

8 CONCLUSION

In this research, we have taken Gujarati script as base to study handwritten text recognition system for Indian languages specifically without head line and having diacritical mark involved in writing system. The segmentation of Gujarati text document into lines and further into characters is the major challenges of the system. Due to the nature of Gujarati characters shapes, large character set, and writing system of Gujarati text, we proposed segmentation method which segment text documents into core characters and diacritic marks.

The method is based on combination approach based on connected component and partial projection histogram. The knowledge of Gujarati construction is used to trace characters and generate text. The line segmentation rate achieved with the said approach is 98.32% and the character segmentation rate is 82.39%.

For recognition of core characters we have created trained model of isolated Gujarati character using neural network. To train the neural network, dataset consisting of 13500 images of isolated characters is used. We have identified feature set with only 90 elements based on water reservoir and radial histogram concept which gives recognition rate of 91.49% for 54 character class consisting of 49 basic alphabet and numerals and 5 part of the character. This indicates that, further study is required to improve the performance including adding more character class.

The character recognition result shows that the feature extraction method used in present work performed well. The lowest rate of recognition is 69.2% which is at higher side. Further combining these features with some other feature may increase rate of recognition significantly.

Future work:

We have restricted our self with segmentation, feature extraction and recognition stages. There are other stages like preprocessing and post processing which need to be added so that higher accuracy can be obtained. Following are few suggestion for future work:

- Conjunct characters are integral part of Gujarati writing system. We have experimented with half form of some of the character, which is giving good recognition rate. The same strategy can be implemented to recognize conjunct if segmented.
- Confusion of certain characters like “ Ya jə”, “ cha tʃə” can be reduced by may be adding new features.
- There is scope for improving overall text recognition accuracy.