





correction can be achieved by (i) estimating the skew angle, and (ii) rotating the image by the skew angle in the opposite direction [3].

**C. Line Segmentation Algorithm:**

In this section we used a robust algorithm for segmentation of individual text lines in Odia printed documents. Both foreground and background information is used here for accurate line segmentation. The Algorithm takes input as Odia text printed documents. The printed document is in the form of image file, like jpg, tiff or png formats. It produces the output as text line segment of Odia file. The Algorithm first read the image and then converts it in to gray scaled image.

As we know the region between any two lines is white but the pixels present there are of different intensities. So here we assumed the pixel values ranging from 200 to 255 as white pixels. As the gray scaled images are stored in the memory in the form of matrices and cell values of the matrix varies from 0 to 255, where zero signifies as black and 255 as white. This Algorithm scans through the document and counts the white line between two successive text lines. Choose the candidate line as middle line of two successive text lines, and then replace that with black line.

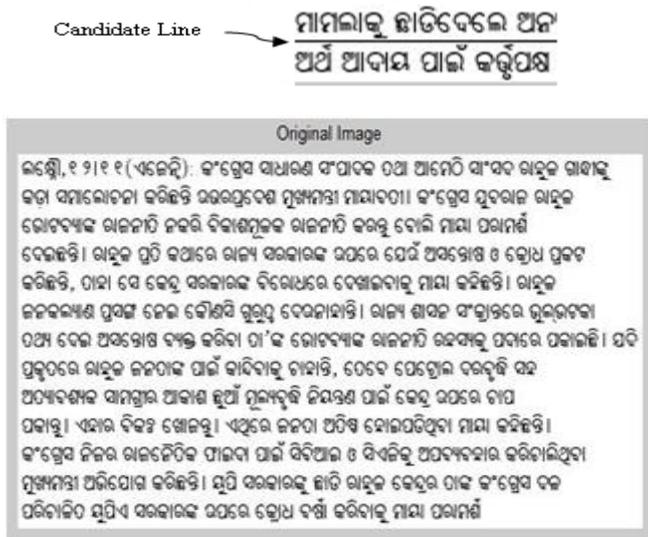


Figure 4.1 Printed Odia Document as Input

**Algorithm**

- Step 1: Read an Odia text document Image
- Step 2: If it is an RGB Image then convert it into gray scale image where the pixel values are from 0 to 255
- Step 3: Store the value of gray scale pixel in a two dimensional array I
- Step 4: For each row of the matrix I do
- Step 5: Count=0
- Step 6: For each row of j<sup>th</sup> element do
- Step 7: if I(i,j)>200 then
- Step 8: Increment count by one
- Step 9: end if
- Step 10: end for
- Step 11: if count value is equal to the sizeof row then
- Step 12: Increment line by one
- Step 13: Assign bline to zero
- Step 14: else if line>2 then
- Step 15: Increment bline by one
- Step 16: if bline==1 then

- Step 17: for each cell of the row do
- Step 18: Fill the black points at the middle of two Odia line
- Step 19: end for
- Step 20: Assign line to zero
- Step 21: end if
- Step 22: else
- Step 23: Increment bline by one
- Step 24: end if
- Step 25: end for

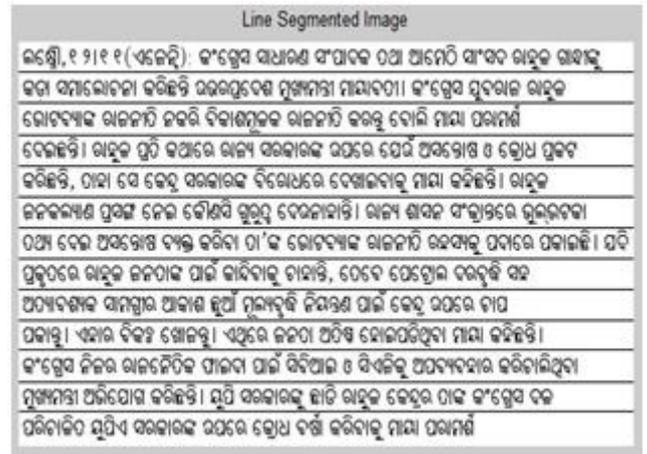


Figure 4.2 Text Line Segmented Image

**V. CONCLUSIONS**

In Optical Character Recognition (OCR), the text lines in a document must be segmented properly before recognition. Line segmentation is the preliminary and essential requirement of OCR systems. In this paper we segmented individual Odia text line accurately. The experiment can be further extended to character segmentation and then character recognition through feature extraction.

**VI. REFERENCES**

- [1] S Mohanty and H N Das Bebartha, "A Novel Approach for Bilingual (English - Odia) Script. Identification and Recognition in a Printed Document", International Journal of Image Processing (IJIP), Volume (4): Issue (2)
- [2] N Tripathy and U Pal, "Handwriting segmentation of unconstrained Odia text", S<sup>adha</sup>n a Vol. 31, Part 6, pp. 755-769, December 2006
- [3] Nallapareddy Priyanka, Srikanta Pal, Ranju Mandal, "Line and Word Segmentation Approach for Printed Documents", IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition", RTIPPR, 2010.
- [4] U. Pal and B.B. Chaudhuri, "Indian script character
- [5] Recognition: A Survey", Pattern Recognition, vol. 37, pp.1887-1899, 2004.
- [6] B. B. Chaudhuri and U. Pal, "A complete printed Bangla OCR system", Pattern Recognition, vol.31, pp.531-549, 1998.
- [7] Vijay Kumar, Pankaj K.Senegar, "Segmentation of Printed Text in Devnagari Script and Gurmukhi Script ", IJCA: International Journal of Computer Applications, Vol.3,pp. 24-29, 2010.

- [8] U. Pal and Sagarika Datta, "Segmentation of Bangla Unconstrained Handwritten Text", Proc. 7th Int. Conf. on Document Analysis and Recognition, pp.1128-1132, 2003.
- [9] K. Wong, R. Casey and F. Wahl "Document Analysis System ", IBM j.Res . Dev., 26(6), pp.647-656, 1982.
- [10] F. Hones and J. Litcher, "Layout extraction of mixed mode documents", Machine Vision Application, vol. 7, pp. 237–246, 1994.
- [11] G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis system for technical journals", Computer, vol. 25, pp. 10-22, 1992.
- [12] H. Yan, "Skew correction of document images using interline cross-correlation", CVGIP: Graph. Models Image Process, vol. 55, pp. 538–543, 1993.
- [13] G. Magy, Twenty years of Document Analysis in PAMI, IEEE Trans. In PAMI, Vol.22, pp. 38-61, 2000.