,

- To extract representative and discriminative options from character patches.
- It combines many feature detectors (Harris-Corner, highest Stable Extremal Regions (MSER), and dense sampling) and histogram of orientated Gradients (HOG) descriptors.
- To generate a binary categoryifier for every character class in text retrieval.
- To novel stroke configuration from character boundary and skeleton to model character structure.

II. RELATED WORKS

Agrawal Hun Kiernan et al., describe the piracy of digital assets like software package, images, video, audio, and text has long been a priority for homeowners of those assets. Protection of those assets is typically based mostly upon the insertion of digital watermarks into the info The watermarking software package introduces little errors into the article being watermarked. These intentional errors are known as marks and every one the marks along represent the watermark. The marks should not have a big impact on the quality of the info and that they ought to be placed in such some way that a malicious user cannot destroy them while not creating the info less helpful. Thus, watermarking doesn't stop repeating, however, it deters unlawful repeating by providing a way for establishing the initial possession of a decentralized copy.

Claudio Lucchese et al., describe the sharing is a crucial side of scientific or business collaboration. However, knowledge homeowners are a concern with the protection of their rights on the datasets, that is several cases are obtained when pricey and arduous procedures. the benefit of knowledge exchange through the medico has combined the requirement to assemble technological mechanisms for effectively protective one's intellectual or pragmatic property. Trajectories occur applications like GPS following experiments, video and motion capture knowledge, and even image shapes may be thought of as 2-dimensional trajectories. we offer possession assurances on such datasets victimization watermarking principles. whereas there's fashionable literature on watermarking for transmission datasets, previous work is primarily involved with watermarking one object and not a group of objects. Here, we tend to think about the watermarking drawback from a replacement perspective, by specializing in the extra maintenance of the inter-relationship between objects.

Vıctor R. Doncel, Nikos, et al., describes a watermark may be a hidden data among a digital signal, used primarily for copyright protection of transmission knowledge. Its main options are the physical property of the obligatory medications and its persistence against process (attacks) which will lead to its removal, either on purpose or accidentally. A general framework for digital watermarking has been given, whereas provides a superb summary of the watermarking principles and techniques. Digital watermarking has been in the main applied to still image, audio, and video knowledge. However, very little work has been worn out watermarking vector graphics knowledge, that is generally utilized in Geographic data Systems (GIS) or in laptop motor-assisted style (CAD).

Lukas mathematician and Jiri Matas., during this paper, we tend to gift associate end-to-end real-time1 text localization and recognition methodology that achieves state-of-the-art results on normal datasets. The period performance is achieved by motility the character detection downside as associate economical sequent choice from the set of Extremal Regions (ERs). The ER detector is powerful against blur, low distinction, and illumination, color and texture variation2. Its complexness is O(2pN), wherever p denotes a variety of channels (projections) used. within the 1st stage of the classification, the likelihood of every ER being a personality is calculable victimization novel options calculated with O(1) complexness and solely ERs with the regionally highest likelihood are chosen for the second stage, wherever the classification is improved victimization a lot of computationally valuable options. An extremely economical thorough search with feedback loops (adapted from [15]) is then applied to cluster ERs into words and choose the foremost probable character segmentation

Rakesh Agrawal and Kraut Kiernan et al., describe watermarking information relations to discourage their piracy, establish the distinctive characteristics of relative information that cause new challenges for watermarking, and supply fascinating properties of a watermarking system for relative information. A watermark may be applied to any info relation having attributes that area unit specified changes in a very few of their values don't affect the applications. a good watermarking technique engaged

for relative information. this method ensures that some bit positions of a number of the attributes of a number of the tuples contain specific values. The tuples attribute among tuples, bit positions in associate attribute, and specific bit values area unit all algorithmically determined underneath the management of a personal key far-famed solely to the owner of the information.

N. F. Johnson, Z. Duric describes the piracy of digital assets like software system, images, video, audio, and text has long been a priority for homeowners of those assets. The protection of those assets is typically based mostly upon the insertion of digital watermarks into the information. The watermarking computer code introduces tiny errors into the thing being watermarked. These intentional errors are referred to as marks and every one the marks along represent the watermark. The marks should not have a major impact on the utility of the information and that they ought to be placed in such some way that a malicious user cannot destroy them while not creating the information less helpful. Thus, watermarking doesn't forestall repetition, however, it deters bootleg repetition by providing a way for establishing the initial possession of a decentralized copy. The increasing use of information based on applications on the far side "behind-the-firewalls data processing" is making an analogous would like for watermarking databases

III. SCENE TEXT RECOGANIZATION MODEL

In the projected effective algorithms is text recognition from detected text regions in the scene image. In the scene text detection method, we tend to apply the ways bestowed in our previous work. Multi Pixel region-based analysis is adopted to extract text regions and phase text characters in pictures, supported color uniformity and horizontal alignment of text characters.

