International Journal of Computer Science Trends and Technology (IJCST) – Volume 5 Issue 5, Sep – Oct 2017

RESEARCH ARTICLE OPEN ACCESS

Comparison of Segmentation based on Threshold and K-Means Method

R.Spurgen Ratheash, Dr.M.Mohmed Sathik
Department of Information Technology, Principal
Sadakathullah Appa College, Tirunelveli
Tamil Nadu - India

ABSTRACT

In MRI brain images segmentation, extraction and detection of tumor infected area from the basic brain image properties are the primary, tedious and time taking process. The accuracy of separation the tumor area is based on the experience of clinical experts or radiologists. So, we need computer aided technology to overcome those limitations. In this study, we do automatic methods to reduce the complexity and improve the performance of MRI bran image segmentation. We have investigated many algorithms are available in medical imaging research area such as K-means clustering, Threshold technique, FCMeans, Watershed and Hierarchical Clustering (WHC) and so on. The proposed method compares Threshold technique and K-means clustering technique. The experimental results of proposed techniques have been evaluated and validated for performance and quality analysis on magnetic resonance brain images, based on segmented area, min and major axis and process time for the segmentation. The experimental results achieved more accuracy, less running time and high resolution.

Keywords:- Threshold technique, k-means clustering

I. INTRODUCTION

The uncontrolled growth of cancerous cells in the brain is called as tumours. The brain cells are identified by Benign or Malignant. Malignant is an active cancerous cells with rapid growth in the brain. The Benign cells are not the dangerous cancerous cells. The Benign cells can be converted into Malignant cells but the Malignant cells never become Benign cells.[1] This study about the segmentation of abnormal brain cells among normal brain properties such as Gray Matter (GM), White Matter (WM), and CerebroSpinalFluid (CSF) in magnetic resonance (MR) images using Threshold technique[2] and K-means Clustering technique.

The Digital image segmentation is employed automatic detection of brain tumor from MRI brain imaging modalities, Segmentation is necessary and important step in image analysis; it is a process of separating an image into different regions, blocks or clusters sharing common and identical properties, such as contrast, patterns of pixels, and distance around the boundary of the region, and gray level.

II. LITERATURE SURVEY

Automatic identifying and extraction of brain tumor has proposed by the techniques like Threshold, K-Means Clustering, Fuzzy Clustering Means (FCM), Pulse Couple Neural Network(PCNN) algorithm, Expectation Maximization (EM) segmentation algorithm, Watershed and Hierarchical Clustering (WHC) algorithm, support vector machine (SVM), artificial neural network (ANN) algorithm[3]. The above literature survey has revealed that some of the techniques are invented to obtain segmentation the brain area from the skull area; some of the techniques are invented to obtain feature extraction and some of the techniques are invented to obtain classification only[4]. Threshold Technique based on image intensity and K-means clustering algorithm based on the clusters[5] in MRI brain images, the K-means clustering method gives an effective segmentation of tumor region. This analysis on combined approach could not be conducted in any published literature.

III. THRESHOLD

In Threshold technique is based on histogram to identify the infected areas by deep and sharp valley between two peaks representing objects and background respectively.

The threshold can be chosen at the bottom of this valley. However, for most MR images, it is often difficult to detect the valley bottom precisely when the valley is flat and broad, imbued with noise, or when the two peaks are extremely unequal in height, often producing no traceable valley[6,7] .The threshold method can choose the value and separate the object from its background.

