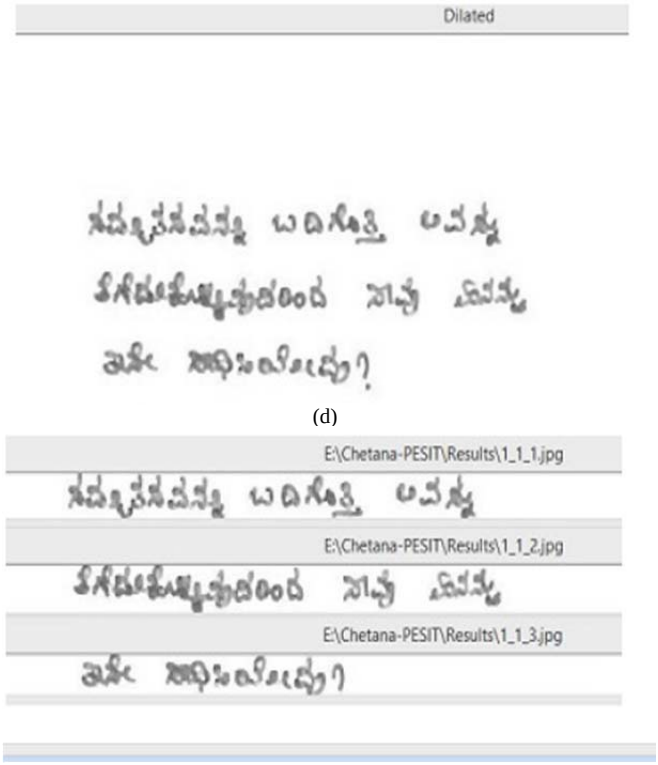
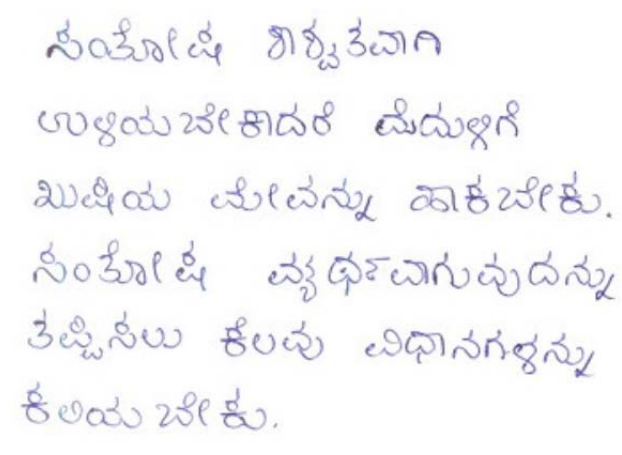


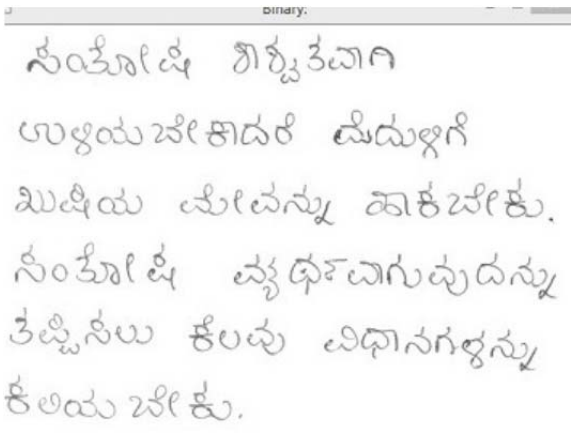
(d)

(e)

Fig. 7 Results of segmentation using morphology with Projection Profile (a)original document (b) binarised document (c) Eroded image (d) Dilated image (e)segmented lines



