TEXT EXTRACTION ALGORITHMS AND METHODS FOR NATURAL IMAGES: A REVIEW

Kannav Dhawan, Akriti Tyagi, Akshay Goyal, Ishank Sharma

ABSTRACT

Text extraction plays a significant role in getting important information which is useful in various fields like medical, engineering, security department etc. and is one of the recent growing technique to be enhanced further. Text Extraction is the process of detection, localization followed by the extraction of the required text to be analyzed. The text is hard to be identified and perceived because of differences in text properties. This review paper studies and analyses different algorithms or techniques previously proposed by researchers for text extraction and illustrates the shortcomings and limitations of text extraction approaches with respect to speed, recall and precision rate of their respective models.

Index Terms— Text Extraction, Detection, Texture based. Edge, Connected Component, Morphological.

1. INTRODUCTION

Extracting text from Images and videos is an important issue in applications such as document processing, Image assortment, video retrieval etc. Images are loosely classified as document, caption and scene text images; caption images are referred as artificial images in which the text overlaps an image or video, document images contain the scanned paper or books having figures, tables and text which can have varying text properties; whereas scene text images are the graphical text images which can have the landscapes with billboards or road signs with blurred text usually vertically or horizontally aligned. The formation of Image includes perceptual and the semantic content where perceptual content refers to color, texture and shape of text whereas picture objects, events and their corresponding relation forms the semantic content. Texts have a definite height and width range and certain proximity between its characters within a text string. Text also possesses properties such as contrast, background, orientation and alignment. Usually texts embedded in the image or a frame apprehend media contexts like title, dates, character names, story abstract etc. Researchers have developed multiple text extraction techniques to avert the issues of varying and scalable text properties which are mentioned above. Region based text extraction approach is based on sharp variation in certain key properties like color intensity, gray scale of text and its immediate background [1], while there is little or no variation of color intensity within text [4]. Text recognition depends on the various properties of the image text which can be color, alignment, size, edge and the compression format. Since all text are made readable from distance, they have sharp edges which forms basis of region-based text extraction. If we threshold image at intensity level which lies between that of the text and its corresponding immediate background, we will be able to extract the required text. This method is only limited to simple backgrounds with no complexity involved [1]. One significant method based on region-based approach has been introduced and is known as edge-based text extraction method. As its name signifies it extracts the text based on sharp variations in color intensity compared with immediate background [1], [6], [7], [8]. There are certain characteristic properties of text which makes it detectable in images namely edge strength, density and the orientation variance [1]. This technique can be used for speedy text extraction (localize and extract) in different spatial conditions which makes it one of the most standardized technique in the given domain. The similar valued connected pixel clusters with a linkage of 4 and 8 pixels for each of the four faces and the corners respectively in a 2-D image are the connected components which guarantees the enclosed pixels results in a single character [15]. This method deploys the bottom-up approach for each component until all the components are traversed [16].

The texture-based method uses various kinds of texture and its properties to extract and evaluate a text from a complex image. Various types of methods used for this text extraction approach are Wavelets, Fourier Transform and Gabor filters, DCT Transform Wavelet [20]. Mathematical Morphology is a topological and geometrical based approach for Image Analysis [11]. Main purpose of Morphological Image processing is to remove unwanted artifacts from an image or improve its clarity by modifying the pixels in an image. Morphological algorithms language is based on the concepts of set theory, topology, lattice algebra and function. For Binary or Grey scale image analysis and evaluation, morphological operators are used which performs fundamental operations such as Erosion, Dilation, Opening and Closing.

