



Comparative Analysis of Segmentation and Recognition Techniques for Offline Handwritten Words

Monika Kohli¹, Satish Kumar²

¹Research Scholar, Department of Computer Science and Applications, Panjab University, Chandigarh.

²Associate Professor, Department of Computer Applications, Panjab University, SSG Regional Centre Hoshiarpur, Punjab.

monikakrajotia@gmail.com¹

Abstract

A Pre-processing is the initial and vital phase in optical character recognition is the Pre-processing. Segmentation deals with the extraction of individual component from a document image. Number of techniques like projection profile, connected components, gaps between characters/components is reported in the literature for component extraction followed by feature extraction and recognition of the individual component. These techniques gives good results if components are isolated but fails if components are touched, shadowed or skewed. A novel technique is required to address such issues to enhance the recognition rate. The problem of segmentation for Roman script cursive handwriting is addressed by various authors but not enough addressed for Indian script especially Devanagari script. This paper is a review which is confined to offline handwritten script domain. It attempt to review various techniques for character segmentation considering touching characters for offline handwritten words in Devanagari script and scripts sharing similar characteristics (like Bangla, Gurumukhi), database used and their accuracy reported in the literature.

Keywords: Devanagari script, OCR (Optical Character Recognition) Segmentation, Touching characters

1. Introduction

OCR (Optical Character Recognition) is a conversion process which converts printed or handwritten data in the form of image, online or offline into machine encoded form. The purpose of converting data images into digital format is to edit and search data electronically, and store the digitized data in a compact way. ICR (Intelligent Character Recognition) more precise than OCR as different styles and fonts are made to learn by the computer system with major application as Automated Form processing. It has major advantages in term of speed, accuracy and cost. It reduces error as data entry (manually) is the

likelihood of typographical errors.

Devanagari script is widely used in northern and western part of India. There is more than 300 million user of the script and has various applications. Segmentation-based or holistic approached are used in literature for the recognition of Devanagari script. Recognition of number of languages is done using these approaches. Both approaches have shortcomings associated with them. However, Holistic approach does not give good results(Shaw, Parui, and Shridhar 2008) as per literature survey. Segmentation approach gives better results but Segmentation of Devanagari script is difficult because of presence of large character set which

include vowels, consonants, compound characters and modifiers. Poor segmentation contributes to recognition error. (Shaw 2008a)HMM has been used in recognizing handwritten words but reported with some success and that too with pre-segmented letters. According to literature survey, various techniques are found in number of research papers in offline handwritten character recognition in Latin and other Asian languages but a few papers are available in Devanagari script (Hindi). One of the reasons can be the non-availability of standard databases of handwritten text/words/characters. Large character set of a language poses another difficulty. Hindi, Marathi, Nepali, Konkani, Sindhi, Kashmiri etc. are various languages that belong to Devanagari script. Punjabi, Bengali, Marathi are the languages of other script that shares characteristics with Devanagari script.

This paper is divided into 8 sections. Section 2, 3 deals with the need of segmentation and various difficulties faces while segmenting word into individual components. Section 4 gives various techniques used in the literature for segmentation used in different scripting languages. Database used in the literature by different authors in their respective work is discussed in Section 5. Section 6 consists of validation and testing. Section 7 comprises of a brief discussion on the techniques used. Section 8 gives conclusion and future scope in the specified area of research. A comprehensive bibliography which includes most relevant papers related to the segmentation of offline Handwritten scripts is added to provide outline for development in the concerned field.

2. Need of Segmentation:

Holistic approach reduces the accuracy results as compared to segmentation approach(Shaw, Parui, and Shridhar 2008)[2][4]. Segmentation reduces the complexity of recognition. If word is properly segmented, then no. of classes used in the recognition system will be equal to the no. of characters and not more. Line segmentation followed by word segmentation give way to character level segmentation in a text image. Different level of segmentation is discussed in (Mehul et al. 2014). Character level segmentation is the lowest level of segmentation which presents

fundamental challenges due to variability in handwritten data.

