Handwritten Devanagari Word Detection and Localization using Morphological Image Processing

2023 10th International Conference on Signal Processing and Integrated Networks (SPIN) | 978-1-6654-9099-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/SPIN57001.2023.10116577

Mamta Bisht
Electronics and Communication Engineering
Jaypee Institute of Information Technology
Noida, India
17402010@mail.jiit.ac.in

Richa Gupta
Electronics and Communication Engineering
Jaypee Institute of Information Technology
Noida, India
richa.gupta@jiit.ac.in

***Abstract*- In text analysis systems, text detection and localization techniques are preliminary steps for further text recognition tasks. Finding the specific text information in the image is important as it has many significant real-time applications. The available techniques for text detection and localization work well for printed text, but their performance is not good for handwritten text present in an image. Also, the available modern deep learning techniques require large datasets for training. So, if small handwritten data images are available, then image processing and an effective morphology-based technique for text detection and localization can be used. In this paper, these techniques are used for handwritten documents written in the Devanagari script.**

Keywords- Handwritten text; Devanagari script; text detection and localization; morphological image processing.

I. INTRODUCTION

The text analysis system is an active research area in computer vision and its general procedure is presented in Fig. 1 [1]. Text information is an important representation tool, and every step used in a text analysis system plays an important role for further processing steps such as script identification [2] and text recognition [3], [4]. The layout analysis, i.e., word, line, or paragraph segmentation, is necessary before proceeding to the transcription of the document image. Various factors, like different orientations, backgrounds, fonts, sizes, and uneven lighting conditions, etc., cause challenges in text analysis systems [5].

