ISSN No:-2456-2165

Ancient Tamil Translator System

Durga Devi.G
Department of Electronics and Communication Engineering,
KCG College of Technology
Chennai, Tamil Nadu, India

Dhivya.R
Department of Electronics and Communication Engineering,
KCG College of Technology
Chennai, Tamil Nadu, India

Mahalakshmi.S
Department of Electronics and Communication Engineering,
KCG College of Technology
Chennai, Tamil Nadu, India

Feba Sosa Abraham
Department of Electronics and Communication Engineering,
KCG College of Technology
Chennai, Tamil Nadu, India

Abstract*:-* A simple method of identifying the characters of ancient Tamil inscription. This method is most specifically proposed for the Vattezhuthu type of Tamil characters. This type of characters were used from 8th to 15th century is not similar to modern Tamil and cannot be understood easily. These inscriptions contain some of the important information such as life of kings, their contributions to the society and also information related to medical field. Hence in an effort to retrieve information from these inscriptions and also from the written scripts this paper has been developed. This method translates the ancient character to modern Tamil which can be understood by todays' generation using non-subsampled contour-let transform (NSCT) and probabilistic neural network (PNN). Median filter is used for removing noise in the image.

Keywords*:- Median filter, non-subsampled contour-let transform, probabilistic neural network.*

I. INTRODUCTION

Tamil is one of the ancient languages which had evolved in its form greatly from the ancient times. We can find great difference between the ancient Tamil words and the modern Tamil words. This difference makes it difficult for us to understand the ancient Tamil scriptures and inscriptions. And there also exists many different styles of writing which has led to more complexity in understanding our ancient writings. Hence we need some translations to understand our ancient Tamil that can be converted into modern Tamil language.

Therefore we are in need to create a database that consists a collection of ancient Tamil language style writings so that we could know how the language has evolved which also helps us in translating ancient Tamil to modern Tamil language.

Image recognition, in the context of machine vision, is the ability of software to identify objects, places, people, writing and actions in images. It is achieved in combination with camera and artificial intelligence technique. Software for image recognition requires deep machine learning. Image recognition algorithms can function on 3D models.

There are many types of image formats that are available which on conversion may lead to loss of data. The main aim of image conversion is to retrieve the information in the image without any loss.

II. RELATED WORK

Various efforts have been taken in translating the ancient Tamil of which have been successfully proposed for Brahmi type of letters. The other two important format of ancient Tamil are Vattazhuthu and Grantha which has been posing a great difficulty for translation due to lack of database for these letters. Grantha letters have also been translated with minimum training of database.

Although all the efforts did not yield full efficiency for translating the letters phases of previous works done are continued to improve the efficiency. The proposed method is an effort to translate Vattezhuthu letters.

III. PROPOSED METHOD

In this method, the input image is first gray converted and then pre-processed using the median filter to remove the noises. Then the contour-let transform is used to enhance the image with respect to the center object which is followed by the morphological operation. It performs four operations namely dilation, erosion, filling and taking details of regional properties. Using the information provided by the regional properties textural feature extraction is done for each segmented image. The feature of dataset is also extracted and trained using Back Propagation Neural Network method and the features are compared and the

corresponding modern Tamil letter is seen as the output. Fig.1 illustrates the methodology.

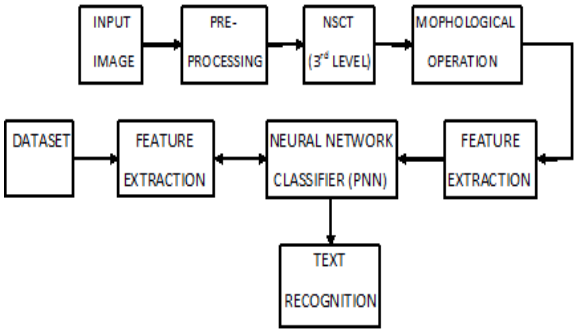

Fig 1:- Block diagram of proposed method

A. Pre-Processing

Pre-processing is done to enhance the image for further operations and to remove the unwanted noise from the image. First the input image is gray converted and then the median filter is applied. This filter enhances the image without any loss in the edges of the image. This center value of the block and adjusts the intensity value to the median value i.e. the center value. This process is repeated for all the blocks in the image and the image intensity is equalized without any odd blur in the image.

B. NSCT

NSCT stands for the Non Subsampled Contour-let Transform. It is evolved from contour-let transform due to shift invariant property since it contains both up-sampling and down-sampling constraints. In NSCT only up-sampling is performed using non-subsampled pyramidal and directional filters. It helps in image denoising, despeckling and enhancement. Since image contains mostly salt and pepper noise it is most efficient in removing it.

C. Morphological Operations

These operations are related to the shape of the images. The four morphological operations used here are dilation, erosion, filling and taking details of regional properties. Dilation is adding the pixels to the image to remove discontinuities. Erosion is adding pixels to the images to adjust the pixel value. Filling is used to adjust the intensity of the image. Regional properties gives information about the grayscale, contour plot of the image, data about the binary image, histogram data, pixel values, correlation and mean values of the image.

D. Feature Extraction

Gray level co-occurrence matrix (GLCM) is used to extract the textural features of the image. The statistics calculated are contrast, correlation, energy and homogeneity. The matrix is formed using the function graycomatrix and the statistics are calculated using the function graycoprops.

