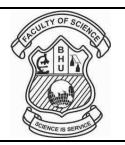


#### Volume 66, Issue 2, 2022

## **Journal of Scientific Research**

of

## The Banaras Hindu University



# Segmentation of Lines and Words of Handwritten Devanagari Text using Connected Components with Statistics Method

 $Viiav\ More^{1[0000\text{-}0002\text{-}2037\text{-}7134]}\ and\ Madan\ Kharat^{2[0000\text{-}0001\text{-}8862\text{-}9775]}$ 

<sup>1</sup>MET's Institute of Engineering, Bhujbal Knowledge City, Nashik, affiliated to Savitribai Phule Pune University, Pune. vbmore2005@rediffmail.com

<sup>2</sup>Dept. of Computer Engineering, MET's Institute of Engineering, Bhujbal Knowledge City, Nashik. mukharat@rediffmail.com

Abstract. The pre-processing activities for handwritten Devanagari text recognition includes an significant step called Segmentation. The segmentation accuracy of Devanagari text characters depends entirely on the accurately segmented lines and words in the handwritten documents. The process of segmenting lines and words correctly leads to many issues. More detailed information is lagging on the segmentation of lines and words from Devanagari text documents, whereas it is available more for other script documents in the literature. Here, we accomplished the task of segmenting the lines and words using Connected Components with Statistics Method on PHDIndic\_11 dataset. Experimentation using above mentioned method resulted in line segmentation accuracy of 91.91% and word segmentation accuracy of 72.89% which outperforms over Global threshold and Otsu's optimum threshold methods.

Keywords: Connected component, Devanagari, preprocessing, segmentation, statistics.

#### 1. Introduction

Segmentation is one of the most important and basic tasks of preprocessing during image processing and hence is a basic task in recognition of Devanagari text. The segment is usually a single character or even a part of the character as we are interested to find the pattern of the character.

Handwritten Devanagari text documents are preprocessed first and then recognition is carried out after extracting useful features. Devanagari handwritten document script extraction technique is used in many application domains.

The handwritten Devanagari text segmentation process includes several issues and associated challenges as well. As a result, correct recognition helps in recognizing the correct character.

#### 1.1 Issues and Challenges

**Issues.** In Devanagari script, we have connected and composite characters which is a mojor issue as shown in Figure 1. As we are concentrating on handwritten characters, overlapping of lines may maximize the complexity of segmentation as shown in Figure 2. Third and foremost

complexity arises due to ignorance caused due to 'Anuswara' which is a point appearing at the top of a character, which may be non-connectivity in nature as in Figure 3. Fourth complexity also arise due to 'Ardha Chandra' as shown in Figure 4 a half-moon-like character, and various alike characters in Devanagari text.

Challenges. Apart from the issues discussed above, increased complexity of segmentation due to large amount of variations in writing style shown in Figure 5 is a challenge. Second challenge is due to overlapped hand written lines of text as shown in Figure 2 may confuse segmentation process and leads to issues in identifying character boundaries correctly. Degraded historical handwritten document segmentation is another great challenge. Limited work in the domain of segmentation on Devanagari text is another added challenge which is also a scope for research.

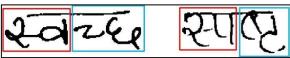


Fig. 1. Composite Characters

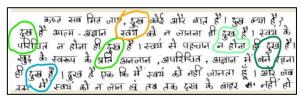


Fig. 2. Overlapped Handwritten Lines of Text

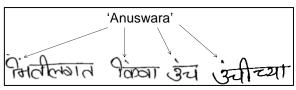


Fig. 3. 'Anuswara' Characters

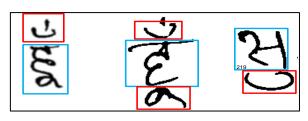


Fig. 4. 'Ardha Chandra' (half moon) Characters



Fig. 5. Writing Style Variations

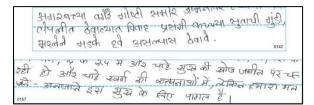


Fig. 6. Sloped and Curved Lines

# 2. Literature Survey

Babu, Subith et al. [1] converted handwritten image components of in the form of graph with end and joint points as vertices of polygon curve. By representing a handwritten image as a graph and processing with graph theory, author found 85% accuracy in segmentation over 250 handwritten character images. The limitation of this methodology was lies in variations in handwriting documents. They found that, expected results will hardly meet expectation in case of variations in handwriting documents.