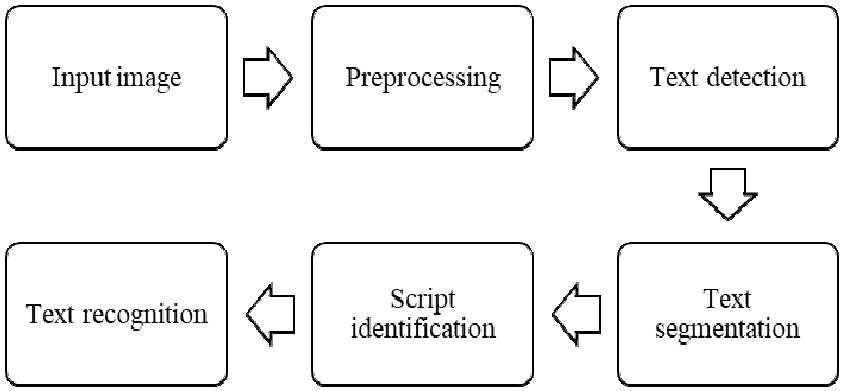


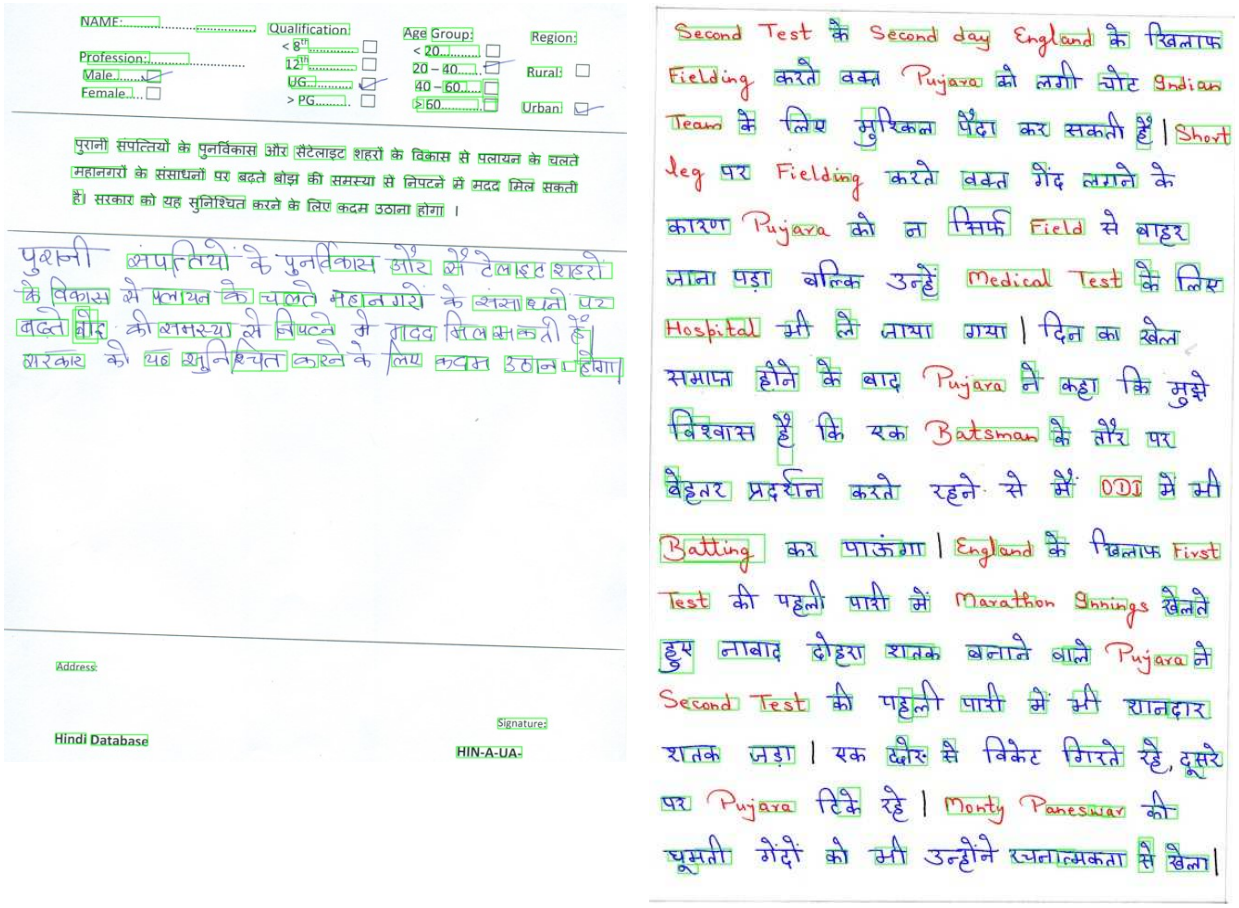
Fig. 1. General procedure of text analysis system

Nowadays, with the arrival of deep learning-based models, the performance of text detection has improved significantly. But these models require large datasets for training and further testing datasets for model evaluation, and if we have limited datasets, then the deep learning-based model causes poor performance. Hence, for small datasets, we can think of using simple image processing and morphological operations for text detection. Another thing is that the text detection models that are available work well for printed text but not for handwritten text.

The performance of one of the most popular deep learning-based pre-trained EAST models is shown in Fig. 2 for two handwritten documents in Devanagari script. It is clear that the model works well for the printed parts of text, but the performance is poor for handwritten text. This motivates researchers to explore text detection and localization work in handwritten documents specifically written in non-Latin scripts such as Devanagari script. In this paper, an image processing and morphology-based method is proposed to detect and localize text on publicly available document images written in Hindi (Devanagari script), which is one of the most popular as well as official language of India.

978-1-6654-9099-3/23/$31.00 ©2023 IEEE

Authorized licensed use limited to: Sardar Patel Institute of Technology. Downloaded on April 06,2024 at 06:36:39 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 2. Text detection and localization using pre-trained EAST model for handwritten text in Devanagari script: (a) CALAM handwritten text dataset (b) Cmaterdb dataset 1.5.1

II. Related Work

The existing work for text detection can be broadly classified into two parts: traditional methods and deep learning-based methods. Traditional methods manually extract features to distinguish text from background [6]–[11], whereas deep learning-based methods extract effective features from training datasets during the learning process [12]–[19].