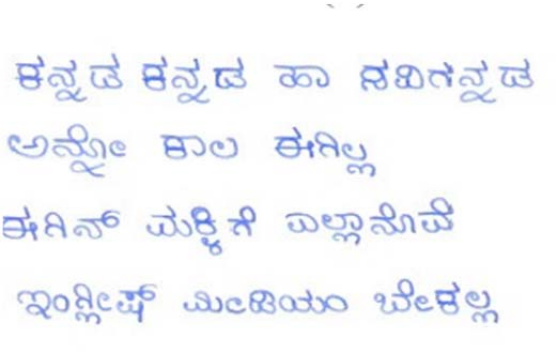
(a)



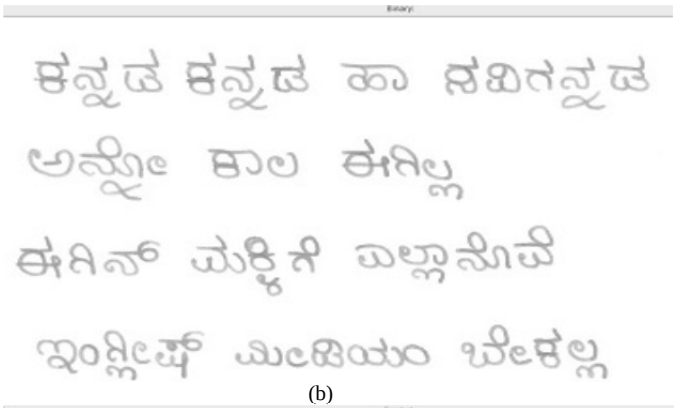
(b)

(c)

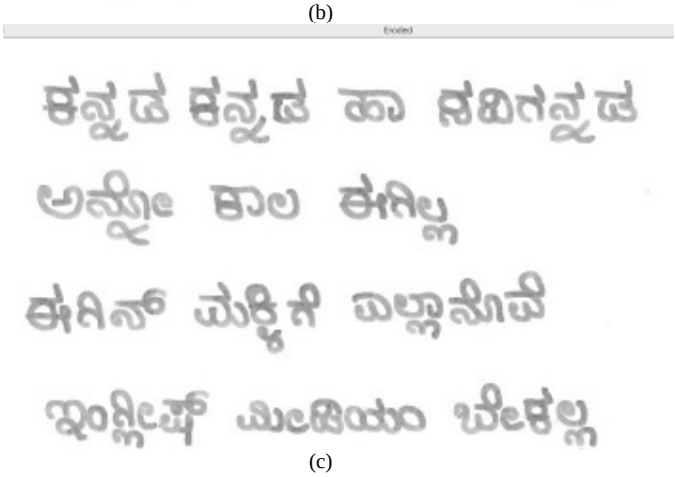
Fig. 8 Results of segmentation using Median Segmentation (a) original document(b) binarised document (c) segmented lines



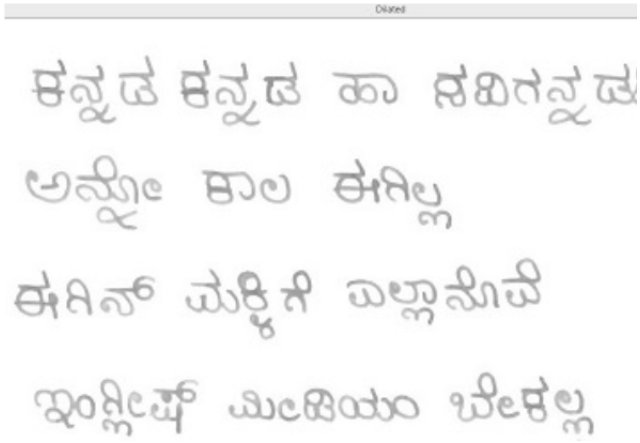
(a)



(b)



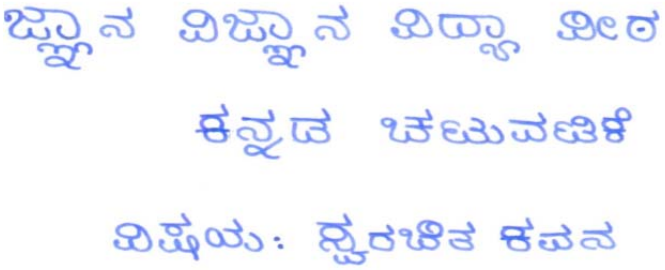
(c)