Bhattad, Anmol J et al. [2] used fuzzy logic membership function to detect segmentation. Author applies segmentation of characters by using first non-black pixel location for each column, and considered various other parameters like thickness, count, run length count of non-black pixels. Author claimed 93.39% accuracy Devanagari characters segmentation.

Thakral, B. et al. [3] experimented for conjunct and overlapped characters in which the author uses the Cluster Detection technique. Header lines are removed from the word using horizontal projection profile, the gap between characters is found then located the pixel cluster for cluster separation. The author claimed 95% accuracy of touching and conjunct characters segmentation and found 88% segmentation accuracy for overlapping characters too.

Bhujade, Vaishali G. et al. [4] experimented with Devanagari handwritten characters segmentation. The author used the projection method for the detection of segmentation. Though primitive operations are described very well by authors in their paper, result of the accuracy percent obtained is not mentioned.

Kapoor, Shuchi et al. [5] identified joint points and formed bounding boxes covering the identified joint points and then applied segmentation to touching characters based on their height and width. Average segmentation accuracy they found to be 75.75% for Devanagari text.

Palakollu, Saiprakash et al. [6] worked to straighten the header line, and then the upper modifier is separated, along with separation of consonant and lower modifier. The author used the horizontal projection profile technique for segmentation of words and segmentation accuracy claimed is 89.90%.

Garg, Naresh Kumar et al. [7] use structural properties of text to segment the half characters in handwritten Hindi text. An author claiming the achievement of the segmentation accuracy as 83.02% for half characters in handwritten text.

Ladwani, Vandana M. et al. [8] uses morphological operations for segmentation. A neighborhood tracing algorithm was used by the author for segmentation. Segmentation accuracy was found to be 54.66%.

Garg, Naresh Kumar et al. [9] segmentation technique used was based on structure approach for handwritten Hindi text. Vertical projection profile i.e. the histogram of an input image is used, where the zero valley peaks indicate space between words and characters. Accuracy claimed by the author is 79.12% on the segmentation of characters.

Sarkar, Ram, et al. [10] feature-based approach to identify Matra pixels from a word image, and the value so obtained is normalized to the maximum longest run value of any pixel within the word image. The author proposed bell-shaped membership functions to map the horizontalness feature values of each row to determine its belongingness in the Matra region. The segmentation accuracy achieved is 94.8%.

Garg, Naresh Kumar et al. [11] author worked on the method which detects header line, detects baseline, and follows the contour. Skew correction, thinning, or noise removal was not done on the data. The method is applicable only for line segmentation. Line segmentation accuracy found by the author is 97% on handwritten Gurmukhi script. Since character segmentation is not addressed, the paper may help in limited scope to the character segmentation domain.

Otsu Nobuyuki[12] proposed a new terminology in segmentation that he suggested using the optimum threshold selection method using Gray-Level Histograms. He suggested automatic threshold selection as an unsupervised method for image segmentation. An optimal threshold is picked up using the non-parametric method by the discriminant criterion from the gray-level histogram.

Sk Md Obaidullah et al. [13] provided PHDIndic\_11: page-level handwritten document image dataset of 11 official Indic scripts for script identification. PHDIndic\_11 is a page-level handwritten document image dataset of 11 official Indic scripts for script identification. These scripts are Bangla, Devanagari, Roman, Urdu, Oriya, Gurumukhi, Gujarati, Tamil, Telugu, Malayalam, and Kannada. Devanagari scripts of 220 different documents written by different writers from this data set are considered for the proposed work.

More Vijay et al. [14] studied and reviewed various techniques of handwritten Devanagari text recognition. The author performed a comparative study of various techniques which can be extended further by the researcher community.

Vijay More et al.[15] observed that the segmentation accuracy of Devanagari text characters depends on the segmented lines and words from handwritten documents. The author identified many issues and challenges for segmenting lines and words from these handwritten documents. Global threshold and Otsu's optimum threshold methods were experimented by the author and found 85.12% segmentation accuracy.

## 3. Proposed Approach

In most research cases, the segmentation accuracy for a single component (line, word, character) was observed very low. Instead, researchers mentioned aggregate performance. This work is experimented separately on line and word segmentation. Their results are mentioned separately for a clear understanding of the results obtained. These results are compared with other researchers' results. In this work, experimentation is done using connected component analysis as an improved method for segmentation of lines and words specifically and their results obtained are analyzed with the results obtained by the author [15].