Stroke Width Transform (SWT) [6] and Maximally Stable Extremal Regions (MSER) [7] are popular conventional methods for manually designing features for text detection tasks. The local symmetry property of text is also considered by Zhang et al. [20] to design features for text detection tasks in natural scene images. Busta et al. [21] also worked on fast text detection and improved the FAST key point detector for stroke extraction. Dhar et al. [22] proposed a gradient morphology-based method for multilingual scene text detection. Recently, deep learning-based methods have been showing significant performance for text detection problems [1], [23], [24]. Tian et al. [19] presented a Connectionist Text Proposal Network (CTPN) and Zhou et al. [18] presented an Efficient and Accurate Scene Text detection (EAST) pipeline for natural images. In the deep learning era, work on irregular scene detection is also found [25].

Text detection work is mostly found for the Latin script, with less for other scripts. India is a country where many languages are spoken and Hindi (Devanagari) is one of the most popular languages. Although a small piece of work is found by Bhattacharya et al. [26], Raj and Ghosh [27], and Rahul et al. [28] for text detection in scene images, it is limited to printed text only. Therefore, the text detection work is required to extend to handwritten text in Devanagari script too.

III. Methodology

After converting the input image to grayscale, it is binarized using adaptive thresholding to separate the background and foreground portions of the document image. Then, opening morphology with 3×3 structuring element of ones is applied to the pre-processed image for noise removal. Connected component analysis with a connectivity value of eight is also applied for the elimination of noise. Then a dilation morphological operation is used to add boundary pixel values so that text contours can be decided for segmentation. In the dilation process, the filter size was varied from 7×25 to 15×25 and it was observed that the 7x25 value provides good contours, so it was selected for this work. The block diagram of the used methodology is shown in Fig. 3.

Authorized licensed use limited to: Sardar Patel Institute of Technology. Downloaded on April 06,2024 at 06:36:39 UTC from IEEE Xplore. Restrictions apply.

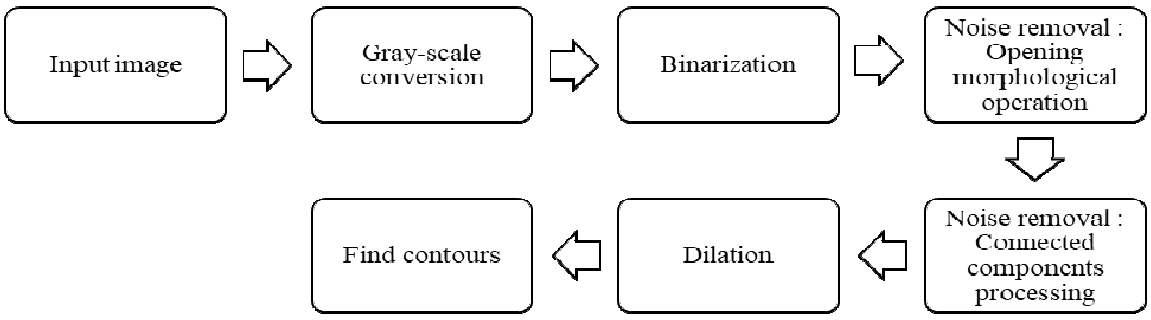


Fig. 3. Flow chart of the proposed method

Typically, handwritten documents are digitized by first segmenting the document into lines and then segmenting those lines into words. Some researchers further segment words into characters [29]. Here, the proposed approach is capable of detecting text as a word and localizing it with image processing and morphological operations.