2. LITERATURE REVIEW

2.1 Region Based Approach Analysis

In thesis, John Canny [2] attempted to identify a group of three edge detection criteria that gave exact or close to desirable properties of an edge operator. The basic criterion for good detection revolves around the idea that detector should be able to accurately spot the edges and should not identify non-edges as edges. Mathematically it means POE (probability of error) should be minimized to increase efficiency of detector. Since the probability decreases by increasing the signal to noise ratio and same is our goal in this criterion. The points identified for edges should be close to the middle point of true edge for good localization. The last criterion is there should be single response on searching any individual edge. On analysis John Canny found the first two criteria for selection of edge detector are self-contradictory at certain points. It is evident in case of

broad operators because even though having high signal to noise ratio, they performed poorly in localization. According to the second observation, both criteria combined do not fully cover the expectations of good edge detector. Now mathematically, good detection means maximizing the SNR (signal to noise ratio). So instead individually maximizing the value of good detection and good localization, we maximize the product of the two to simplify the analysis [3]. We can derive edge detectors which provide optimal performance using numerical optimization.

$$\text{SNR}(f) * \text{Localization}(f) - \sum \mu_i P_i(f)$$

$\sum \mu_i P_i(f)$ is the numerical optimization part where if any constraint is violated P_i value becomes positive and if most of the constraints are being followed μ_i has a large positive value [2]. This method is also called penalty method and gives us the most optimal response wherein we maximize the above equation. Now for the final optimization step we remove noise from the edge detector operator output using Wiener filtering. We find the noise component of the signal and reduce some percentile of it [3]. Xiaoqing and Jagath [10] proposed a basic and quick text extraction algorithm based on statistic approach which worked by measuring the difference in the spatial intensities. The major disadvantage of this algorithm was in uneven environmental lighting conditions it fared poorly. Xiaoqing and Jagath [11] then proposed a sturdier edge-based algorithm which made indoor image text extraction much more effective by taking in account most of the environmental conditions like types of fonts and their sizes, styles, intensity variations, alignment of texts in images, spatial conditions, shadows and reflection of surroundings [11]. As an improved version, Xiaoqing and Jagath [12] suggested a multiscale edge-based text extraction algorithm which overcome all the disadvantages and could be efficiently used for both well-lit and less lit spatial conditions based automatic text extraction. Its computation is highly efficient because text blocks are the unit of analysis. Jing et al. [13] in their approach of text detection analyzed the properties of edges, used histogram of Oriented Gradient for effective localization of blocks of character using intrinsic characteristics of edges followed by Graph based spectrum to find the global linkage of text properties which aided in grouping, hence lead to much more effective text detection. This algorithm is sturdy to the color and text size variations. Zhihu and Jinsong proposed much improved edge-based text extraction algorithm which not only use heuristic rules concerning shape and texture but also included improved stroke width rules to filter out non - text regions [14]. Stroke width rules were improved by reduction of computing time of Stroke Width Transform (SWT) by limiting its range and exceptional accuracy of edge point search by increasing the pixels used for search which made algorithm robust.

2.2 Connected Component Approach Analysis

A robust approach to detect horizontal text deploying color reduction technique followed by the localization using geometrical properties and projection profile analysis outputting text boxes ready to be recognized by OCR has been given by Gllavata et al [15]. Their approach generated edge image using contrast segmentation which separates text and image using contrast of the text pixels and the background pixels proceeded by the analysis of the horizontal projection to locate the text regions giving a recall of around 88.7% with a precision rate of 83.9%. In [16], Chen and Yuille proposed an algorithm which performs statistical analysis on the text region to indicate the low entropy features. They used 4 strong classifiers having 79 features developed and trained by feeding the weak classifier as log-likelihood ratio test which in turn was made by using the joint probabilities to detect the feature response into AdaBoost Machine learning algorithm to train a robust cascade classifier which selects regions for adaptive binarization and application of extension algorithm. The application of strong classifier resulted in 2–5 false positives in images of 2,048 − 1,536 pixels. This method resulted in correct text detection of 93% within a span of 3 seconds of running the algorithm for the images of size 2048x1536. Still less improved adaptive binarization and untrained AdaBoost for vertical text left the small, blurred and shadowed text undetected. A novel and data-dependent Stroke width image operator to find the value of stroke width for each image pixel for text detection irrespective of scale, language and font of the text was proposed by the Epshtein et al. [17] eliminating the requirement of multi-scale computation. For every single pixel, stroke width transform converts the grey values of the image to an array consisting of similar stroke widths for extracting the text by the stroke width variance measurement. Being a multilingual approach, it is independent of the features like size, edges, gradient and color. The detection of the curved text is still an issue with this approach which can be solved by letter grouping using the directional orientation of the strokes. Koo and Kim [18] proposed an algorithm using machine learning classifiers which sorts out between the candidate text and non-text regions. It basically employs the stable extremal region algorithm to find the connected components which are partitioned into clusters by AdaBoost classifier to spot the candidate text regions which are normalized in the proceeding steps. They have developed an approach to exploit multi-channel information. Due to the optimal pixel-level classification of deep neural networks, Jiang et al. [19] proposed two task network with integrated bottom-up cues for deep scene text detection where the first task using canonical connected component analysis does the pixel labelling to generate word proposals whereas the second step outputs the bundle of text candidates to verify the word proposals generated in first step. They have achieved a recall rate of 0.915 in ICDAR text localization process.

