International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

A Survey on Approaches for Recognition of Dravidian Tulu Script

Amrutha[1], Shwetha Kamth[2]

[1]Dept OF CS&E, MITE, Moodabidri, Mangalore
[2] Dept OF CS&E, MITE, Moodabidri, Mangalore

Abstract - Handwritten script recognition is a interesting task that has been widely researched in the arena of computer vision and image processing. Tulu language is a Dravidian language spoken in coastal regions of Southern part of India. Recognition of Tulu handwritten scripts is important for preserving cultural heritage, historical documents, and facilitating communication among native speakers. This survey paper provides an overview of the existing tactics for recognition of Tulu scripts. We present a comprehensive review of the literature on various methods such as feature extraction, segmentation, classification, and deep learning techniques used for Tulu script recognition. The paper also discusses the challenges associated with Tulu script recognition and highlights the future research directions in this area.

Key Words: Tulu, Handwritten script recognition, Feature extraction, Segmentation, Classification, Deep learning

1.INTRODUCTION

Handwritten script recognition is a challenging task due to the variability in writing styles, non-uniformity in letter formation, and distortions introduced during the writing process. Tulu is the main Dravidian language which is spoken in the coastal regions of south India. Recognition of Tulu handwritten scripts is important for preserving cultural heritage, historical documents, and facilitating communication among native speakers. The development of efficient Tulu script recognition systems is essential for digitization of Tulu manuscripts, automatic translation, and education purposes

2. FEATURES OF TULU SCRIPT:

The Dravidian Tulu language spoken in the coastal regions of Kerala and Karnataka in India. The Tulu script, also known as the Tigalari script, is an abugida writing system which is used to write the Tulu language. Recognition of Tulu scripts has been a exciting task due to the variability in handwriting styles and the complexity of the script.

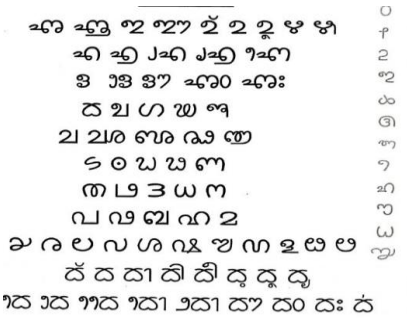

Fig -1: Tulu alphabet

Fig- 2: Vowels of Tulu script

The Tulu-Tigalari script is used on the numerous historical Tulu inscriptions that have been discovered near Udupi and Kundapura. The Tulu script was factually used by Brahmins and Havyaka Brahmins of Tulu Nadu to compose Vedas and Sanskrit writings. Through the Grantha script, the Tulu writing is inclined from the Brahmi script. It is the Malayalam script's sibling script. Tulu script has 5 long and 5 short vowels (*a*, *ā*, *e*, *ē*, *u*, *ū*, *i*, *ī*, *o*, *ō*) are common in Dravidian languages. There are totally 36 consonants in Tulu.

3. LITERATURE REVIEW

The recognition of Tulu handwritten scripts has been an energetic area of research in current years. Various tactics have been projected for recognition of Tulu scripts. One of the earliest works on Tulu script recognition was proposed by Hegde et al. (2010) which used a combination of morphological operations and HMM-based recognition for recognizing Tulu characters. Later, Deepthi(2014) used a method based on contour-based feature removal for recognition of Tulu characters. Kini et al. (2018) proposed a hybrid method grounded on a combination of contour-based features and SVM classification for Tulu script recognition.

.In 2014, a research paper titled "Recognition of Handwritten Tulu Characters Using Neural Networks" was published by S. Umesh and K. K. Achary. The study used a backpropagation neural network to recognize handwritten Tulu characters, achieving an accuracy of 93%.

In 2019, a research paper titled "Offline Handwritten Tulu Character Recognition Using Deep Learning Techniques" was published by P. V. Rajeeva and K. S. Sesh Kumar. The study used a deep convolutional neural network (CNN) to recognize handwritten Tulu characters, achieving an accuracy of 97.5%.

In 2020, another research paper titled "Tulu Handwritten Character Recognition Using Convolutional Neural Network and Recurrent Neural Network" was published by K. G. Arun and K. K. Achary. The study used a grouping of CNN and recurrent neural network (RNN) to recognize handwritten Tulu typescripts, achieving an accuracy of 98.6%.

In 2021, a research paper titled "Offline Tulu Script Recognition using CNN and Capsule Network" was published by J. S. Shetty and K. S. Sesh Kumar. The study used a combination of CNN and capsule network to recognize offline Tulu handwritten scripts, achieving an accuracy of 98.4%.