IV. Dataset Description

This work uses three datasets of handwritten documents. Firstly, CALAM handwritten text dataset of 50 samples is used, which was made available at the International Conference on Computer Vision & Image Processing 2019 at Malaviya National Institute of Technology, Jaipur. The second is the Cmaterdb dataset version 1.5.1 [30] benchmark dataset, comprising of 150 Devanagari documents developed by Jadavpur University, Kolkata. These documents were gathered from a variety of sources, including class notes from students of various ages, document pages written by various individuals, and so on. The third is the PHDIndic_11 benchmark handwritten document image dataset, and we chose the Devanagari script from 11 Indian regional scripts present in the dataset. This selected database consists of 220 Devanagari handwritten documents.

V. Results

The process of the proposed method is presented in Fig. 4 with one handwritten text sample, and it looks well performed. Since the dataset is not formed according to the standard text detection datasets, evaluation metrics such as precision, recall, and f-score are not possible to evaluate. As the work to detect and localize text for the Devanagari script is less, especially for handwritten text formats, we attempted this work using image processing and morphological operations. However, we tried to calculate the percentage of accuracy for the CALAM handwritten text dataset in terms of the number of detected words divided by the total handwritten words as 91.47% which is satisfactory. Fig. 5 presents the qualitative results of the proposed method on selected three datasets.

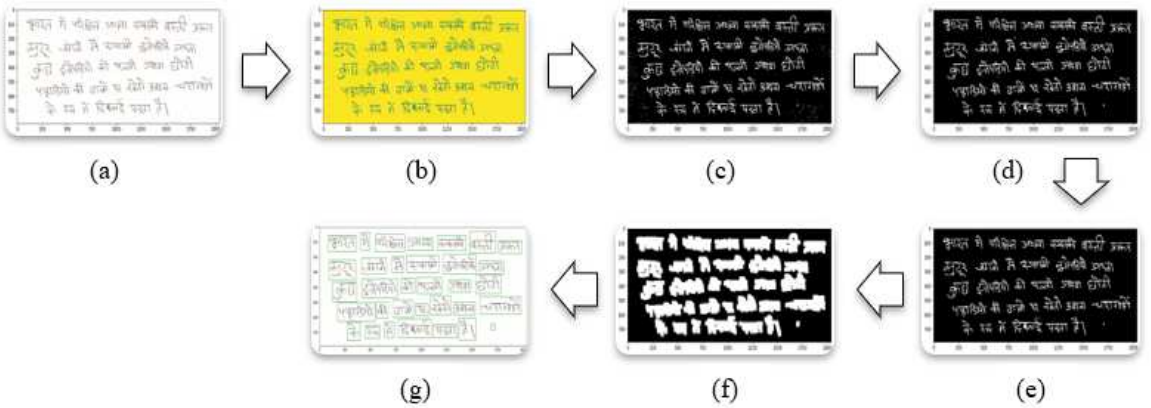
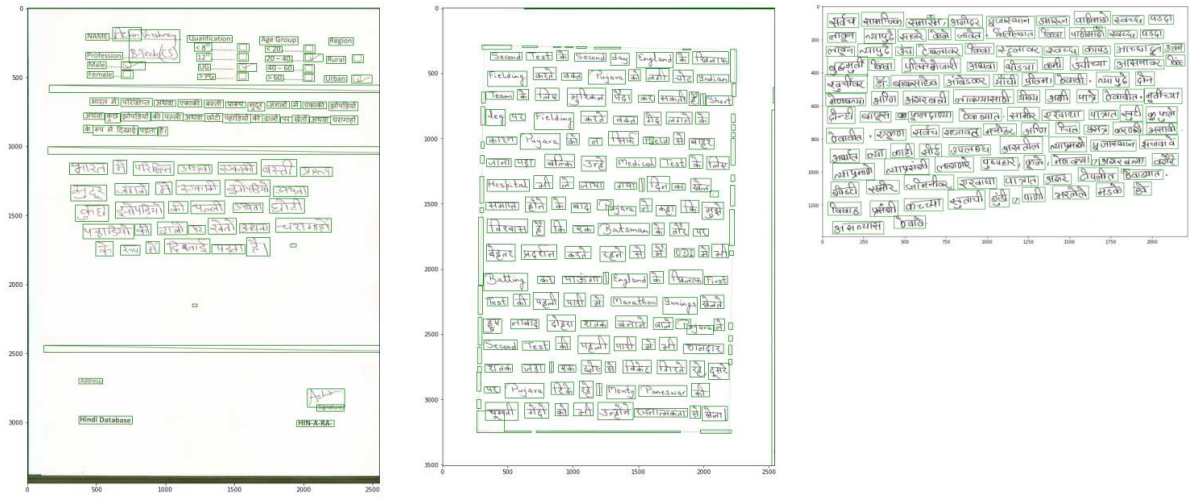


