RESEARCH REVIEW JOURNALS
Volume-03
Issue-11
November -2018

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com [UGC Listed Journal]

# Online Handwriting Recognition for Smaller Units in Devanagari and Tamil Scripts

Sukhdeep Singh

*Assistant Professor in Computer Science, Mata Sahib Kaur Girls College, Talwandi Bhai, Ferozepur, Punjab (India)*

**ARTICLE DETAILS**

**Article History**
*Published Online:* 10 November 2018

**Keywords**
Indic, Tamil, Devanagari, Online handwriting recognition

**ABSTRACT**

The research for handwriting in Indic scripts is at early stage as compared to Japanese, Chinese, Korean, Latin and Indic scripts. It is especially limited for larger units' online handwriting in Indic scripts. In the present study, we have demonstrated the available work for smaller units in two Indic scripts: Devanagari and Tamil. In these studies, it has been observed, the work done for smaller units in these scripts is much more as compared to larger units work in Devanagari and Tamil scripts. The statistical recognition techniques have been successfully used in these results. The smaller units results presented in present study have been obtained from premier journals and conferences of artificial intelligence, pattern recognition and handwriting recognition as Pattern Analysis and Machine Intelligence (PAMI), Pattern Recognition(PR), Pattern Recognition Letters (PRL), International Journal on Document Analysis and Recognition(IJDAR) , International Conference on Document Analysis and Recognition(ICDAR) and International Conference on Frontiers in Handwriting Recognition (ICFHR) etcetera.

## 1. Introduction

The 21st century is the modern era of technology. The modernization has also resulted in technological advancements. The technological advancements have been made in a variety of ways. One of the innovations in technological advancement is the machine learning. The machine learning has been proven a boon for devising new methods of exchanging information between computers and human beings. The speech recognition has been successfully employed for exchanging information between computers and human beings. But this way of exchanging information has certain drawbacks. To remove these drawbacks, a new area of research is active from past few years and this area is the handwriting recognition. The handwriting recognition refers to a process where handwritten information and text is understood by the computer. Further the handwriting can be understood by scanning the paper or while writing on the screen of computer, the later one is called online handwriting recognition. The online handwriting recognition uses the x-y coordinate values and time information of handwritten strokes for identifying the text. So the fundamental unit of online handwriting is the pixel points and then pixel coordinate values are used to form strokes. The stroke is the pen down to pen up activity. The strokes combinations are further used to form characters, and a group of characters form words. The present study has been done for online handwritten characters.

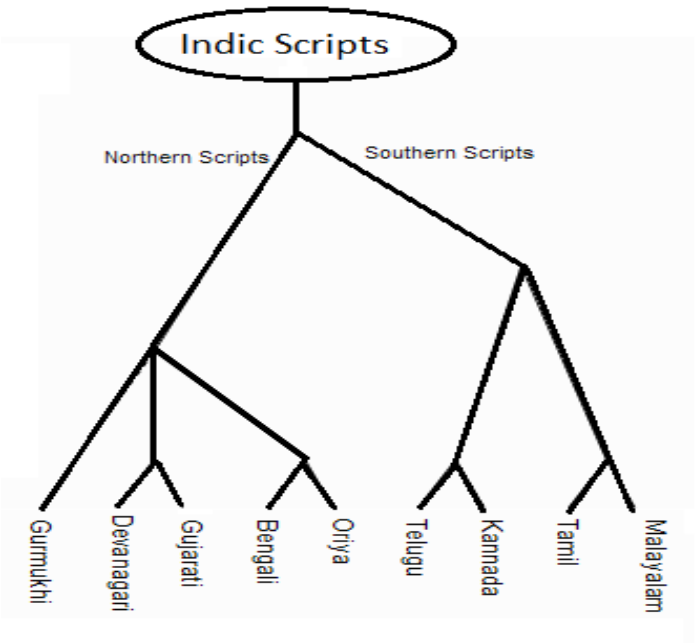


**Figure 1. Indic scripts**

As the spoken languages used by human beings change over different geographical regions on the earth, in the same way the scripts used by human beings for conveying information also over different regions across the globe. In

view to see the scope of present study, these scripts can be divided into various classes. If we look at non Indic scripts, the scripts like Japanese, Korean, Chinese, Latin and Arabic etcetera fall into the family of non- Indic scripts. The Indic scripts are those scripts that are spoken by the people who reside around the Indus river and these scripts are also called Brahmi scripts. The Indic scripts are shown in figure 1. Devanagari, Tamil, Telugu, Gujarati, Gurmukhi, Kannada and Bengali etcetera are popular Indic scripts. These all scripts are related to each other in a variety of ways. So these are not isolated scripts and share many properties with each other. So the work done for any of these scripts, also affect the future work of rest scripts. Among these scripts, the present study has considered the two popular scripts: Devanagari and Tamil. Overall, the present study has surveyed and analyzed the recent years' work done for online handwritten character recognition in these scripts.