#### 3.1 Connected Component Analysis

Connected components scan an image and group a set of pixels into components based on pixel connectivity. All pixels in a connected component having similar pixel intensity values and are connected. Once all groups have been determined, each pixel is labeled with a gray-level or a color (color labeling) according to the component it was assigned to. Pixel connectivity describes a relation between two or more pixels. For two pixels to be connected they have to fulfill certain conditions on the pixel brightness and spatial adjacency.

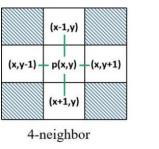
For two connected pixels, their pixel values must belong to same set of values V.

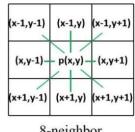
For a grayscale image,  $V=\{22,23,...40\}$ , range of graylevels For a binary image,  $V=\{1\}$ , either 1 or 0.

#### 3.2 Adjacency Criteria

For a pixel p with the coordinates (x,y) the set of adjacency pixels are calculated by using 4-neighbor or 8-neighbor formula:

$$\begin{split} &N_4(p) {=} \{(x{+}1{,}y){,}(x{-}1{,}y){,}(x{,}y{+}1){,}(x{,}y{-}1)\} \\ &N_8(p) {=} N_4(p) \stackrel{.}{E} \left\{(x{+}1{,}y{+}1){,}(x{+}1{,}y{-}1){,}(x{-}1{,}y{+}1){,}(x{-}1{,}y{-}1)\right\} \end{split}$$





neighbor 8-neighbor
(a) (b)

Fig. 7. Adjacency Criteria a) 4-neighbor, b) 8-neighbor

#### 3.3 Statistics includes

- The leftmost (x) coordinate which is the inclusive start of the bounding box in the horizontal direction.
- The topmost (y) coordinate which is the inclusive start of

- the bounding box in the vertical direction.
- The horizontal size of the bounding box.
- The vertical size of the bounding box.
- The total area (in pixels) of the connected component.

#### 4. Dataset

Dataset used for implementation of the proposed method is PHDIndic\_11 [13]. This dataset consists of 220 images of handwritten documents in Devanagari script. This dataset is used by author [15] in their implementation. These handwritten documents are written by different writers. Example handwritten image document is given in figure 8.

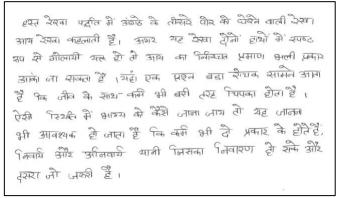


Fig. 8. Sample Handwritten Image Document

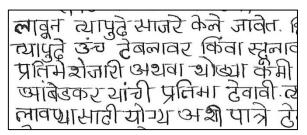
These handwritten image documents are 96 dpi with 8 bit depth. Author [15] categorized these documents in seven categories as partial header lines, congested lines, slope documents, no header lines, sloped with congested lines documents, normal image documents and slope with no header lines documents. Results obtained by author [15] is also categorized according to these different categories.

Experimentation is carried out on these documents and obtained results are categorized according to these document categories.

#### 5. Results

Document category wise segmentation analysis

a) Congested lines. There are 27 such documents identified which are categories as congested lines. Sample of congested line document is p\_dev\_0074.tif and is shown in figure 9.



**Fig. 9.** Congested line sample document [p\_dev\_0074,tif]

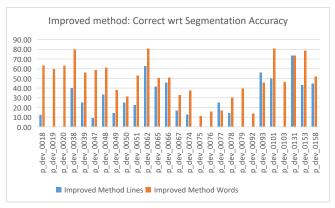


Fig. 10. Correct with respect to segmentation accuracy

Figure 10 is the correct with respect to segmentation accuracy, i.e. correctly segmented lines or words with respect to how many number of line and words segmented, which is found 24.68% for line segmentation and 48.83% for word segmentation.

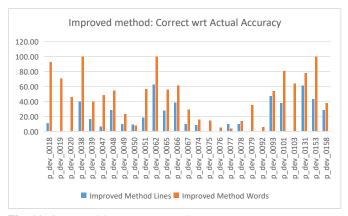


Fig. 11. Correct with respect to actual accuracy

Figure 11 is the correct with respect to actual accuracy, i.e. correctly segmented lines or words with respect to how many number of line and words actual available, which is found 19.42% for line segmentation and 47.92% for word segmentation.