More recently, deep learning-based approaches have gained popularity for Tulu script recognition. Kumari(2019) proposed a way based on(CNN) for recognition of Tulu characters. Similarly, Hegde(2021) proposed a deep learning approach for Tulu script recognition using a combination of long and short-term memory (LSTM) networks and CNN.

4. METHODOLOGY:

Methodology basically used for recognizing Dravidian tulu script can be divided into the following steps Data collection: Collect a dataset of Tulu script images from various sources. The dataset should cover a wide range of variations in handwriting styles, image quality, and backgrounds.

Pre-processing: Pre-process the dataset images by converting them into grayscale, applying noise reduction filters, and resizing them to a standard size.

Feature extraction: Extract relevant features from the pre-processed images, such as shape, texture, and intensity. This can be done using techniques called Scale-Invariant Feature Transform (SIFT), Speeded Up Robust Feature (SURF), or Local Binary Patterns (LBP).

Training: Train a classifier on the extracted features using a suitable machine learning system, such as SVM, Random Forests, CNN. The classifier should be trained to distinguish between different Tulu script characters.

Testing: Evaluate the performance of the classifier on a distinct trial dataset. This can be completed by calculating metrics such as accurateness, F1 score ,precision, recall.

Improvements: If the performance of the classifier is not acceptable, improve the methodology by altering the parameters of the feature extraction and machine learning algorithms or by using more advanced techniques.

5. CHALLENGES AND FUTURE DIRECTIONS:

Despite the progress made in Tulu script recognition, there are several tasks that **to be taken care**. one of the biggest obstacles is the absence of a huge dataset of Tulu handwritten scripts. Another challenge is the high variability in handwriting styles among Tulu writers. Additionally, the presence of noise, blur, and other distortions in the handwritten scripts makes recognition difficult.

Future research directions in Tulu script recognition include the development of more robust feature extraction techniques that can handle the variability in handwriting styles, the creation of a larger dataset of Tulu handwritten scripts, and the exploration of new deep learning architectures for Tulu script recognition.

Table -1: Accuracy of recognizing Tulu scripts by using various techniques

Author	Technique	Accuracy
S. Umesh and K. K. Achary.	Neural Networks	93%
P. V. Rajeeva and K. S. Sesh Kumar	Deep Learning and Neural Networks	97.5%
K. G. Arun and K. K. Achary	Convolutional Neural Network and Recurrent Neural Network	98.6
J. S. Shetty and K. S. Sesh Kumar	CNN and Capsule Network	98.4%

6. CONCLUSIONS

Tulu script recognition is an important study area that has gained increasing attention in recent years. Various approaches, including traditional feature extraction and deep learning-based methods, have been proposed for recognition of Tulu handwritten scripts. However, there are still several tasks that need to be addressed, including the lack of a big dataset, the high variability in handwriting styles, and the presence of noise and distortions. The development of more robust and accurate

Tulu script recognition systems is essential for preserving cultural heritage and facilitating communication among native speakers.

REFERENCES

1. 1. Vaishnavi Murthy Yerkadithaya(2012), "*Tending to the Inheritance of Tulu Script*".

2. Anitha Marry M O,Chacko and Dhanya P M," *Multiple Classifier System for Offline Malayalam Character Recognition 2014 ICICT(Elsevier)*".

3. Tao, Dapeng, Xu Lin,Lianwen Jin and Xuelong Li. "*Principal component 2-D long short-term memory for font recognition on single Chinese characters*' 2016 IEEE transactions on cybernetics 46(IEEE) 3 pp 756-765.

4. Prasad, Binod Kumar, and Goutam Sanyal. "*A hybrid feature extraction scheme for Off-line English numeral recognition*." In Convergence of Technology (I2CT), 2014 International Conference for, pp. 1-5. IEEE, 2014.

5. Joshi, Dhanashree, and Sarika Pansare. "*Combination of Multiple Image Features along with KNN Classifier for Classification of Marathi Barakhadi*." In Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on, pp. 607-610. IEEE, 2015.

6 . P J Antony; C K Savitha(2016), "*A framework for recognition of handwritten South Dravidian Tulu script*", 2016 Conference on Advances in Signal Processing (CASP).

7. Manoj Balaji Jagadeeshan, Ganesh N Sharma(2019), "*Recognition of Off-line Kannada Handwritten Characters by Deep Learning using Capsule Network*", International Journal of Engineering and Advanced Technology (IJEAT) (pp.2249-8958).

8. C K Savitha, P J Antony(2019), "*PNN and Deep Learning Based Character Recognition System for Tulu Manuscripts*", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019.

9. Sachin Bhat, Gumstha Seshikala(2021), "*Character Recognition of Tulu Script Using Convolutional Neural Network",* Advances in Artificial Intelligence and Data Engineering (pp.121-131).