3. Difficulties in Segmentation:

The horizontal line (Shirorekha) used in scripts like Devanagari (Hindi), Bangla, Gurumukhi (Punjabi), Marathi, Nepali makes segmentation problem more difficult. Spaces between the characters in handwritten data may vary which makes segmentation a difficult problem.

- Large character set which includes consonants, vowels, modifiers, compound characters in script makes segmentation more complicated.
- Different shapes/writing style/device used for writing further complicated the process of segmentation. Cursive Nature of handwriting make characters connected to each other.
- Characters sharing similar contours.
- Location of contact point at any elevation and non-linear boundary(Lu and Shridhar 1996).
- Finding junction path to segment touched components.

4. Literature Survey:

Various survey papers are available in the literature Authors in (Lu and Shridhar 1996)(Jayadevan et al. 2011b)(Soumen Bag and Ankit Krishna 2013)(Yarman-Vural 2001) discusses various algorithm for various techniques available for the purpose of segmentation and recognition of segmented components in handwritten text or word image.

4.1. English

The first survey that focuses on touched character is given by Tanzilsaba et. al. (Saba, Sulong, and Rehman 2010a). Various approached used for segmentation, segmentation rate, test data used for experiment till 2010 is provided in the survey.

Paper by chen et.al. (Chen 1994) used HMM(Hidden Markov Model) – stochastic network for unconstrained word recognition A segmentation approach is followed using morphology and heuristics based segmentation. The proposed algorithm used modified viterbi algorithm to search for best path. The resulted are obtained by applying the algorithm on 1583 images (1489 training images and 94 test images). The algorithm successfully segmented 95.6% of the trained images.

Authors used junction based approach and fuzzy features for the segmentation of touching string (Jayarathna and Bandara 2006). Character skeleton is used to find junction point i.e. pixel having more than three or more neighbouring points. Authors in (Saba, Sulong, and Rehman 2010b) proposed segmentation of touched characters in roman cursive characters based on Genetic algorithm. Experiments on cursive handwritten words are performed on unconstrained 300 word images. The results are tested on IAM benchmark database and up to 89.76% accuracy is obtained.

Contour based approach used by Ventzislavet. al. (Alexandrov 2004) used geometrical and structural information for finding critical point on the contours. Another technique for segmentation of offline cursive handwritten words is given by F. Kurniawanet. al. (Kurniawan et al. 2011) applied particularly for touching characters problem. Self-Organizing feature maps are employed to locate the segmentation path. The database is created using CCC cursive handwriting database. Two different samples were extracted and merged to form touching characters pair. 123 samples of the touching pair are considered for experimentation.

Chun Ki Cheng & Michael Blumenstein (Cheng and Blumenstein, n.d.) Used enhanced Heuristic Segmenter (EHS) and confidence values to locate possible segmentation points using ligature and global features of cursive handwriting. Authors in (Choudhary, Rishi, and Ahlawat 2013) proposes vertical segmentation algorithm which used thinned word image for finding the segmentation points. A single pixel image of the stroke is obtained. Ligatures are detected using the geometrical feature of characters.

Hybrid HMM (Hidden Markov Model)/ANN (Artificial Neural Network) is used by M.J. Castro et. al (Castro-bleda, Gorbe-moya, and Zamora-martinez 2011) for the recognition of unconstrained offline handwritten text. The structural part has been modelled with Markov chains and to estimate the emission probability is done using MLP (Multilayer Perceptron). Over segmentation is dealt by S. Sagar et. al. (Sagar and Dixit 2018) using Potentially segmented column (PSC) using HMM approach in cursive handwritten words. Namrata Dave (Dave and Colledge 2015) discusses various methodologies

and levels for segmentation of a text based image. Various techniques are reviewed with its limitations.

4.2. Numerical String

Ghazaliet. al. (Sulong, Rehman, and Saba 2010) used hybrid approach for segmentation of handwritten touched numeral string. More than 92% recognition results are obtained using 1,316 numeral strings. Results are obtained using NIST SD19 which are further distributed into four classes consisting of 2 digit (370), 3 digit (285), 4 digit (345), 5 digit (316) strings.