E. Neural Network Classifier

The type of neural network used is Probabilistic Neural Network (PNN). It calculates the probability distribution function of each class and compares it with the input class and the allocates the highest probability rate for the output.

IV. SIMULATION AND RESULTS

The training of database takes time to generate the pdf function for the class to classify the outputs. The input is given as Vattezhuthu font. Fig.2 and Fig.3 represents the input and output image of a single letter. Fig.4 and Fig.5 represents the input and output image of the group of letters.

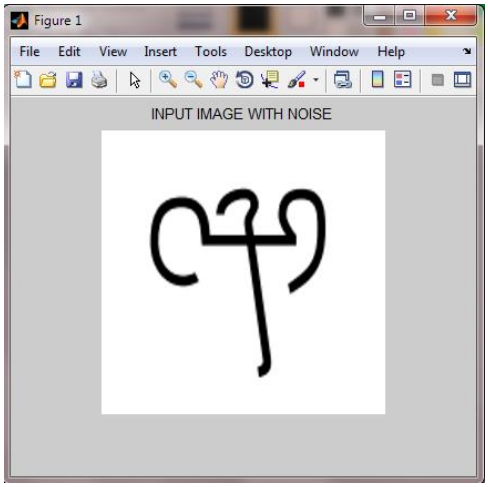

Fig 2:- Input of single letter

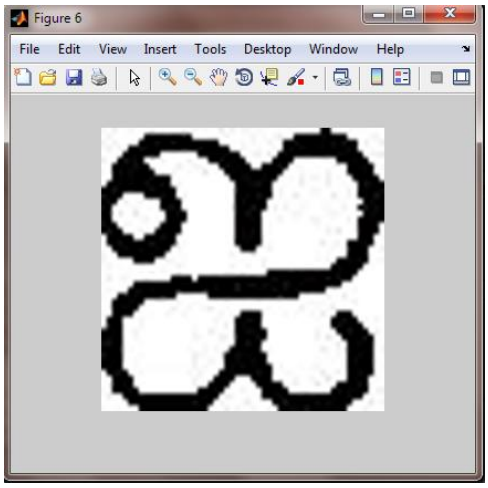

Fig 3:- Output of single letter

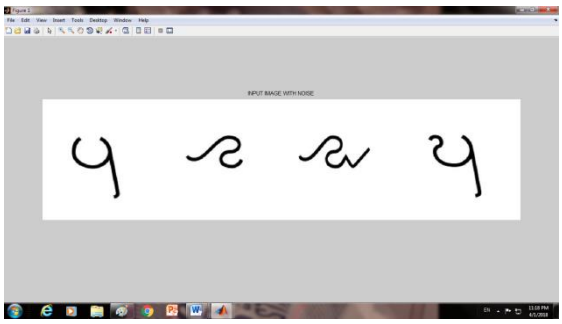

Fig 4:- Input image of group of letters

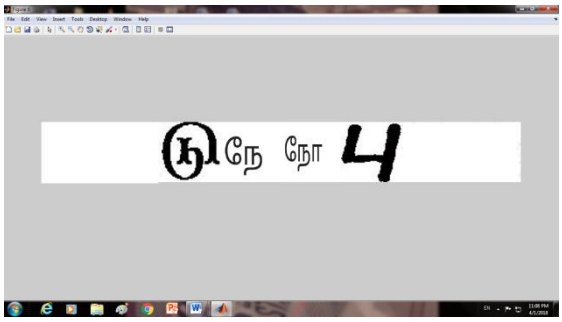

Fig 5:- Output image of group of letters

V. CONCLUSION

This work concentrates on translating the Vattezhuthu style of ancient Tamil. Further, the work can be improved by adding more details to the database using handwritten scripts and improving the algorithm for image recognition for efficient translation of the whole image. The work can be further developed by combining the works done for all the three format of ancient Tamil letters to create an application which would be useful for people to understand the inscriptions of Tamil language.

REFERENCES

[1]. Yoshinari Tsuzuki, Kei Sawada, Kei Hashimoto, Yoshihiko Nankaku and Keiichi Tokuda- "Image Recognition based on Discriminative Models using features generated from separable lattice HMMs"(ICASSP-2017)

[2]. R.Giridharan, E.K. Vellingiriraj and P. Balasubramanie- "Identification of Tamil Ancient Characters and Information retrieval from temple epigraphy using image zoning" (ICRTIT 2016).

[3]. M.Mahalakshmi, Malathi Sharavanan- "Ancient Tamil Script and Recognition and Translation using LabVIEW" (ICCSP 2013).

[4]. K. Sawada, A. Tamamori, K. Hashimoto, Y. Nankaku, and K. Tokuda, "Bayesian approach to image recognition based on separable lattice hidden Markov models," IEICE Transactions on Information and Systems, vol.E99-D, no.12, pp3119–3131, 2016.

[5]. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma,Z. Huang, A. Karpathy, A. Khosla, M.Bernstein, A.C. Berg, and L. Fei-Fei,"ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision,vol. 115, no. 3, pp. 211–252, 2015.

[6]. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R.B. Girshick, S. Guadarrama, and T. Darrell,"Caffe: Con-volutional architecture for fast feature embedding," ACM international conference on Multimedia, pp. 675–678, 2014.

[7]. S. Rajkumar and Dr. Subbiah Bharathi-"Century Identification and Recognition of Ancient Tamil Character Recognition",IJCA(0975-8887), Volume 26-No.4, July 2011.