2.3 Texture Based Approach Analysis

A novel texture-based method for detecting texts in images proposed by Kim et al. [21] where a support vector machine (SVM) is used to analyze the textural properties of texts. No external text extraction module is used, rather the raw pixels that make up the textural pattern are fed directly to the SVM, which works in higher-dimensional spaces. Later, text regions are identified by applying a continuously adaptive mean shift algorithm (CAMSHIFT) to the output of the text analyzed. Combining both CAMSHIFT and SVMs produces sturdy and efficient text detection, as time-consumption for texture analyses for less relevant pixels are restricted, leaving only a small part of the input image that is required to be texture-analyzed. Zaravi et al. [22] proposed a new powerful and efficient algorithm to automatic text extraction from colored book and journal covers using wavelet transform and then

detecting edge from detail wavelet coefficient using dynamic threshold. Their method was relatively simple yet effective and algorithm was tested on a very small dataset (80 images). However, result was satisfying approximate 91.2 percent hit rate. The proposed method by Kim et al. [21] works properly in extracting texts from complex and textured backgrounds and produced a lot better performance than other techniques. However, proposed texture classifier did encounter problems classifying very small text or text with a low contrast. Navjot Kaur [23] extended the technique used for text extraction from document Images using 2-D Haar Wavelet. She implemented an effective method using the fact that in text regions horizontal, vertical and diagonal edges are mixed together while they are separately distributed in non-text regions. She also emphasized that there's a wide scope for improvement in text segmentation techniques.

Aradhya et al. [24] combined the wavelet transforms and Gabor filter to extract sharpened edges and textural features of an image. The proposed method detects the text of different font sizes, complex background and contrast. Furthermore, the system performance surpassed the existing methods in terms of detection accuracy. Yeotikar et al. [25] worked on the implementation of following methods: Spatial domain filter, 2D-Fast Fourier Transform, 2D Fourier Transform, Image Enhancement and Edge Reinforcement. Proposed methods are used for purposes like selection and separation of text element from document image, removing noise from the document. Later in 2015 Yeotikar et al. [26] proposed an efficient and robust method by using three feature extraction techniques i.e. Gabor, Wavelet and Hough to detect text objects from document images. The performance of the proposed method is tested on NIST document Image dataset and it outperformed all the compared state-of-the-art and baseline algorithms. Hiremath et al. [27] proposed a novel system which uses a texture-based method. Various steps which are involved in their proposed system are pre-processing, feature extraction, classification, merging of text. This system has attained a detection rate of 99% when compared to connected component method and is proved to be more robust while dealing with complicated backgrounds and other image properties which poses a challenge in extracting text like different text font size, color etc. This technique accomplished excellent results for natural scene images (each of which resized to 300×300), detects text regions which are not linearized and incorporated in applications such as helping physically challenged persons, tourists guide license plate reading and so on.