In (Elnagar and Alhadj 2003), feature are extracted from connected string numerals after pre-processing and thinning. Potential segmentation points are extracted based on deepest/highest valley/hill. Recognition rate of 96% is reported in the paper. NIST Database 19, CEDAR is used for experimentation.

Water reservoir technique and morphological features are used in (U. Pal, A. Behaid, C. Choisy 2003) to generate the segmentation points. The proposed scheme has 94.8% accuracy. Two digits touching components are considered. French bank cheques data is used for the experimentation.

(Kyung Kim, Ho Kim, and Suen 2002) used Ligature analysis is done to extract different touching types and structural features of contour to find break points in a word image. Recognition rate of 92.5% is obtained. Experimentation is done on the benchmark database (NIST SD19) on 3500 touching pairs of digits

4.3. Gurumukhi

Segmentation of a touching character in an offline handwritten Gurumukhi document is discussed in (Kumar 2014). Authors applied water reservoir approach to identify connected components and then further used the concept for segmentation with accuracy of 93.51%.

Authors in (Modi, N., & Jindal 2013) deals with text line segmentation considering overlapped and connected components. The proposed technique is applied on 30 sample document images with 289 lines and 75.78% accuracy is reported.

In (Kaur, Singh, and Rani 2015), the broken and overlapped character problem and applied projection profile with neighbouring pixel for touching components (characters) in Gurumukhi

script is discussed. Neighbouring pixel approach is applied in (Mangla, n.d.) For segmentation which included broken and touching word in Gurumukhi script. Database consisting of 50 words for isolated, touching and broken words is taken and accuracy of 97%, 95% and 95% respectively is reported.

Authors in (Sharma and Lehal 2006) proposed an iterative technique to segment words. Presence of headline, aspect ratio of characters, vertical and horizontal projection profiles are used as a characteristic feature to segment the words.

4.4. Kannada

In (Mamatha and Srikantamurthy 2012), segmentation scheme for unconstrained handwritten Kannada scripts is proposed. Segmentation of words and character is accomplished using projection profiles and morphological operations. 82.35% accuracy for words segmentation and 73.08% accuracy for characters segmentation respectively are reported. Author proposed (Venkatesh, Majjagi, and Vijayasenan 2014) implicit segment for character segmentation along with recognition using HMM. Thinning, branch-points and mean points used in (Naveena and Manjunath Aradhya 2012) are used to find segmentation points. Author used expectation-maximization for learning mixture of Gaussians.

4.5. Oriya

Tripathy and U.pal in (Tripathy and Pal 2004) proposed segmentation technique for Oriya handwritten text. Unconstrained text is used for experimentation. Oriya handwritten text into individual characters. Projection profile is used for line segmentation and structural features are used for word segmentation. Segmentation of isolated and touched characters is proposed using water reservoir, structural and topological based features. 96.7% accuracy is obtained using the proposed algorithm for two-character touching strings. 1840 touching components is prepared consisting of two or more characters touching each other while writing (two-character, three-characters or more than three characters touching each other) The accuracy for segmentation of 96.7%, 95.1% and 93.3% respectively is reported.

4.6. Bangla

U. pal. and Sagarika Datta (U. Pal, A. Behaid, C. Choisy 2003) used water reservoir technique and structural features for the purpose of segmentation of touching characters in Bangla. 1430 Bangla touching string images are taken for evaluation and 95.97% accuracy is observed.

A novel technique is proposed by Soumenet. al. in (Bag, S., & Krishna 2015) in handwritten Bangla documents for segmentation. Technique uses the isothetic covers properties for vertex characterization. Success rate of 96.4% is obtained on different handwritten data obtained from different individuals. For cursive handwritten words segmentations, S. Basuet. al. (Basu, Chaudhuri, and Kundu 2002) segment the word into its components using two phase segmentation.