Fig. 4. Process of proposed work (a) Input image (b) Gray scale conversion (c) Binarization (d) Opening morphological operation (e) Connected component processing (f) Dilation (g) Find contours

Authorized licensed use limited to: Sardar Patel Institute of Technology. Downloaded on April 06,2024 at 06:36:39 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c)

Fig. 5. Qualitative results of proposed method on datasets as: (a) CALAM handwritten text dataset (b) Cmaterdb dataset 1.5.1 (c) Devanagari script of PHDIndic_11

VI. Conclusion and Future Scope

The work is presented for text detection and segmentation in Devanagari handwritten document images using image processing and morphological operations. Although the performance of the proposed method is good for all three datasets, it has some limitations. This work assumes that the sample is written on the plan sheet. When the horizontal lines are drawn before writing or when handwriting is done with less space between words, then the proposed method performs poorly for handwritten word detection and localization.

This work also limits the performance if the outline is drawn from handwritten samples, as in the Cmaterdb dataset version 1.5.1. Although the work is added in preprocessing steps to remove outlines, some portion of the line remains in samples for further steps due to the non-horizontal writing style, resulting in some false detection.

In the future, work to avoid false detection may be explored, and word recognition work may be added to the segmented words. Further restructuring work can also be framed so that the document image looks like an editable document file.

References

[1] S. Mahajan and R. Rani, "Text detection and localization in scene images: a broad review," Artificial Intelligence Review, vol. 54, no. 6, pp. 4317–4377, 2021.

[2] M. Bisht and R. Gupta, "Fine-tuned pre-trained model for script recognition," International Journal of Mathematical, Engineering and Management Sciences, vol. 6, no. 5, pp. 1297-1314, 2021.

[3] M. Bisht and R. Gupta, "Multiclass recognition of offline handwritten Devanagari characters using CNN," International Journal of Mathematical, Engineering and Management Sciences, vol. 5, pp. 1429–1439, 2020.

[4] M. Bisht and R. Gupta, "Offline handwritten Devanagari modified character recognition using convolutional neural network," Sādhanā, vol. 46, no. 1, pp. 1–4, 2021.

[5] A. Kaur, R. Dhir, and G. S. Lehal, "A survey on camera-captured scene text detection and extraction: towards Gurmukhi script," International Journal of Multimedia Information Retrieval, vol. 6, no. 2, pp. 115–142, 2017.

[6] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 2963–2970.

[7] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images," in Asian conference on computer vision, 2010, pp. 770–783.

[8] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images," in 2012 IEEE conference on computer vision and pattern recognition, 2012, pp. 1083–1090.

[9] W. Huang, Z. Lin, J. Yang, and J. Wang, "Text localization in natural images using stroke feature transform and text covariance descriptors," in Proceedings of the IEEE international conference on computer vision, 2013, pp. 1241–1248.

[10] L. Neumann and J. Matas, "Real-time scene text localization and recognition," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 3538–3545.

[11] X.-C. Yin, X. Yin, K. Huang, and H.-W. Hao, "Robust text detection in natural scene images," IEEE transactions on pattern analysis and machine intelligence, vol. 36, no. 5, pp. 970–983, 2013.