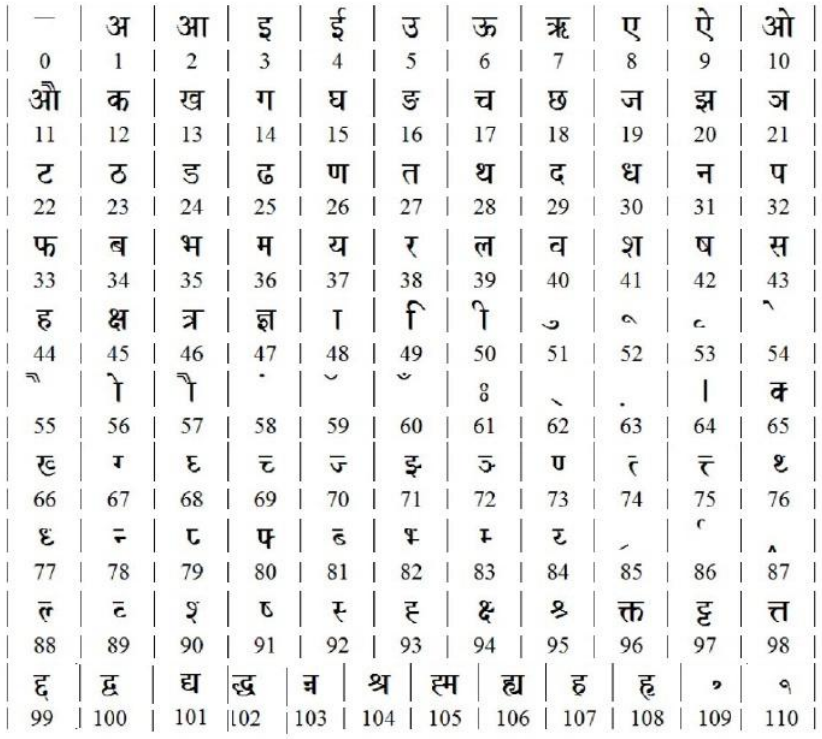


**Figure 2. Devanagari script symbol set**

## 2. Symbol Set

Many Brahmi scripts have more than two hundred and fifty alphabets where these alphabets include both basic and compound characters[1]. The current work has been donefor online Tamil and Devanagari handwriting recognition for smaller units: characters. These are widely used scripts in India. Theseare not alone/isolated script; these sharemany characteristics with otherBrahmi scripts. For example, the stroke/symbol order variation challenge is common for most of Indic scripts and it is also present in these two scripts.For online handwritten characters or words recognition in these two scripts, the primary phase is to identify the fundamental/primary units of the script are known as symbols. The symbol sets forDevanagari and Tamilscripts arepresented in figure 2 and 3, respectively. The description of Devanagari script symbols is given in table 1, and the description of Tamil script symbols is given in table 2.

**Table 1. Devanagari script symbols description**

| Sr. no. | Symbol number | Symbol description |
|---|---|---|
| 1 | 0 | Headline |
| 2 | 1-11 | Independent vowels |
| 3 | 12-44 | Consonants having implicit vowel sounds |
| 4 | 48-63, 109-110 | Consonants |
| 5 | 64 | Sentence ending |
| 6 | 65 and 95 | Matras |
| 7 | 45-47, 96-108 | Conjuncts |

 


**Figure 3. Tamil script symbol set**

**Table 2. Tamil script symbols description**

| Sr. no. | Symbol number | Symbol description |
|---|---|---|
| 1 | 0-10 | Vowels |
| 2 | 11 | Special symbol |
| 3 | 12-33 | Consonants having implicit vowel sounds |
| 4 | 72-88 | Matras |
| 5 | 72 | Coverts consonants to half characters |
| 6 | 81 and 82 | Conjuncts |
| 7 | 83 | Dot |

### 3. Literature Survey

To recognize handwrittentext in Devanagari and Tamil scripts in online handwriting mode, it is very important to studythe existing workfor online handwritten text recognition in these two scripts. The existing study for thesetwo scripts' online handwriting recognition has been done for smaller units and larger units. For smaller units, a great work has been done for online Devanagari and Tamil handwriting recognition. These smaller units include strokes and characters. Considering the use of smaller units online handwritingrecognition for larger units, the present study analyzed the related work for Devanagari and Tamil scripts for smaller units as characters.Like otherBrahmi scripts, a greatwork for Devanagari and Tamil scripts online handwriting recognition has been done in recent two decades. The table 1 shows the Devanagari and Tamil scripts online handwriting recognition work in the recent past.