4.7. Hindi

Large character set and irregularities poses difficulty in Hindi text segmentation is discussed by Garget. al. (Kumar Garg, Kaur, and K. Jindal 2011). The survey by Sharma et. al divulges that many papers are available for the line, word and character segmentation of Devanagari script but a few papers consider segmentation of touching character. Ashwin and Milind (Ramteke and Rane 2012) used connected component and projection profile for segmentation of offline handwritten words which is composed of isolated characters, but words containing touched characters are not considered. S. Kapoor and V. Verma (Kapoor and Verma 2014) devised a technique to segment the touching characters by identifying the joint points and structural properties. Hanmandluet. al. (Hanmandlu M., Pooja Agrawal 2001) used structural features for segmentation of handwritten Hindi words. The paper covers various issues using hierarchical segmentation approach like headline detection, separating upper and lower modifiers, identifying conjunct. 78% accuracy is reported by applying structural feature in hierarchical order.

Morphological operations used by author in (Ladwani and Malik 2010) and applied on 100 words with 57% accuracy reported for segmentation of top modifiers, 55% for lower modifiers and 52% accuracy rate for middle zone characters.

A script independent approach for character

segmentation is given by Ram sarkar et.al (Sarkar 2010). Various scripts like Bangla, Gurumukhi, Syloti, Devanagari share characteristics like presence of Shirrekha, presence of vowels and modifiers. Fuzzy features (horizontalness and verticalness) are used for Matra identification and to find segmentation point. Sample images of 400 words are used for experimentation and success rate of 95.41%, 93.61%, 91.23% and 92.37% or Bangla, Devanagari, Gurumukhi, Syloti respectively is reported.

5. Database

Non-availability of touching character database necessitates authors to create database for validation and testing their respective work. (Jayadevan et al. 2011a) A legal amount database consists of 26720 handwritten Hindi and Marathi words are used.

To encourage and promote research in Devanagari script, ICDAR make the database available on request for the handwritten word database consisting of 10070 samples of legal amount. Database[1] consists of 39,700 samples of town names collected from 436 writers. These handwritten samples are unconstrained and lexicon of size 100 is used. The training and test databases refereed in the paper [2] consists of 22500 and 17200 images respectively. Handwritten words of 100 word classes collected from 436 different writers are obtained. (Madhav Goyal 2013) The database of 1380 words has been used collected from 15 writers. Work is on isolated characters in Hindi. Author used 1907 handwritten words for testing proposed algorithm in (Sharma and Lehal 2006) consisting of 389 sets of connected characters. Database used in (Hanmandlu M., Pooja Agrawal 2001) consists of 1000 handwritten Hindi words, 500 handwritten samples each of 20 words in different styles.

6. Validation and Testing

Authors used various techniques to address the segmentation of line, text or words into characters. Due to non-availability of standard database, researchers need to create their respective databases and apply their techniques. Validating or comparison of their result obtained is not possible because no standard or benchmark database is available. As per the literature survey, results obtained using various techniques of segmentation are verified manually.

7. Discussion

Projection profile, connected components, structural properties, recognition and segmentation using Neural Network are the various techniques is applied by authors for segmentation of words into characters. Thinning algorithm is applied to the word image and candidate segmentation points are found. Contour based approach is applied in many papers to find valley and crest points. Approach like Water-reservoir is used in many Indian scripts with similar touching patterns. This approach was conceived by U.pal et.al. (U. Pal, A. Behaid, C. Choisy 2003) for segmenting touching numerals. Later the same approach was applied for Bangla, Oriya, Punjabi and Thai touching characters. A brief comparison is show in Table 1. According to literature survey, a few papers are available which consider segmentation of touching or fused characters in offline handwritten data as compare to online handwritten or printed data.

Conclusion and Future directions

Numbers of techniques are proposed for segmentation of text into their constituent components and authors used their respective self-created database for testing their proposed technique. This is the major challenge faced by the researchers in optical character recognition due to unavailability of the benchmark database. Performance evaluation in number of languages is done manually due to the lack of benchmark database. Contributing in the database is one way to aid the research community to fulfil the problem of non-availability of benchmark database. Isolated character and word database for Devanagari script (Hindi and Marathi) is available with CEDAR. Word database is available online and word database is made available on request. Techniques given in the literature for segmentation and recognition of online handwritten/printed or offline printed text are applied and tested under some text constraints. Enhanced techniques are required to address the problem segmentation in case of offline handwritten text or data. Further, a robust technique is required to segment and recognize script so that it can be applied to unconstrained handwritten text.