In-text recognition method, we tend to style two schemes of scene text recognition. the primary one is coaching a personality recognizer to predict the class of a personality in a picture patch. The other is coaching a binary character categorified for every character class to predict the existence of this class in a picture patch. The two schemes are compatible with two promising applications associated with scene text, that area unit text understanding, and text retrieval.

Text understanding is to amass text info from natural scenes to grasp close settings and objects. Text retrieval is to verify whether or not a bit of text info exists in the natural scene. These two applications will be widely utilized in sensible mobile devices. The projected methodology combines scene text detection and scenes text recognition algorithms recognition the character recognizer, text understanding is in a position to supply close text info for mobile applications and by the character categorified of every character class, text retrieval is in a position to assist explore for expect objects from setting.

Similar to different ways, our projected feature illustration is predicated on the various text format feature descriptors schemes. Different in different ways, our methodology combines the low-level feature descriptors with stroke configuration to model text character structure.

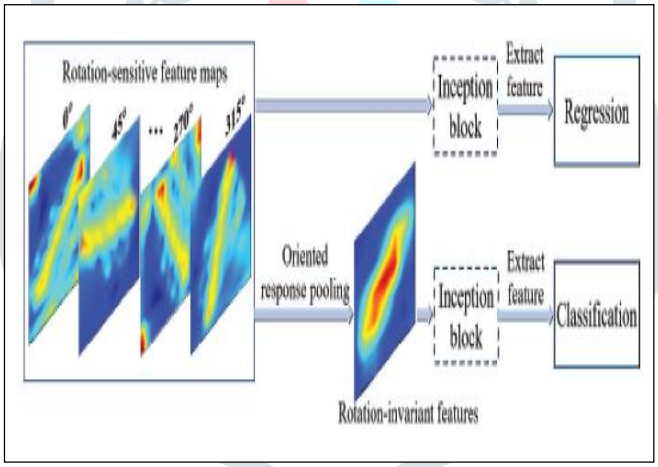

Fig 3.1. Text Segmentation Process

The target of existing model optical character recognition (OCR) is especially document images nonheritable by a scanner. Since even recent scanners have enough resolution for text image acquisition, the popularity rates of the many OCR methods will simply reach ninety-nine. Compared to ancient OCR, however, STR is tougher, that are mentioned as follows,

Texts are typically scattered within the scene image, and there are no previous data concerning their location. For scanned documents, the number of text lines, line spacing and even the amount of words are renowned. For scene texts, however, we tend to cannot directly apply segmentation ways for document pictures since there's no such info rule.

Scene texts typically have a style of sizes, fonts, and orientations. Targets in scene image might contain adorned or specially-designed characters, like presentation slides on a screen, script slogans on the wall, and messages on the digital structure. Such texts with multifarious appearance are difficultly recognized by ancient OCR engines

The quality of the scene image nonheritable by digital devices is probably poor. At present, scene text covers a wide variety of applications coupled to wearable cameras or huge urban captures that are tough or undesired ready for management. Therefore, characters and their backgrounds typically have a low distinction or perspective distortion, which ends in issue for localization and recognition.

There are several character-like patterns (non-character) in a scene image. Since the background of the scene image is usually complicated, there are several ambiguous objects like leaves, windows or icons that are very similar to characters or words. Moreover, generally scene texts connect with different objects, that simply end up in confusing patterns.

A. CC code

CC ways initial extract candidate parts from the image, then filter non-text parts mistreatment manually designed rules or mechanically trained classifiers.

Compared to SW ways, such ways are additional efficient and robust. There are two representative ways, i.e.,

- Stroke dimension rework (SWT)
- Maximally stable extremal regions (MSER)

SWT operator to reason the dimension of the foremost possible stroke for image constituent. the smart edge detector is 1st wont to realize edges in an image. MSER technique desires less prior information and is additional strong to language and minded text. Character candidates are clustered into text candidates by the single-link agglomeration algorithmic rule, whose distance weights and agglomeration threshold will be mechanically learned

B. Linguistics Segmentation Model

Scene text detection as a linguistics segmentation drawback. They use an FCN model supported holistically-nested edge detection (HED) to supply world maps, together with data of text region, individual characters, and their relationship. The projected algorithmic rule may discover multi-oriented and curving texts in scene images. The cascaded convolutional text networks (CCTN), that uses 2 networks to implement coarse-to-fine segmentation for scene image.