2.4 Mathematical Morphology Review

A novel morphological text extraction technique (incorporating fast filling and Hit and Miss algorithm) from color images, grey-scale images and video has been proposed by Hasan and Karam, with a few contrast, color and size restriction assumptions. Their technique is robust to noise, insensitive to skew and text orientation [28]. Wong and Chen [29] developed a robust algorithm for text extraction from color video using line scan approach, which eventually also worked well with 'high quality' text and noisy images after JPEG Compression – Decompression. However multiple limitations and drawbacks of [28], [29] have been discussed and resolved experimentally in [30]. In [30], it is claimed that Hasan et al. [28] algorithm is sensitive to clustered/complex backgrounds, fragmented text lines and Wong et al. algorithm [29] is sensitive to text fonts and increases the possibility of false alarms for text candidates. To overcome, [30] proposed a morphological approach with a x-projection technique and a text recovery algorithm which is successful in detecting all kinds of text lines even under clustered backgrounds. Omar et al. [31] proposed a new skeletonization algorithm (or iterative thinning algorithm) which eliminated the problems such as skeleton width, spurious branches, distortion and tolerance to invariance rotation and delivered high quality performance for any shape binary images with any rotation. However, a disadvantage of using skeletonization algorithm is that it leaves considerable amount of fuzz, short offshoots that stick out from the sides of longer segments [32].

With the rapid advances in digital technology, most of the researcher's interest has inclined towards deep learning and they've started proposing unique and much faster techniques using CNN (Convolutional Neural Network). CNN's have faced many achievements in the areas of handwriting recognition, visual object recognition and character recognition. By incorporating large multi-layer CNN's, an end to end text recognition systems approach has been proposed which achieved state-of-the-art performance on standard benchmarks [33]. Alimi et al. [34] developed an architecture called as Morph-CNN (combination of Morphological filters and CNN) for digit recognition where they are successfully able to recognize the label from Image feature map which is not the case with images resulting from morphological operators. This represent the power of Neural Networks and indicates the departure from previous traditional text extraction systems which are generally relied on intricate graphical models or hand engineered systems. [35] illustrated a CNN based text extraction approach where they also considered multi-oriented text and Thai texts and delivered excellent results. To make the world safer and a better place to live, researchers keep on exploring neural network applications in different fields. N. Wang et al. [36] proposed a morphological based text extraction approach for vehicle license plate detection using BPNN (Back Propagation Neural Network). Jongbae Kim [37] developed a morphological method for the 'real time detection' of small vehicles and their license plate using R-CNN (Region Based Convolution neural network) however they had to keep size of the input image limited to maintain a readable resolution for license plate character recognition. Yang et al. [38] constructed a robust and precise tracking based multi-orientation text detection method from videos using dynamic programming and a hybrid machine learning filter designed with CNN, AdaBoost and Bayesian classifiers. Impressively, their proposed system won the first place of "Video Text Detection" in the ICDAR 2015 Robust Reading Competition. On the other hand, their tracking technique could be improved in order to deal with challenges like severe motion blur and occlusions that results in the degradation of text tracking and authors showed their concern regarding computational cost of multi-oriented text detection.

CNN is a hot and demanding topic in Machine Learning and Computer Vision [39]. Various improvements have been made in past few years by using CNN based models with Morphological Operations for Text Extraction and proactive research on multiple models for achieving better performance

and results are already in place.

3. LITERATURE ANALYSIS

Edge-based text extraction is a standard technique having quick and effective extraction. Although current Edge based techniques work in different environmental conditions, they are not that sturdy when flawed text detection, localization and extraction is considered. Major short coming is that in most of the cases it cannot handle large volume of data. Further if the spatial conditions are not well lit, most of edge-based techniques result are poor in performance. Also, complex backgrounds may imitate text blocks and make edge-based text extraction highly inaccurate. All these shortcomings can be tackled by non-classical technique called deep learning (deep neural networks) which allows unsupervised learning of characteristics to improve the accuracy of text extraction algorithms without manual modelling.

Although the connected component optimal localization is achieved deploying the pixel-level bottom cues and object-level top cues unification but still the task interaction in the multi-task approaches is still an issue which is partially solved by the introduction of the consistency loss which forces the multiple tasks to comply with each other. The optimized localization and extraction can be achieved by using the modified hybrid network of ParseNet and Fast R-CNN with a range of varied parameters to achieve the best recall rate. The shortest path problems of clustering the pixels can be solved by either using Dijkstra algorithm by connecting each pixel at the shortest distance from the previous one or by using Shortest path faster algorithm which is an improvement of bellman-ford algorithm.