Table 1: Water Reservoir Segmentation approaches analysis

Author	Approach Used	No. of words/characters/Numerals	Accuracy
U.Pal et. al. (Pal and Datta 2003)	Water reservoir(Bangla)	1430	95.97%
Tripathy et. al.(Tripathy and Pal 2004)	Water reservoir(Oriya)	Two-characters touching(1485) & three- characters touching(311)	96.7% & 95.1% resp.
M. Kumar et. al.(Kumar, Jindal, and Sharma 2014)	Water reservoir(Punjabi)	Words collected from 300 handwritten Gurumukhi script document.	93.51%

References

- [1] Alexandrov, Ventzislav. 2004. "Using Critical Points in Contours for Segmentation of Touching Characters." *Proceedings of the 5th International Conference on Computer Systems and Technologies - CompSysTech '04*, 1. <https://doi.org/10.1145/1050330.1050366>.
- [2] Bag, S., & Krishna, A. 2015. "Character Segmentation of Hindi Unconstrained Handwritten Words." *International Workshop on Combinatorial Image Analysis. Springer, Cham, 2015* 9448: 247–60. <https://doi.org/10.1007/978-3-319-26145-4>.
- [3] Basu, S, C Chaudhuri, and M Kundu. 2002. "Segmentation of Offline Handwritten Bengali Script Component Segment," 171–74.
- [4] Castro-bleda, Maria Jose, Jorge Gorbemoya, and Francisco Zamora-martinez. 2011. "Improving Offline Handwritten Text Recognition with Hybrid HMM / ANN Models." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33 (4): 767–79.
- [5] Chen, Mou-Yen. 1994. "Off-Line Handwritten Word Recognition Using a Hidden Markov Model Type Stochastic Network." *Ieee Transactions on Pattern Analysis and Machine Intelligence* 16 (5): 481–96. <https://doi.org/10.1109/34.291449>.
- [6] Cheng, Chun Ki, and Michael Blumenstein. n.d. "Improving the Segmentation of Cursive Handwritten Words Using Ligature Detection and Neural Validation." *Proceedings of the 4th Asia Pacific International Symposium on Information Technology (APIS 2005), Gold Coast, Australia. 2005.*, 56–59.
- [7] Choudhary, Amit, Rahul Rishi, and Savita Ahlawat. 2013. "New Character Segmentation Approach for off-Line Cursive Handwritten Words." *Procedia Computer Science* 17: 88–95. <https://doi.org/10.1016/j.procs.2013.05.013>.
- [8] Dave, Namrata, and G H Patel College. 2015. "Segmentation Methods for Hand Written Character Recognition" 8 (4): 155–64.
- [9] Elnagar, Ashraf, and Reda Alhadjj. 2003. "Segmentation of Connected Handwritten Numeral Strings." *Pattern Recognition* 36 (3): 625–34. [https://doi.org/10.1016/S0031-3203\(02\)00097-3](https://doi.org/10.1016/S0031-3203(02)00097-3).
- [10] Hanmandlu M., Pooja Agrawal, Brejesh Lall. 2001. "Segmentation of Handwritten Hindi Text: A Structural Approach" 22 (4): 1–20.
- [11] Jayadevan, R., S. R. Kolhe, P. M. Patil, and Umapada Pal. 2011a. "Database Development and Recognition of Handwritten Devanagari Legal Amount Words." *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, no. September: 304–8. <https://doi.org/10.1109/ICDAR.2011.69>.
- [12] Jayadevan, R, Satish R Kolhe, Pradeep M Patil, and Umapada Pal. 2011b. "Offline Recognition of Devanagari Script: A Survey" 41 (6): 782–96.