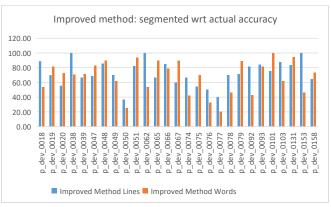


Fig. 12. Segmented with respect to actual accuracy

Figure 12 is the Segmented with respect to actual accuracy, i.e. number of lines and words segmented with respect to how many number of line and words actual available, which is found 72.71% for line segmentation and 67.06% for word segmentation.

**b)** No-header lines. There are 6 documents categories as no-header lines. Sample of no-header line document is p\_dev\_0176.tif and is shown in figure 13.

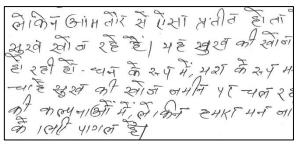


Fig. 13. No-header line sample document [p\_dev\_0176.tif]

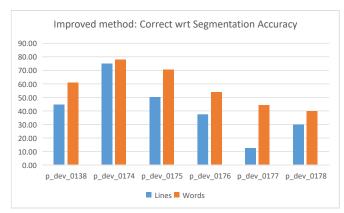


Fig. 14. Correct with respect to segmentation accuracy

Representation of figure 14 is the correct with respect to segmentation accuracy, i.e. correctly segmented lines or words with respect to how many number of line and words segmented, which is found 41.57% for line segmentation and 57.83% for word segmentation.

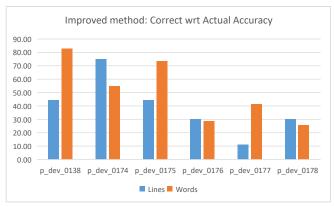


Fig. 15. Correct with respect to actual accuracy

Figure 15 is the correct with respect to actual accuracy, i.e. correctly segmented lines or words with respect to how many number of line and words actual available, which is found 39.17% for line segmentation and 51.07% for word segmentation.

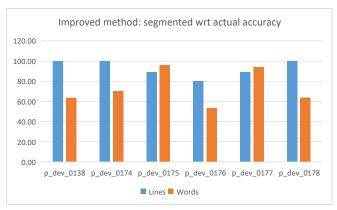


Fig. 16. Segmented with respect to actual accuracy

Figure 16 is the Segmented with respect to actual accuracy, i.e. number of lines and words segmented with respect to how many number of line and words actual available, which is found 92.96% for line segmentation and 73.41% for word segmentation.

c) Partial header lines. There are 9 such documents identified which are categories as partial header lines. Sample of partial header line document is p\_dev\_0008.tif and is shown in figure 17.

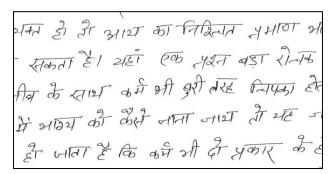


Fig. 17. Partial header line sample document [p\_dev\_0008.tif]

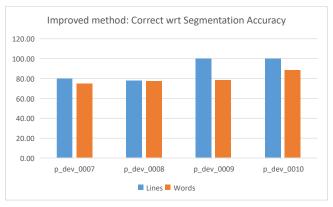


Fig. 18. Correct with respect to Segmentation Accuracy

Representation of figure 18 is the correct with respect to segmentation accuracy which is found 89.44% for line segmentation and 79.72% for word segmentation.

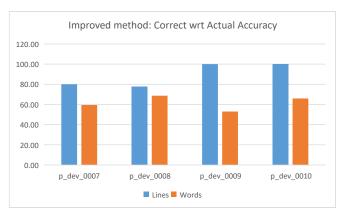


Fig. 19. Correct with respect to Actual Accuracy

Figure 19 is the correct with respect to actual accuracy which is found 89.44% for line segmentation and 61.77% for word segmentation.

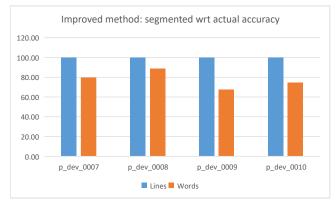


Fig. 20. Segmented with respect to actual accuracy: Lines , Words

Figure 20 is the Segmented with respect to actual accuracy which is found 100.00% for line segmentation and 77.65% for word segmentation.

**d) Slant lines.** There are about 37 handwritten documents observed and categorized as slant documents. Sample slant document is p\_dev\_0096.tif and is shown in figure 21.