Researchers have worked on various methods to extract the text in an image like CAMSHIFT algorithm, Gabor, Wavelet, Hough and so on which comes under the texture-based approach. Many challenges were resolved as the research progressed in this area and said approach being more skillful produced better results for natural scene images. Multiple successful experiments on datasets with text having different textural properties, low contrast has been completed. Also, noise in the document using wavelet transform has proved that it gives more hit rate compared to connected component method. Feature extraction has been improved by Wavelet method. However, hybridization of DCT and DWT gives better generalization performance.

Most of the studies focus on contrast enhancement for Image processing and there are various methods available for the same such as Histogram Modification methods, Spatial Domain and Frequency Domain Filtering etc. All these methods have their unique characteristics however they all usually enhance the entire structure in an image without any discrimination. Mathematical morphology or Morphological based approaches resolved this issue and proved to be an efficient method in enhancing specific part (any shape or size) of an Image and not its surrounding area.

4. CONCLUSION

In this paper, we have reviewed multiple text extraction techniques proposed by researchers and illustrated the peak differences between the approaches citing the precision, accuracy and feasibility of the algorithm for the distinct images having variable properties and backgrounds. The in-depth analysis of the methods and models suggested that there are certain limitations of the algorithms being used to create the respective model at steps which were resolved by bringing in the application of multi-layered neural networks (CNN's) and dynamic programming. Recently, researchers have proposed a method to identify the Fake news on web by training the algorithms with a huge set of data whose performance is being analyzed. However, researchers are still paying a huge amount of attention to the problems associated with text extraction and enhancement by incorporating multiple optimal algorithms to achieve state of the art results and performance.

5. REFERENCES

[1] Divya gera, Neelu Jain, "Comparison of Text Extraction Techniques- A Review", International Journal of Innovative Research in Computer and Communication Engineering, (An ISO 3297: 2007 Certified Organization), Vol. 3, Issue 2, February 2015.

[2] John Canny, "Finding edges and lines in images", MIT Libraries, DSpace@MIT, Artificial Intelligence Lab Publications, AI Technical Reports (1964-2004).

[3] John Canny (Member, IEEE), "A Computational Approach to Edge Detection", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. PAMI-8, No.6, November 1986.

[4] S. Hassanzadeh, H. Pourghassem, "Fast Logo Detection Based on Morphological Features in Document Image", 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, 2011, pp. 283-286.

[5] A.N. Lai, G.S. Lee, "Binarization by Local K-means Clustering for Korean Text Extraction", IEEE Symposium on Signal Processing and Information Technology, 2008, pp. 117-122.

[6] H. Anoual, D. Aboutajdine, S.E. Ensias, A.J. Enset, "Features Extraction for Text Detection and Localization",5th International Symposium on I/V Communication and Mobile Network, IEEE, 2010, pp. 1-4.

[7] S.V. Seeri, S. Giraddi and Prashant. B. M, "A Novel Approach for Kannada Text Extraction", Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering, 2012, pp. 444-448. 16.

[8] Xiaoqing Liu and Jagath Samarabandu, "Multiscale Edge Based Text Extraction From Complex Images", 1-4244-0367-7/06/$20.00 ©2006 IEEE, ICME 2006.

[9] X. Liu and J. Samarabandu, "A simple and fast text localization algorithm for indoor mobile robot navigation," in Proc. of the SPIE- IS&T Electronic Imaging 2005, San Jose, California, USA, Jan. 2005, vol. SPIE vol. 5672, pp. 139–150.

[10] X. Liu and J. Samarabandu, "An edge-based text re- gion extraction algorithm for indoor mobile robot navigation," in Proc. of the IEEE International Conference on Mechatronics and Automation (ICMA2005), Niagara Falls, Canada, July 2005, pp. 701–706.

[11] Jing Zhang and Rangachar Kasturi, "Text Detection

Using Edge Gradient and Graph Spectrum", 2010 International Conference on Pattern Recognition, IEEE Computer Society.

[12] Zhihu Huang and Jinsong Leng, "Text extraction in natural scenes using region-based method", Journal of Digital Information Management, Volume 12, Number 4, August 2014.

[13] Y. Gupta, S. Sharma and T. Bedwal. Article: Text Extraction Techniques. IJCA Proceedings on National Seminar on Future Trends and Innovations in Computer Engineering NSFTICE 2015(1):10-12, August 2016.