- [13] Jayarathna, U.K.S., and G.E.M.D.C. Bandara. 2006. "A Junction Based Segmentation Algorithm for Offline Handwritten Connected Character Segmentation." 2006 *International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*, 147–53. <https://doi.org/10.1109/CIMCA.2006.11>.
- [14] Kapoor, Shuchi, and Vivek Verma. 2014. "Fragmentation of Handwritten Touching Characters in Devanagari Script." *International Journal of Information Technology, Modeling and Computing* 2 (1): 11–21.
- [15] Kaur, Akashdeep, Paramjeet Singh, and Shaveta Rani. 2015. "Segmentation of Broken and Isolated Characters in Handwritten Gurumukhi Word Using Neighboring Pixel Technique." *Transactions on Networks and Communications* 3 (2): 37–42. <https://doi.org/10.14738/tnc.32.1094>.
- [16] Kumar, Munish. 2014. "Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition." *International Journal of Information Technology and Computer Science (IJITCS)* 6.2 (2014):, no. January: 58–63. <https://doi.org/10.5815/ijitcs.2014.02.08>.
- [17] Kumar, Munish, M. K. Jindal, and R. K. Sharma. 2014. "Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition." *International Journal of Information Technology and Computer Science* 6 (2): 58–63. <https://doi.org/10.5815/ijitcs.2014.02.08>.
- [18] Kumar Garg, Naresh, Lakhwinder Kaur, and M. K. Jindal. 2011. "The Hazards in Segmentation of Handwritten Hindi Text." *International Journal of Computer Applications* 29 (2): 30–34. <https://doi.org/10.5120/3536-4830>.
- [19] Kurniawan, F., Mohd Shafry Mohd Rahim, Daut Daman, Amjad Rehman, Dzulkifli Mohamad, and Siti Mariyam Shamsuddin. 2011. "Region-Based Touched Character Segmentation in Handwritten Words." *International Journal of Innovative Computing, Information and Control* 7 (6): 3107–20.
- [20] Kyung Kim, Kye, Jin Ho Kim, and Ching Y. Suen. 2002. "Segmentation-Based Recognition of Handwritten Touching Pairs of Digits Using Structural Features." *Pattern Recognition Letters* 23 (1–3): 13–24. [https://doi.org/10.1016/S0167-8655\(01\)00114-3](https://doi.org/10.1016/S0167-8655(01)00114-3).
- [21] Ladwani, Vandana M., and Latesh Malik. 2010. "Novel Approach to Segmentation of Handwritten Devnagari Word." *Proceedings - 3rd International Conference on Emerging Trends in Engineering and Technology, ICETET* 2010, 219–24. <https://doi.org/10.1109/ICETET.2010.143>.
- [22] Lu, Yi, and M. Shridhar. 1996. "Character Segmentation in Handwritten Words - An Overview." *Pattern Recognition* 29 (1): 77–96. [https://doi.org/10.1016/0031-3203\(95\)00072-0](https://doi.org/10.1016/0031-3203(95)00072-0).
- [23] Madhav Goyal, Garg Naresh. 2013. "Techniques for Recognition of Handwritten Devanagari Numeral" 4 (9): 5697.
- [24] Mamatha, H R, and K Srikantamurthy. 2012. "Morphological Operations and Projection Profiles Based Segmentation of Handwritten Kannada Document." *International Journal of Applied Information Systems (IJ AIS)* 4 (5): 13–19.
- [25] Mangla, Parika. n.d. "Mangla, Parika, and Harleen Kaur. 'An End Detection Algorithm for Segmentation of Broken and Touching Characters in Handwritten Gurumukhi word.' Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2014 3rd Internatio."
- [26] Mehul, Gupta, Patel Ankita, Dave Namrata, Goradia Rahul, and Saurin Sheth. 2014. "Text-Based Image Segmentation Methodology." *Procedia Technology* 14: 465–72. <https://doi.org/10.1016/j.protcy.2014.08.059>.
- [27] Modi, N., & Jindal, K. 2013. "Text Line Detection and Segmentation in Handwritten Gurumukhi Scripts." *IJ of Advanced Research in Computer Science and Software Engineering* (2013). 3 (5): 1075–80.
- [28] Naveena, C., and V. N. Manjunath Aradhya. 2012. "Handwritten Character Segmentation