**Table 1.Online handwritten characterrecognition for Devanagari and Tamil scripts**

| Sr. no. | Authors and references | Year | Script | Classification techniques | Recognition rate (%) |
|---|---|---|---|---|---|
| 1 | Ghosh and Roy[2] | 2015 | Devanagari | ZSDP, Support Vector Machine (SVM) | 90.63 |
| 2 | Ghosh and Roy [2] | 2015 | Devanagari | Characters Zone wise structural anddirectional features (ZSD), SVM | 85.10 |
| 3 | Chowdhury et al. [3] | 2013 | Devanagari | Levenshtein distance metric | 83.95 |
| 4 | Mondal et al. [4] | 2010 | Devanagari | Point-float feature, HMM | 82.43 |
| 5 | Mondal et al. [4] | 2010 | Devanagari | Point-float feature, Multilayer perceptron (MLP) | 83.30 |
| 6 | Mondal et al. [4] | 2010 | Devanagari | Chain-code feature, HMM | 87.13 |
| 7 | Mondal et al. [4] | 2010 | Devanagari | Chain-code feature, MLP | 86.15 |
| 8 | Swethalakshmi et al. [5] | 2007 | Devanagari | SVM | 96.69 (42 classes) and 97.27 (82 classes) |
| 9 | Kunwar et al. [6] | 2014 | Tamil | Bayesian network (BN) | 83.85 |

 

| | | | | | |
|---|---|---|---|---|---|
| 10 | Kunwar et al. [6] | 2014 | Tamil | Random BN (RBN) | 86.10 |
| 11 | Kunwar et al. [6] | 2014 | Tamil | Online RBN (ORBN) | 87.80 |
| 12 | Kunwar et al. [6] | 2014 | Tamil | Semi-supervised ORBN (SSORBN) | 88.48 |
| 13 | Kunwar et al. [6] | 2014 | Tamil | Naive bayesian (NB) | 78.26 |
| 14 | Kunwar et al. [6] | 2014 | Tamil | Support vector machine (SVM) | 90.68 |
| 15 | Kunwar et al. [6] | 2014 | Tamil | Hidden Markov model (HMM) | 87.82 |
| 16 | Chowdhury et al. [3] | 2013 | Tamil | Levenshtein distance metric | 85 |
| 17 | Mondal et al. [4] | 2010 | Tamil | Point-float feature, HMM | 84.67 |
| 18 | Mondal et al. [4] | 2010 | Tamil | Point-float feature,Multilayer perceptron (MLP) | 84.98 |
| 19 | Mondal et al. [4] | 2010 | Tamil | Chain-code feature, HMM | 92.10 |
| 20 | Mondal et al. [4] | 2010 | Tamil | Chain-code feature, MLP | 91.80 |

## 4. Conclusion

This study is step for online handwriting recognition in Indic scripts. This work has demonstrated that a great work has been done for online Devanagari and Tamil handwriting recognition, where work has been done for smaller units (characters) recognition. From the present study, it has been analyzed that promising recognition results (above 80% and 85% recognition accuracy) have been attained for smaller units' recognition in these two scripts. Further, it needs more attention of pattern recognition researchers to carry out work for Indic scripts' recognition in online handwriting mode. For larger units' online handwriting recognition in Indic scripts, a lot of work is to be done in the near future.

## References


1. B. B. Chaudhuri and U. Pal, "An OCR system to read two Indian language scripts: Bangla and Devanagari (Hindi)", In: *Proceedings of 4th International Conference on Document Analysis and Recognition*, 1997.
2. R. Ghosh and P. P. Roy,"Comparison of Zone-Features for Online Bengali and Devanagari Word Recognition Using HMM", In: *Proceedings of 15thInternational Conference on Frontiers in Handwriting Recognition (ICFHR),* 435–440, 2016.
3. S. D. Chowdhury, U. Bhattacharya, and S. K. Parui, "Online Handwriting Recognition Using Levenshtein Distance Metric", In: *Proceeding of 12thInternational Conference on Document Analysis and Recognition*, 79–83, 2013.
4. T. Mondal, U. Bhattacharya, S. K. Parui, and K. Das, "On-line handwriting recognition of Indian scripts - the first benchmark", In: *Proceeding of the 12thInternational Conference on Frontiers in Handwriting Recognition*, 200–205, 2010.
5. H. Swethalakshmi, A. Jayaraman, V. S. Chakravarthy and C. Sekhar, "Online handwritten character recognition of Devanagari and Telugu characters using support vector machine", *In the proceedings of 10th IWFHR*, 367-372, 2006.
6. R. Kunwar, U. Pal, and M. Blumenstein, "Semi-Supervised Online Bayesian Network Learner for Handwritten Characters Recognition", In*: Proceedings of 22ndInternational Conference on PatternRecognition*, 3104–3109, 2014.


 