[12] A. Coates et al., "Text detection and character recognition in scene images with unsupervised feature learning," in 2011 International Conference on Document Analysis and Recognition, 2011, pp. 440–445.

[13] M. Jaderberg, A. Vedaldi, and A. Zisserman, "Deep features for text spotting, 2014," in European Conference on Computer Vision (ECCV), vol. 1, no. 2, pp. 512-528.

[14] W. Huang, Y. Qiao, and X. Tang, "Robust scene text detection with convolution neural network induced mser trees," in European conference on computer vision, 2014, pp. 497–511.

[15] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Reading text in the wild with convolutional neural networks," International journal of computer vision, vol. 116, no. 1, pp. 1–20, 2016.

[16] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic data for text localisation in natural images," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2315–2324.

[17] Z. Zhang, C. Zhang, W. Shen, C. Yao, W. Liu, and X. Bai, "Multi-oriented text detection with fully convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 4159–4167.

[18] X. Zhou et al., "East: an efficient and accurate scene text detector," in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2017, pp. 5551–5560.

[19] Z. Tian, W. Huang, T. He, P. He, and Y. Qiao, "Detecting text in natural image with connectionist text proposal network," in European conference on computer vision, 2016, pp. 56–72.

Authorized licensed use limited to: Sardar Patel Institute of Technology. Downloaded on April 06,2024 at 06:36:39 UTC from IEEE Xplore. Restrictions apply.

[20] Z. Zhang, W. Shen, C. Yao, and X. Bai, "Symmetry-based text line detection in natural scenes," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 2558–2567.
[21] M. Busta, L. Neumann, and J. Matas, "Fastext: Efficient unconstrained scene text detector," in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1206–1214.
[22] D. Dhar et al., "Multilingual scene text detection using gradient morphology," International Journal of Computer Vision and Image Processing (IJCVIP), vol. 10, no. 3, pp. 31–43, 2020.
[23] S. Long, X. He, and C. Yao, "Scene text detection and recognition: The deep learning era," International Journal of Computer Vision, vol. 129, no. 1, pp. 161–184, 2021.
[24] T. Khan, R. Sarkar, and A. F. Mollah, "Deep learning approaches to scene text detection: a comprehensive review," Artificial Intelligence Review, vol. 54, no. 5, pp. 3239–3298, 2021.
[25] Y. Xu, Y. Wang, W. Zhou, Y. Wang, Z. Yang, and X. Bai, "Textfield: Learning a deep direction field for irregular scene text detection," IEEE Transactions on Image Processing, vol. 28, no. 11, pp. 5566–5579, 2019.
[26] U. Bhattacharya, S. K. Parui, and S. Mondal, "Devanagari and bangla text extraction from natural scene images," in 2009 10th International Conference on Document Analysis and Recognition, 2009, pp. 171–175.
[27] H. Raj and R. Ghosh, "Devanagari text extraction from natural scene images," in 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2014, pp. 513–517.
[28] R. Rahul, S. Bhaskaran, J. Amudha, and D. Gupta, "Multilingual text detection and identification from Indian signage boards," in 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 1120–1125.
[29] R. Pramanik and S. Bag, "Segmentation-based recognition system for handwritten Bangla and Devanagari words using conventional classification and transfer learning," IET Image Processing, vol. 14, no. 5, pp. 959–972, 2020.
[30] P. K. Singh, R. Sarkar, N. Das, S. Basu, M. Kundu, and M. Nasipuri, "Benchmark databases of handwritten Bangla-Roman and Devanagari-Roman mixed-script document images," Multimedia Tools and Applications, vol. 77, no. 7, pp. 8441–8473, 2018.

Authorized licensed use limited to: Sardar Patel Institute of Technology. Downloaded on April 06,2024 at 06:36:39 UTC from IEEE Xplore. Restrictions apply.