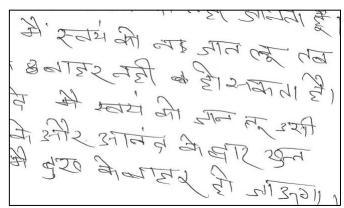


Fig. 21. Slant line sample document [p\_dev\_0096.tif]

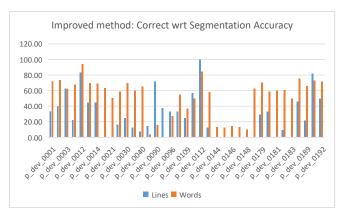


Fig. 22. Correct with respect to Segmentation Accuracy

Figure 22 is the correct with respect to segmentation accuracy which is found 28.31% for line segmentation and 51.84% for word segmentation.

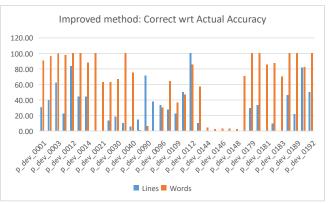


Fig. 23. Correct with respect to Actual Accuracy

Figure 23 is the correct with respect to actual accuracy which is found 27.35% for line segmentation and 64.26% for word segmentation.

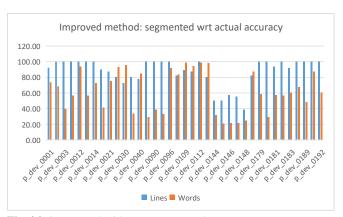


Fig. 24. Segmented with respect to actual accuracy

Figure 24 is the Segmented with respect to actual accuracy which is found 87.47% for line segmentation and 61.69% for word segmentation.

e) Slant with congested lines. There are about 9 handwritten documents observed and categorized as slant with congested lines. Sample Slant with congested line document is p\_dev\_0042.tif and is shown in figure 25.

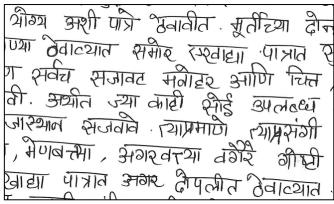


Fig. 25. Slant with congested line sample document [p\_dev\_0042.tif]

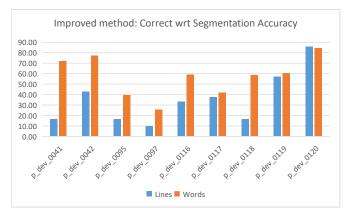


Fig. 26. Correct with respect to Segmentation Accuracy

Figure 26 is the correct with respect to segmentation accuracy which is found 35.17% for line segmentation and 57.63% for word segmentation.

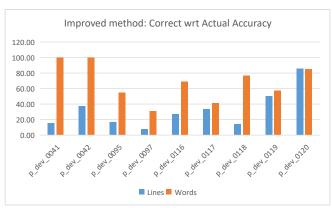


Fig. 27. Correct with respect to Actual Accuracy

Figure 27 is the correct with respect to actual accuracy which is found 31.98% for line segmentation and 68.29% for word segmentation.

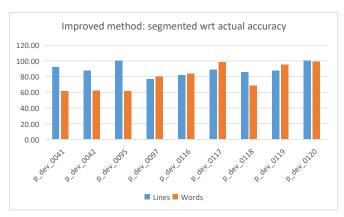


Fig. 28. Segmented with respect to actual accuracy

Figure 28 is the Segmented with respect to actual accuracy which is found 88.96% for line segmentation and 78.88% for word segmentation.

**f) Slant with no-header lines.** There are about 20 handwritten image documents observed which are categorized as slant with no-header lines. Sample slant with no-header lines document is p\_dev\_0214.tif and is shown in figure 29.

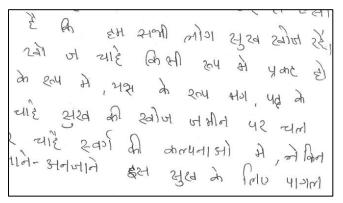


Fig. 29. Slant with no-header line sample document [p\_dev\_0214.tif]

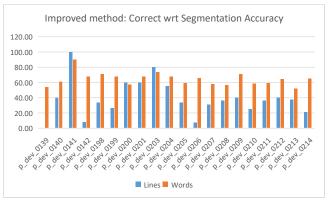


Fig. 30. Correct with respect to Segmentation Accuracy: Lines , Words

Figure 30 is the correct with respect to segmentation accuracy which is found 38.61% for line segmentation and 64.36% for word segmentation.