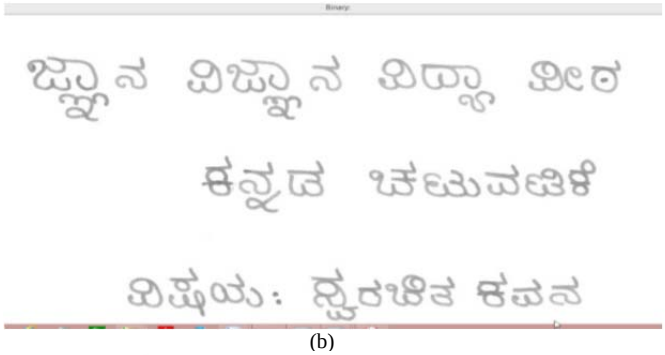
(d)

(e)

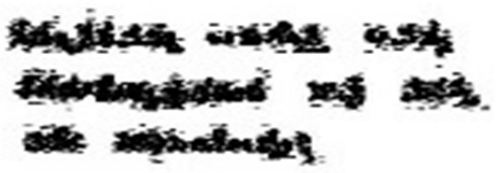
Fig. 9 Results of segmentation using morphology with Median Segmentation (a)original document (b) binarised document (c) Eroded image (d) Dilated image (e)segmented lines



(a)



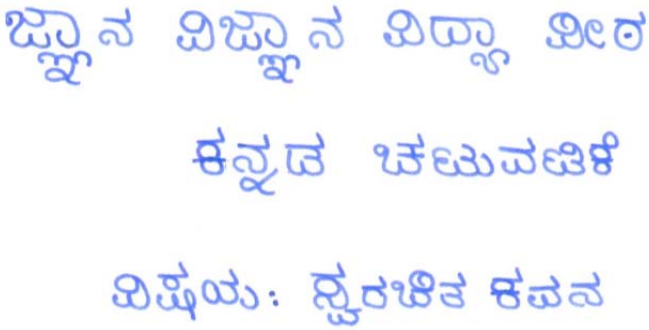
(b)



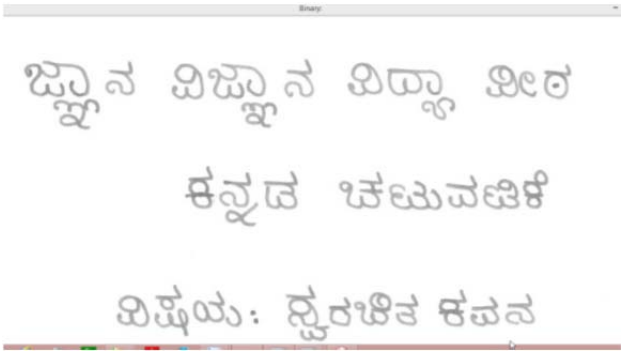
(c)

(d)

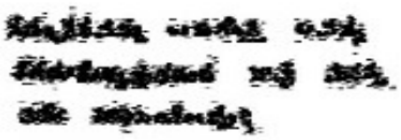
Fig. 10 Results of segmentation using morphology with Median Segmentation (a)original document (b) binarised document (c) Eroded image (d) Dilated image (e)segmented lines



(a)



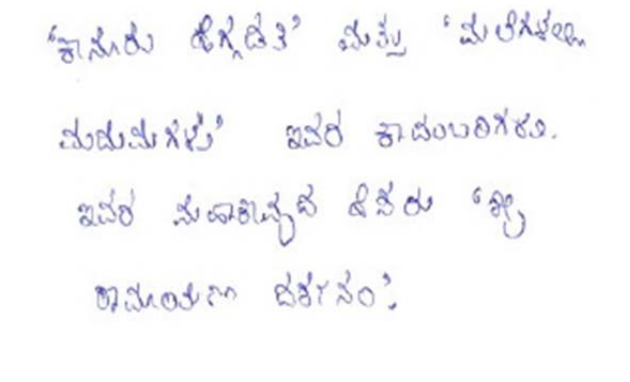
(b)