- for Kannada Scripts.” *Proceedings of the 2012 World Congress on Information and Communication Technologies, WICT 2012*, 144–49.
<https://doi.org/10.1109/WICT.2012.6409065>
- [29] Pal, U., and Sagarika Datta. 2003. “Segmentation of Bangla Unconstrained Handwritten Text.” In *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2003*–Janua:1128–32.
<https://doi.org/10.1109/ICDAR.2003.1227832>.
- [30] Ramteke, Ashwin S, and Milind E Rane. 2012. “Offline Handwritten Devanagari Script Segmentation.” *International Journal of Scientific and Research* 1 (4): 142–45.
<https://doi.org/10.5815/ijitcs.2012.08.04>.
- [31] Saba, Tanzila, Ghazali Sulong, and Amjad Rehman. 2010a. “A Survey on Methods and Strategies on Touched Characters Segmentation.” *International Journal of Research and Reviews in Computer Science (IJRRCS)* 1 (2): 103–14.
- [32]———. 2010b. “Non-Linear Segmentation of Touched Roman Characters Based on Genetic Algorithm.” *International Journal on Computer Science and Engineering* 2 (6): 2167–72.
- [33] Sagar, Sourabh, and Sunanda Dixit. 2018. “HMM Segmentation Approach for Offline Cursive Handwritten Words” 8 (5): 17372–77.
- [34] Sarkar, Ram. 2010. “A Script Independent Technique for Extraction of Characters from Handwritten Word Images” 1 (23): 83–88.
- [35] Sharma, Dharam Veer, and Gurpreet Singh Lehal. 2006. “An Iterative Algorithm for Segmentation of Isolated Handwritten Words in Gurmukhi Script.” In *Proceedings - International Conference on Pattern Recognition*, 2:1022–25.
<https://doi.org/10.1109/ICPR.2006.258>.
- [36] Shaw, Bikash. 2008a. “A Segmentation Based Approach to Offline Handwritten Devanagari Word Recognition.” *Pattern Recognition*, 256–57.
<https://doi.org/10.1109/ICIT.2008.32>.
- [37] 2008b. “Offline Handwritten Devanagari Word Recognition: A Segmentation Based Approach.” *Pattern Recognition, 2008. ICPR 2008. 19th International Conference On. IEEE, 2008*.
- [38] Shaw, Bikash, Swapan Kumar Parui, and Malayappan Shridhar. 2008. “Offline Handwritten Devanagari Word Recognition: A Holistic Approach Based on Directional Chain Code Feature and HMM.” *Proceedings - 11th International Conference on Information Technology, ICIT 2008*, 203–8.
<https://doi.org/10.1109/ICIT.2008.33>.
- [39] Soumen Bag and Ankit Krishna. 2013. “A Survey on Optical Character Recognition for Bangla and Devanagari Scripts.” *Sadhana - Academy Proceedings in Engineering Sciences* 38 (1): 133–68.
<https://doi.org/10.1007/s12046-013-0121-9>.
- [40] Sulong, Ghazali, Amjad Rehman, and Tanzila Saba. 2010. “Improved Offline Connected Script Recognition Based on Hybrid Strategy.” *International Journal of Engineering Science and Technology* 2.6 (2010) 2 (6): 1603–11.
- [41] Tripathy, N., and U. Pal. 2004. “Handwriting Segmentation of Unconstrained Oriya Text.” *Proceedings - International Workshop on Frontiers in Handwriting Recognition, IWFHR 31* (December): 306–11.
<https://doi.org/10.1109/IWFHR.2004.50>.
- [42] U. Pal, A. Behaid, C. Choisy. 2003. “Touching Numeral Segmentation Using Water Reservoir Concept.” *Pattern Recognition Letters* 24 (jan): 261–72.
- [43] Venkatesh, Manasij, Vikas Majjagi, and Deepu Vijayasanen. 2014. “Implicit Segmentation of Kannada Characters in Offline Handwriting Recognition Using Hidden Markov Models.” *Implicit arXiv Preprint arXiv:1410.4341* (2014)., 1–6.
- [44] Yarman-Vural, Nafiz Arica and Fatos T. 2001. “An Overview of Character Recognition Focused on Off-Line Handwriting.” *Indian Academy of Sciences VOL.* 31 (2): 216–33.
<https://doi.org/10.1109/5326.941845>.