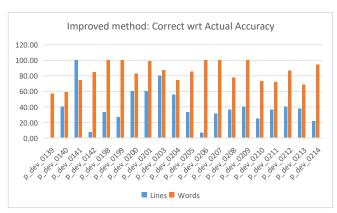


Fig. 31. Correct with respect to Actual Accuracy

Figure 31 is the correct with respect to actual accuracy which is found 38.53% for line segmentation and 83.85% for word segmentation.

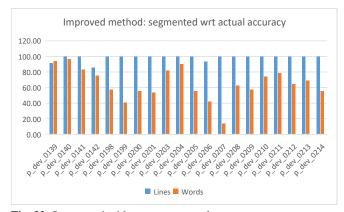


Fig. 32. Segmented with respect to actual accuracy

Figure 32 is the Segmented with respect to actual accuracy which is found 98.54% for line segmentation and 65.04% for word segmentation.

**g) Normal documents.** There are about 117 such handwritten image documents observed as normal documents. Sample normal document is p\_dev\_0217.tif and is shown in figure 33.

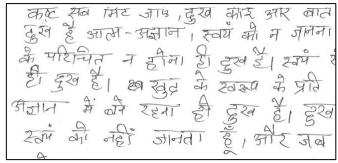


Fig. 33. Sample normal document [p\_dev\_0217.tif]

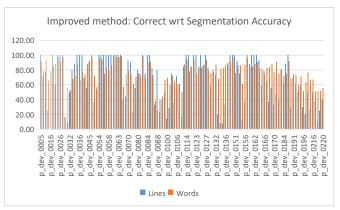


Fig. 34. Correct with respect to Segmentation Accuracy

Figure 34 is the correct with respect to segmentation accuracy which is found 68.02% for line segmentation and 76.34% for word segmentation.

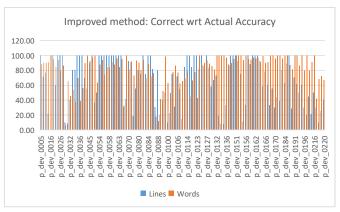


Fig. 35. Correct with respect to Actual Accuracy

Figure 35 is the correct with respect to actual accuracy which is found 66.89% for line segmentation and 82.00% for word segmentation.

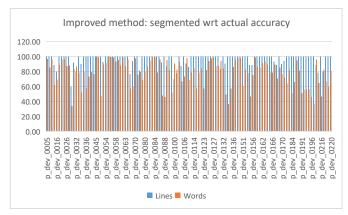


Fig. 36. Segmented with respect to actual accuracy: Lines, Words

Figure 36 is the Segmented with respect to actual accuracy which is found 96.51% for line segmentation and 78.41% for word segmentation.

h) Overall documents. Collection of all category documents is overall documents of the dataset PHDIndic\_11 [13] and which are 220 nos and whose samples are already provided in above results discussion.

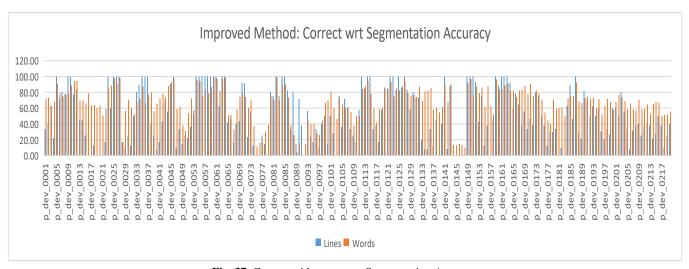


Fig. 37. Correct with respect to Segmentation Accuracy

Figure 37 is the correct with respect to segmentation accuracy which is found 51.67% for line segmentation and 66.54% for word segmentation.

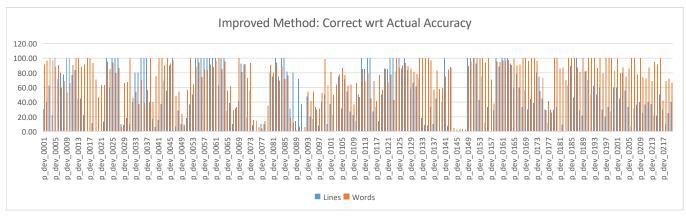


Fig. 38. Correct with respect to Actual Accuracy

Figure 38 is the correct with respect to actual accuracy which is found 50.06% for line segmentation and 73.23% for word segmentation.