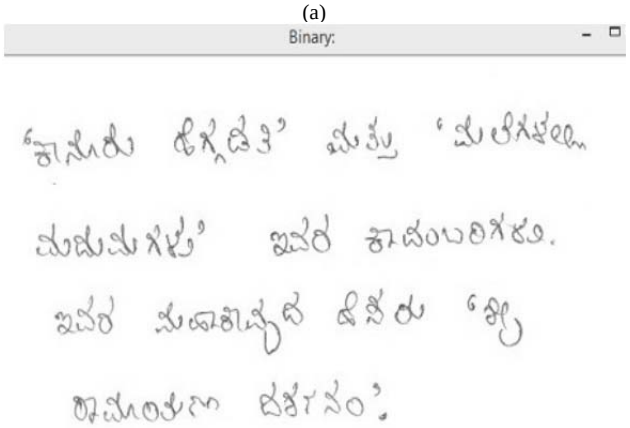
(c)

(d)

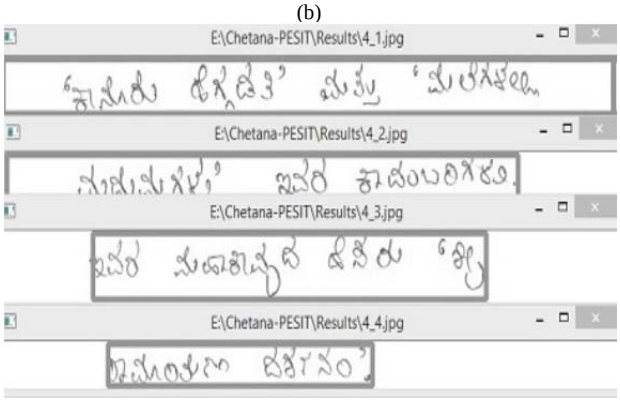
Fig. 11 Results of segmentation using Run length Smearing method (a) original document (b) binarised document (c) Smeared image (d) segmented lines



(a)

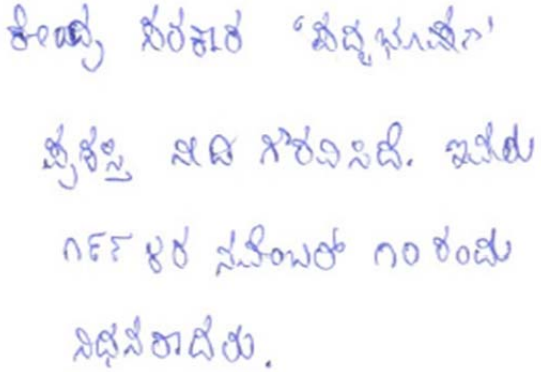


(b)

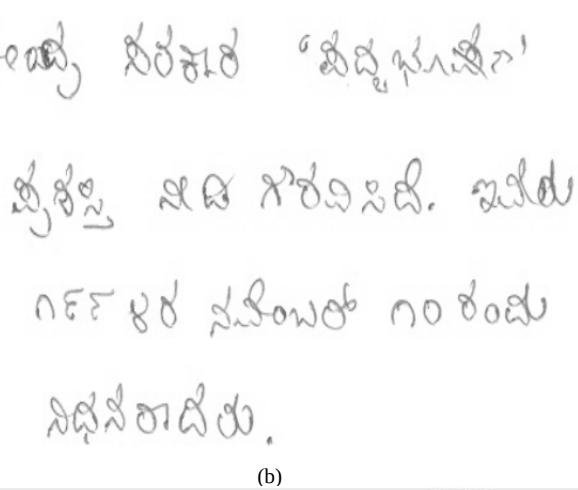


(c)