FCN based mostly model for scene text detection. Multiple channels of pixel-level text score map and pure mathematics may well be generated during this model, which is versatile to turn out either word level or line level predictions. This cluster of ways is appropriate for handling multi oriented text in the real-world scene images. Once text instances in image area unit terribly on the brink of one another, however, merely mistreatment text/non-text linguistics segmentation is difficult to separate them. Therefore, post-processing is usually inevitable to enhance performance.

C. General Object Detection Model

Text attention model, that encodes robust text-specific data employing a pixel-wise text mask. The general object detection model may effectively suppress background interference within the convolutional options. GODB Model describes a text-alignment layer and designed by introducing a grid sampling theme. It aims to reason mounted length convolutional options that exactly align to a detected text region with absolute orientation. The bounding box and segmentation mask of text may well be conjointly expected within the multi-task model

D. Character Classification Model

The deep neural network that's trained on HOG options for character classification. to boost the popularity performance, a two-level language model is adopted: a compact character-level n-gram model is control in RAM and a far larger distributed word-level n-gram model is accessed over the network. The RNNs may well be applied for learning character-level language model while not victimization n-grams. The soft-attention mechanism permits the model to pick out options flexibly for end-to-end coaching.

E. Word Classification Model

WCBM designed a context-aware convolutional perennial network for word recognition. Besides a lexicon wordbook, the information of the input image, like title, tags, and comments, area unit used as a context before enhance the popularity rate

F. Sequence Primarily Analysis Model

A convolutional perennial neural network (CRNN) for image-based sequence recognition. a typical CNN model is initially accustomed to extract a consecutive feature illustration from an input image. Then a bifacial long-short term memory (LSTM) network is connected with the highest convolutional layers to predict a label distribution for every frame of feature sequence. Finally, the connectionist temporal classification (CTC) is applied to search out the label sequence with the very best likelihood conditioned on the per-frame predictions.

G. CRF

Then a CRF model is made on the detection windows to make your mind up character locations. Finally, word recognition is enforced consistent with a value perform outlined by character detection scores, spatial constraints and linguistic Knowledge.

H. Finish Finishing Text Recognizing

Text detection and recognition area unit typically combined to implement text recognizing, instead of being treated as separate tasks. during a unified system, the recognizer not solely produces recognition outputs however conjointly regularizes text detection with its semantic-level awareness.

FCN to search out bounding boxes of text, supported that a RoIRotate operator is introduced to extract correct options from shared feature maps. Finally, the options of text proposals are fed to RNN and agency for text recognition.

I. Text Recognition Protocols

Given cropped word image, word recognition accuracy may be a usually used analysis metric, that is outlined because of the quantitative relation of the properly-recognized word range to the ground truth range. For holistic scene image containing texts, there area unit 2 protocols for analysis, i.e., word spot-

ting and end-to-end. Word recognizing solely examines whether or not the words in lexicon seem in an input image, and it ignores symbols, punctuations, numbers, and words whose length is less than 3. End-to-end protocol considerations each detection and recognition results, and it must acknowledge all the words exactly, regardless of whether or not the lexicon contains these strings.

Fig 3.2. Color Segmentation Process

In the color segmentation method long because the text regions area unit properly localized, ancient ways have already achieved comparatively high cropped word recognition accuracy. However, gift ways conceive to construct associate end-to-end framework while not sophisticated pre or post-processing for each text detection, recognition and extraction method.

A. Dataset

- The add alphabet module is used to add the alphabet character information in the application details such as character, image file path and image and store these information in the "images" file.
- The add number form is used to add the number information in the application details such as character, image file path and image and store these information in the "images" file.
- The add special character form is used to add the special character information in the application details such as special character, image file path and image and store these information in the "images" file.
- In this module, the image file is selected. The image type may be any format (e.g., bmp, jpg, gif, tiff, etc). Both gray scale and color image can be selected.

B. Text Recognition

- In this step, the selected image file is taken for processing. Then threshold value is set so that different size of character other than in training set can also be extracted from the given image. Spaces between characters are also measured so that words can be extracted out.

- In addition, characters can be of any colors in the give image. Using Character Skeleton and Character Boundary extraction, the character is recognized.