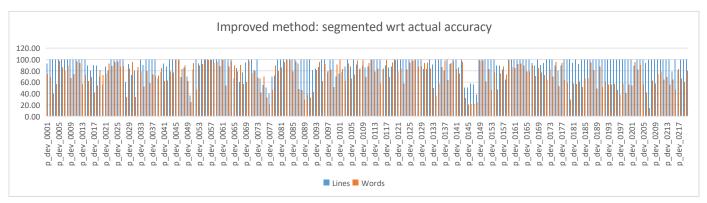


Fig. 39. Segmented with respect to actual accuracy

Figure 39 is the Segmented with respect to actual accuracy which is found 91.91% for line segmentation and 72.89% for word segmentation.

Accuracy comparison of lines and words between Global thresholding method, Otsu's thresholding method, and Connected Component with Statistics method as Improved method is shown in table (Table 1).

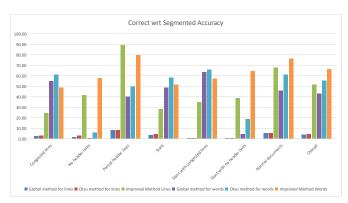
Table 1. Correct with respect to Segmented Accuracy

Document category	Global method for lines	Otsu method for lines	Improved Method Lines	Global method for words	Otsu method for words	Improved Method Words
Congested lines	2.70	3.29	24.68	55.13	61.21	48.83
No header lines	1.85	2.96	41.57	0.83	5.81	57.83
Partial header lines	8.57	8.22	89.44	40.10	49.75	79.72
Slant	3.54	4.41	28.31	48.68	58.35	51.84
Slant with congested lines	0.34	0.15	35.17	63.49	65.97	57.63
Slant with no header lines	0.21	0.21	38.61	4.29	18.86	64.36
Normal documents	5.56	5.56	68.02	46.18	61.03	76.34
Overall	4.12	4.36	51.67	43.25	55.26	66.54

It is seen from (Table 1) that, improved method performed better in all the document categories. If we observe the case of partial header lines document categories, it is found 89.44% segmentation accuracy which is much better than 8.57% and 8.22% accuracy obtained using Global and Otsu's method respectively by author [15]. In word segmentation also the Improved method outperforms more in case of Partial header lines documents in comparison with other category

#### documents.

The table values from (Table 1) are depicted in chart form in figure 40.



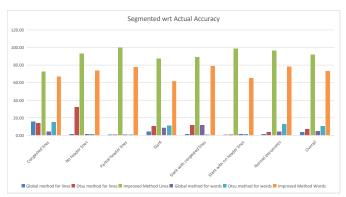
**Fig. 40.** Comparison chart of Connected Component Method vs Global and Otsu's Method Under Correct with respect to Segmented Accuracy

Correct with respect to segmented accuracy comparison (figure 40) shows that Connected Component Method performs better than Global and Otsu's method [15] in most of the document categories.

Table 2.	Segmented	with resi	pect to	Actual	Accuracy	7
----------	-----------	-----------	---------	--------	----------	---

Document category	method for	Otsu method for lines	Method		method for	Improved Method Words
Congested lines	15.45	14.13	72.71	4.52	15.02	67.06
No header lines	0.63	31.97	92.96	0.50	0.69	73.41
Partial header lines	0.31	0.31	100	1.00	1.00	77.65
Slant	4.30	10.64	87.47	8.69	10.97	61.69
Slant with congested lines	0.44	11.44	88.96	11.71	0.94	78.88
Slant with no header lines	0.15	0.15	98.54	0.61	0.61	65.04
Normal documents	1.56	3.88	96.51	4.08	13.08	78.41
Overall	3.51	6.95	91.91	4.75	10.78	72.89

The table values from (Table 2) are mentioned in chart (figure 41) for more aggregate clarity on results.



**Fig. 41.** Comparison Chart of Connected Component Method vs Global and Otsu's Method Under Segmented with respect to Actual Accuracy

Segmented with respect to actual comparison (see figure 41) shows that Connected Component Method outperforms too much better than Global and Otsu's method [15]. In all the document categories, the peak values of connected component method are very much higher than that of Global and Otsu's method. Line segmentation accuracy obtained is 91.91% which is too much better than that of using methods used by author [15] and word segmentation accuracy obtained is 72.89% using connected component method which also better than author [15] results.