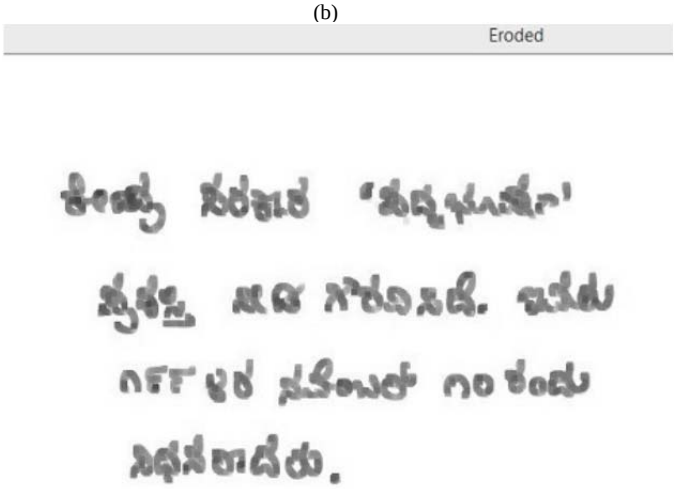
Fig. 12 Results of segmentation using Bounding Box method(a) original document(b) binarised document (c) segmented lines



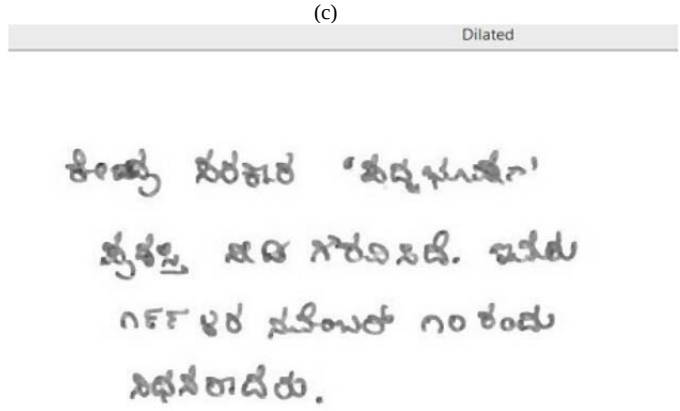
(a)



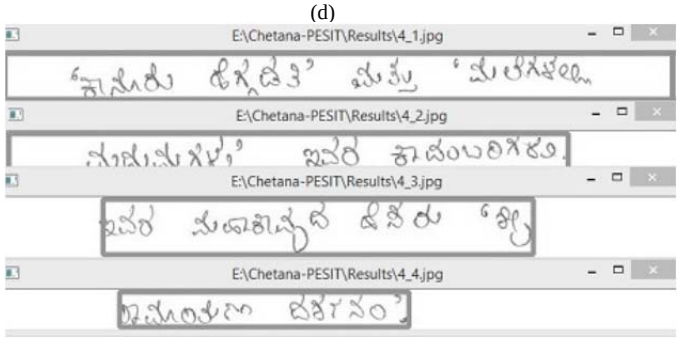
(b)



(c)



(d)



(e)

Fig. 13 Results of segmentation using morphology with Bounding Box method(a)original document (b) binarised document (c) Eroded image (d) Dilated image (e)segmented lines

V. CONCLUSIONS

Developing an OCR for handwritten Kannada documents is quite challenging and prone to errors due to its structural complexity and increased character set in Kannada language. An attempt is made in this direction and extraction of lines is done considering documents with different handwriting styles. But the accuracy obtained from the proposed methods are reduced because we have considered different documents with different handwriting styles. The accuracy for the documents with good handwriting style with less skew would have been much more higher than what we have obtained.

Four different methods like Projection profiles, Run Length smearing method, Median segmentation and Bounding box methods are proposed for text line extraction of Handwritten Kannada Documents. These proposed methods are experimented on two different datasets named as DS1 and DS2 collected from the authors of [4] and [9].

Morphological operations with projection profile gives the best segmentation rate of 93.87% among all other proposed methods because this method works well for clearly separated lines and this method cannot divide the touching or overlapping lines and instead it will merge those lines.

ACKNOWLEDGMENT

We would like to thank Nagabhushan P and Alireza Alaei authors of [9] for providing us with thier dataset for our experimentation.