/*Multi Color Segmentation With Structure Analysis Model */

Input : Alphabet Image, Number Image, Special Character Image
Output : Extraction text given Image
Notation :
Alphabet Image AB,
Numerical Image NU,
Special Character Image SC,
Extraction Text ET
Create AB, NU and Sc Image
If AB image then
Select image and embedded alphabet text
If NU image then
Select image and embedded Number text
If SC image then
Select image and embedded Special Character text
End
End
End
Extraction image ()
If AB image then
AB image and Extraction for given alphabet text
Reorganization Character Skeleton and Character
Boundary extraction ET
End
If NU image then
NU image and Extraction for given Number text
Reorganization Character Skeleton and Character
Boundary extraction ET
End
If SC image then
SC image and Extraction for given Special Character text
Reorganization Character Skeleton and Character
Boundary extraction ET
End
Return ET

The below Figure 3.3 shows the text image file before the Scene Text Recognition.

Fig 3.3 Before- Scene Text Recognition process

The below Figure 3.4 shows the text image file after the Scene Text Recognition.

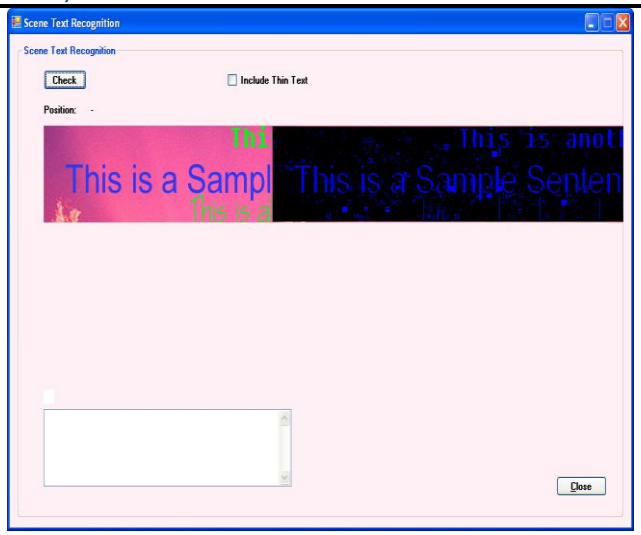

Fig 3.4 After- Scene Text Recognition process

IV.EXPERIMENTAL RESULTS

Table 4.1 describes a training dataset for flower defect CD & SC Model and Multi Color Segmentation Model (MCSM) analysis model. The table contains precision, recall, F-measure and accuracy details are shown

Table 4.1 Training Dataset Metrics Analysis

Techniques	Text Detection		Precision	Recall	F-measure	Accuracy
	Instances	No of Attributes				
CD & SC	850	13- Including class Label	0.715	0.764	0.777	0.782
MCSM	850	13 - Including class Label	0.755	0.803	0.818	0.804

Fig 4.1 describes a training dataset for flower defect CD & SC Model and Multi Color Segmentation Model (MCSM) analysis model. The figure contains precision, recall, F-measure and accuracy details are shown

Table 4.2 describes a training dataset for flower defect CD & SC Model and Multi Color Segmentation Model (MCSM) analysis model. The table contains precision, recall, F-measure and accuracy details are shown

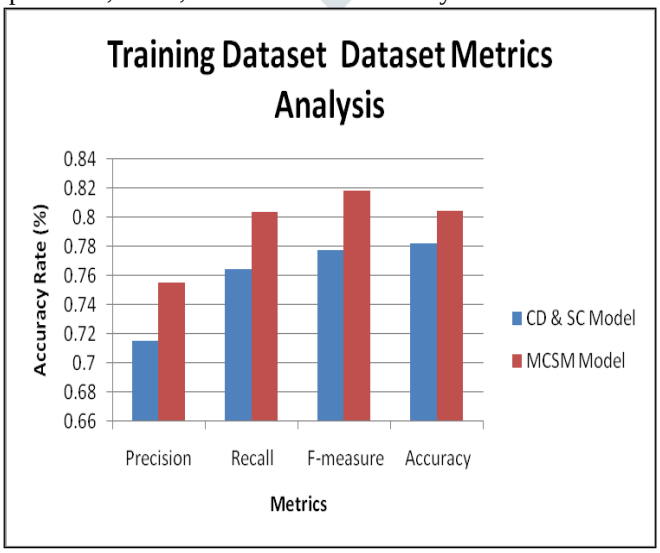

Fig 4.1 Training Dataset Metrics Analysis

Figure 4.2 describes a training dataset for flower defect CD & SC Model and Multi Color Segmentation Model (MCSM) analysis model. The figure contains precision, recall, F-measure and accuracy details are shown

Table 4.2 Test Dataset Metrics Analysis

Techniques	Text Detection		Precision	Recall	F-measure	Accuracy
	Instances	No of Attributes				
CD & SC	500	13-Including class Label	0.709	0.756	0.765	0.798
MCSM	500	13 - Including class Label	0.733	0.796	0.809	0.833