#### 6. Conclusion

The experimental results of the connected component analysis method provide the segmentation of lines and words. In this work, we obtained line segmentation accuracy of 91.91% and word segmentation accuracy of 72.89%. Experimentation results show that the method outperforms on Global threshold [3.51% for lines and 4.75% for words] and Otsu's optimum threshold method [6.95% for lines and 10.78% for words] as per the results obtained in [15].

This work has more scope to extend to many dimensions like the segmentation of composite characters and overlapped lines. Though the work carried out by various researchers in these areas is appreciated, there is still room for improvement.

### 7. References

- [1] S. Babu and M. Jangid, "Touching character segmentation of Devanagari script", International Conference on Computing, Communication and Networking Technologies, ICCCNT, (2016), pp.1-5.
- [2] A. J. Bhattad and B. B. Chaudhuri, "An approach for character segmentation of handwritten Bangia and Devanagari script", IEEE International Advance Computing Conference (IACC), IEEE, (2015), pp.676-680.
- [3] B. Thakral and M. Kumar, "Devanagari handwritten text segmentation for overlapping and conjunct characters- a proficient technique", Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization, IEEE, (2014), pp.1-4.
- [4] V. G. Bhujade and C. M. Meshram, "A technique for segmentation of handwritten Hindi text", International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 2, February - 2014, IJERT, (2014), pp.1491-1495.
- [5] S. Kapoor and V. Verma, "Fragmentation of handwritten touching characters in Devanagari script", International Journal of Information Technology, Modeling and Computing (IJITMC) Vol. 2, No. 1, February 2014, IJITMC, (2014), pp.11-21.
- [6] S. Palakollu, R. Dhir, and R. Rani, "Handwritten Hindi text segmentation techniques for lines and characters", Proceedings of the World Congress on Engineering and Computer Science 2012, Vol I WCECS 2012, October 24-26, 2012, San Francisco, USA, IAENG-WCECS, (2012), pp.640-644.
- [7] N. K. Garg, L. Kaur, and M. K. Jindal, "The segmentation of half characters in handwritten Hindi text", International Conference on Information Systems for Indian Languages (ICISIL) 2011, Communications in Computer and Information Science, vol 139. Springer, Berlin, Heidelberg, Springer, (2011), pp.48-53.
- [8] V. M. Ladwani and L. Malik, "Novel approach to segmentation of handwritten Devnagari word", Third International Conference on Emerging Trends in Engineering and Technology, IEEE, (2010), pp.219-224.
- [9] N. K. Garg, L. Kaur, and M. K. Jindal, "Segmentation of handwritten Hindi text", International Journal of Computer Applications (0975 – 8887), Volume 1 – No. 4, IJCA, (2010), pp.19-23.
- [10] R. Sarkar, B. Sen, N. Das, and S. Basu, "Handwritten Devanagari script segmentation: a non-linear fuzzy approach", IEEE Conference on AI Tools and Engineering (ICAITE-08), Researchgate, (2008), pp.1-6.
- [11] N. K. Garg, L. Kaur, and M. K. Jindal, "A new method for line segmentation of handwritten Hindi text", Seventh International Conference on Information Technology, IEEE, (2010), pp.392-397.
- [12] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Transactions on Systems, MAN, And Cybernetics, Vol. SMC-9, No. 1, January 1979, IEEE, (1979), pp.62-66.
- [13] Sk. Md. Obaidullah, C. Halder, K. C. Santosh, N. Das, and K. Roy, "PHDIndic\_11:Page-level handwritten document image dataset of 11 official Indic scripts for script identification", Springer: Multimedia Tools and Applications January 2018, Volume 77, Issue 2, Springer, (2017), pp.1643-1678.
- [14] V. B. More, M. U. Kharat, and S. V. Gumaste, "Study of different features and classification techniques for recognition of handwritten Devanagari text", International Journal of Engineering & Technology (UAE-Scopus Indexed), 7 (4.19), IJET, (2018), pp.1055-1059.
- [15] Vijay More, Madan Kharat, Shyamrao Gumaste, "Segmentation of Devanagari Handwritten Text Using Thresholding Approach", International Journal of Scientific & Technology Research (IJSTR), Volume 9, Issue 3, 2277-8616, March (2020), pp.6594-6605.

\*\*\*