REFERENCES

[1] Priyadharshini N & Vijaya MS, *Genetic Programming for Document Segmentation and Region Classification using Discipulus Perceptron,(IJARAI)* International Journal of Advanced Research in ArtificialIntelligence,Vol.2,No.2,2013.

[2] Rafael C. Gonzalez, Richard E. Woods and Steven L.Eddins ,*Digital Image Processing using MATLA*,. Indian Edition,2009,pp 348-361.

[3] Pulagam Soujanya, Vijaya Kumar Koppula , Kishore Gaddam and P. Sruthi, *Comparative Study of Text Line Segmentation Algorithms on Low Quality Documents*, Special Issue of International Journal of Computer Science & Informatics **(**IJCSI), ISSN **(**PRINT**)** :22315292,Vol.II,Issue1,2.

[4] Mamatha HR and Srikantamurthy K, *Morphological Operations and Projection Profiles based Segmentation of Handwritten Kannada Document*,International Journal of Applied Information Systems (IJAIS)–ISSN:2249-0868 Foundation of Computer Science FCS,2012.

[5] Laurence Likforman- Sulem , Abderrazak Zahour and Bruno Taconet**,** *Text line segmentation of historical documents: a survey, IJDAR(2007)* 9:123–138 DOI 10.1007/s10032-006-0023-z.

[6] Munish Kumar, R. K. Sharma and M. K. Jindal, *Segmentation of Lines and Words in Handwritten Gurumukhi Script Documents*, Indian Institute of Information Technology Allahabad, India.

[7] Vijaya Kumar Koppula and Atul Negi ,*Using Fringe Maps for Text Line Segmentation in Printed or Handwritten Document Images*,In the proceedings of 2010 Second Vaagdevi International Conference on Information Technology for Real World Problems,2010,pp8388.

[8] Mamatha H R and Srikantamurthy K, *Skew Detection, Correction and Segmentation of Handwritten Kannada Document* , International Journal of Advanced Science and Technology Vol. 48, November,2012.

[9] Nagabhushan P, Alireza Alaei and Umapada pal, *A Benchmark Kannada Handwritten Document Dataset and its Segmentation*,2011 International Conference on Document Analysis and Recognition.

[10] Laurence Likforman- Sulem and Ana hid Hanimyan , *A Hough Based Algorithm for Extracting Text Lines in Handwritten Documents*, Claudie Faure Ecole Nationale SupCrieure des T & communications, CNRS-URA 82046 rue Barrault,1995.

[11] M. Arivazhagan and H. Srinivasan, S. N. Srihari, *A Statistical Approach to Handwritten Line Segmentation*,In Proceedings of SPIE Document Recognition and Retrieval XIV,SanJose,CA,February2007.

[12] A.V. Aho, J.E. Hopcroft and J.D. Ullman, *Data Structures and Algorithms*, Addison- Wesley, 1983.

[13] A. Alaei, U. Pal and P. Nagabhushan, *A new scheme for unconstrained handwritten text-line segmentation*, Pattern Recognition,44(4),2011,pp.917–928.

[14] V. N. Manjunath Aradhya and C Naveena ,*Text Line Segmentation of Unconstrained Handwritten Kannada Script*, In the proceedings of ICCCS'11,2011,pp231-234.

[15] M.K Jindal, R. K. Sharma and G.S. Lehal. ,*Segmentation of Horizontally Overlapping Lines in Printed Indian Script*,.International Journal of Computational Intelligence Research. ISSN 0973-1873 Vol.3, No.4 (2007), pp. 277–286

[16] G. Louloudis, B. Gatos, I. Pratikakis, K.Halatsis, *A Block-Based Hough Transform Mapping for Text Line Detection in Handwritten Documents*, Proceedings of the Tenth International Workshop on Frontiers in Handwriting Recognition, La Baule, Oct. 2006.

[17] B.M.Sagar, Dr.Shobha G and Dr. Ramakanth kumar P, *OCR for printed kannada text to Machine editable format using Database approach*,9th WSEAS International Conference on AUTOMATION and INFORMATION(ICAI'08), Bucharest, Romania, June 24-26,2008.