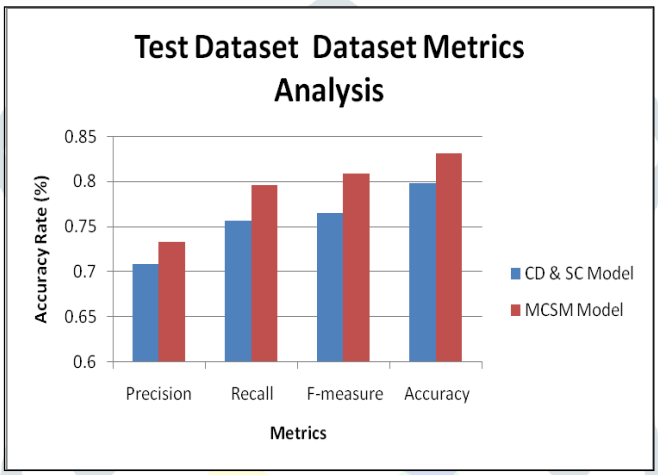

Fig 4.2 Test Dataset Metrics Analysis

V. CONCLUSION

In this paper is detects text regions from natural scene image, and acknowledges text info from the detected text regions. In scene text detection, layout analysis of color decomposition and horizontal alignment is performed to look for image regions of text strings. In scene text recognition, 2 schemes, text understanding, and text retrieval are severally planned to extract text info from close surroundings. The planned system character descriptor is effective to extract representative and discriminative text options for each recognition scheme. The model text character structure for the text retrieval theme designed a unique feature illustration, stroke configuration map, supported boundary and skeleton. Quantitative experimental results demonstrate that our planned technique of scene text recognition outperforms most existing ways.Also, the varied size of character pictures square measure saved in the coaching set. Threshold price is ready so that different size of character aside from in coaching set may also be extracted from the given image. Any image sort is often given as a supply image. areas between characters are measured so that words are often extracted out. additionally, characters are often of any colors within the offer image.

To require full advantage of matter info in natural scenes, it is necessary to scan texts of various orientations. Considering usefulness, it's crucial to develop detection and recognition systems that will handle texts of various languages.Further improvement in detection and recognition accuracy is often achieved if the deep learning framework is used to get and model the characteristics of scene text from a massive volume of information.

REFERENCES

[1] X. Bai, L. J. Latecki, and W.-Y. Liu, "Skeleton pruning by contour partitioning with discrete curve evolution," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 3, pp. 449–462, Mar. 2007.
[2] R. Beaufort and C. Mancas-Thillou, "A weighted finite-state framework for correcting errors in natural scene OCR," in Proc. 9th Int. Conf. Document Anal. Recognit., Sep. 2007, pp. 889–893.
[3] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87–99, Jan. 2004.
[4] Coates et al., "Text detection and character recognition in scene images with unsupervised feature learning," in Proc. ICDAR, Sep. 2011, pp. 440–445.
[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2005, pp. 886–893.
[6] T. de Campos, B. Babu, and M. Varma, "Character recognition in natural images," in Proc. VISAPP, 2009.
[7] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. CVPR, Jun. 2010, pp. 2963–2970.

[8] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
[9] T. Jiang, F. Jurie, and C. Schmid, "Learning shape prior models for object matching," in Proc. CVPR, Jun. 2009, pp. 848–855.
[10] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Johsi, "Text extraction and document image segmentation using matched wavelets and MRF model," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2117–2128, Aug. 2007.
[11] L. J. Latecki and R. Lakamper, "Convexity rule for shape decomposition based on discrete contour evolution," Comput. Vis. Image Understand., vol. 73, no. 3, pp. 441–454, 1999.
[12] Y. Liu, J. Yang, and M. Liu, "Recognition of QR code with mobile phones," in Proc. CCDC, Jul. 2008, pp. 203–206.
[13] S. Lu, L. Li, and C. L. Tan, "Document image retrieval through word shape coding," IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 11, pp. 1913–1918, Nov. 2008.
[14] S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young, "ICDAR 2003 robust reading competitions," in Proc. Int. Conf. Document Anal. Recognit., Aug. 2003, pp. 682–687.
[15] A. Mishra, K. Alahari, and C. V. Jawahar, "Top-down and bottom-up cues for scene text recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 1063–6919.
[16] N. Nikolaou and N. Papamarkos, "Color reduction for complex document images," Int. J. Imag. Syst. Technol., vol. 19, no. 1, pp. 14–26, 2009.