[14] X. Yin, X. Yin, K. Huang and H. Hao, "Robust Text Detection in Natural Scene Images," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, pp. 970-983, May 2014.

[15] J. Gllavata, R. Ewerth and B. Freisleben, "A robust algorithm for text detection in images," International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the, Rome, Italy, 2003, pp. 611-616 Vol.2.

[16] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, Washington, DC, USA, 2004, pp. II-II.

[17] B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 2010, pp. 2963-2970.s

[18] H. I. Koo and D. H. Kim, "Scene Text Detection via Connected Component Clustering and Nontext Filtering," in *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2296-2305, June 2013.

[19] F.Jiang, Z.Hao & X.Liu Deep Scene Text Detection with Connected Component Proposals (2017).

[20]

[21] Kwang In Kim, Keechul Jung and Jin Hyung Kim, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1631-1639, Dec. 2003.

[22] Zaravi, D. & Rostami, Habib & Malahzaheh, A. & Mortazavi, S.S.. (2011). Journals subheadlines text extraction using wavelet thresholding and new projection profile. World Academy of Science, Engineering and Technology. 73. 686-689

[23] Kaur, Navjot. "Modified Method of Document Text Extraction from Document Images Using Haar DWT." (2012).

[24] Aradhya, V. N. Manjunath et al. "A Robust Multilingual Text Detection Approach Based on Transforms and Wavelet Entropy." (2012).

[25]

[26] Yeotikar, Vikas K. et al. "Text Extraction from Document Images Using Gabor, Wavelet and Hough Technique: A Novel Approach." (2015).

[27] Hiremath, Tejashvini & Patil, Shivanand & Malemath, Virendra. (2015). Detection and Extraction of Text in Images using DWT. 4. 2015.10.17148/IJARCCE.2015.46115

[28] Hasan, Y.M.Y., Karam, L.J.: Morphological text extraction from images. IEEE Trans. Image Process. **9**(11), 1978–1983 (2000).

[29] Wong, E.K., Chen, M.: A new robust algorithm for video text extraction. Patt. Recognit. **36**(6), 1397–1406 (2003).

[30] Wu, Jui-Chen & Hsieh, Jun-Wei & Chen, Yung-Sheng. (2008). Morphology-based text line extraction. Mach. Vis. Appl.. 19. 195-207. 10.1007/s00138-007-0092-0.

[31] Waleed Abu-Ain, Siti Norul Huda Sheikh Abdullah, Bilal Bataineh, Tarik Abu-Ain, Khairuddin Omar, Skeletonization Algorithm for Binary Images, Procedia Technology, Volume 11, 2013, Pages 704-709, ISSN 2212-0173.

[32]

[33] T. Wng, D. J. Wu, A. Coates and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, 2012, pp. 3304-3308.

[34] D. Mellouli, T. M. Hamdani, J. J. Sanchez-Medina, M. Ben Ayed and A. M. Alimi, "Morphological Convolutional Neural Network Architecture for Digit Recognition," in IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2876-2885, Sept. 2019.

[35] T. Kobchaisawat and T. H. Chalidabhongse, "A method for multi-oriented Thai text localization in natural scene images using Convolutional Neural Network," 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, 2015, pp. 220-225.

[36] Chuin, M.W.; Jian, H.L. License plate recognition system. In Proceedings of the Conference on Fuzzy Systems and Knowledge Discovery, Zhangjiajie, China, 15–17 August 2015; pp. 1708–1710.

[37] Kim, JongBae. (2019). Automatic Vehicle License Plate Extraction Using Region-Based Convolutional Neural Networks and Morphological Operations. Symmetry. 11. 882. 10.3390/sym11070882.

[38] C. Yang et al., "Tracking Based Multi-Orientation Scene Text Detection: A Unified Framework With Dynamic Programming," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3235-3248, July 2017.

[39] Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks, Yushi Chen, Member, IEEE, Hanlu Jiang, Chunyang Li, Xiuping Jia, Senior Member, IEEE, and Pedram Ghamisi, Member, IEEE, 2016.