Let the pixels of a given picture be represented in L gray levels $[1, 2, \ldots, L]$. The number of pixels at level i is denoted by n_i and the total number of pixels by $N = n_1 + n_2 + \ldots\ldots + n_L$. In order to simplify the discussion, the gray-level histogram is normalized and regarded as a probability distribution:

$$P_i = n_i / N,\ P_i \geq 0,\ \sum_{i=1}^{L} P_i = 1 \qquad (1)$$

Now suppose that we dichotomize the pixels into two classes Co and C1 (background and objects, or vice versa) by a threshold at level K; C_0 denotes pixels with levels [$1, .., k$], and C_1 denotes pixels with levels [$k+1, ... , L$]. Then the probabilities of class occurrence and the class mean levels, respectively, are given by

$$\omega_0=P_r(C_0) \quad (2)$$
$$\omega_1=P_r(C_1) \quad (3)$$
$$\mu_0=\sum_{i=1}^{k} i \ P_r(i / C_0) \quad (4)$$
$$\mu_1= \sum_{i=0}^{n} i \ P_r(i / C_1) \quad (5)$$

where

$$\omega(k)= \sum_{i=1}^{k} P_i \quad (6)$$

and

$$\omega(k)= \sum_{i=1}^{k} i \ P_i \quad (7)$$

are the zeroth and the first-order cumulative moments of the histogram up to the kth level, respectively, and

$$\mu_r=\mu (L)= \sum_{i=1}^{k} i \ P_i \quad (8)$$

is the total mean level of the original picture. We can easily verify the following relation for any choice of k

$$\omega_0\mu_0+\omega_1\mu_1+...+\mu_r, \quad \omega_0+\omega_1= 1 \quad (9)$$

The Class variances are given by

$$\sigma_0^2 =\sum_{i=1}^{k}(i - \mu_0)^2 P_r(i / C_0) \quad (10)$$
$$\sigma_1^2 =\sum_{i=k+1}^{k}(i - \mu_1)^2 P_r(i / C_1) \quad (11)$$

These require second-order cumulative moments (statistics).

In order to evaluate the "goodness" of the threshold (at level k), we shall introduce the following discriminant criterion measures (or measures of class separability) used in the discriminant analysis [5]:

$$\lambda= \sigma_B^2/\sigma_W^2, \quad K= \sigma_T^2/\sigma_w^2, \quad \eta = \sigma_B^2/\sigma_T^2 \quad (12)$$

where

$$\sigma_W^2 = \omega_0\sigma_0^2 + \omega_1\sigma_1^2 \quad (13)$$
$$\sigma_B^2 = \omega_0 (\mu_0 - \mu_r)^2 + \omega_1 (\mu_1 - \mu_r)^2$$
$$= \omega_0\omega_1 (\mu_1 - \mu_0)^2 \quad (14)$$

(due to (9)) and

$$\sigma_T^2 = \sum_{i=1}^{L}(i - \mu_T)^2P_i \quad (15)$$

are the within-class variance, the between-class variance, and the total variance of levels, respectively. Then our problem is reduced to an optimization problem to search for a threshold k that maximizes one of the object functions (the criterion measures) in (12).

This standpoint is motivated by a conjecture that well-threshold giving the best separation of classes in gray levels would be the best threshold.

The discriminant criteria maximizing λ, k and η respectively, for k are, however, equivalent to one another; e.g., $k=\lambda+1$ and $\eta=\lambda$ $(\lambda+1)$ in terms of λ., because the following basic relation always holds:

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2 \quad (16)$$

It is noticed that σ_W^2 and σ_B^2 are functions of threshold level k, but σ_T^2 is independent of k.[8] It is also noted that σ_W^2 is based on the second-order statistics (class variances), while σ_B^2 is based on the first-order statistics (class means) [9]. Therefore, η is the simplest measure with respect to k. Thus we adopt η as the criterion measure to evaluate the "goodness" (or separability) of the threshold at level k.

The optimal threshold k^* that maximizes η, or equivalently σ_B^2 maximizes is selected in the following sequential search by using the simple cumulative quantities (6) and (7), or explicitly using (2)-(5):

$$\eta(k) = \sigma_B^2(k)/ \sigma_T^2 \quad (17)$$
$$\sigma_B^2(k) = \frac{[\mu_r \ \omega(k) - \mu(k)]2}{\omega(k)[1-\mu(k)]} \quad (18)$$

and the optimal threshold k^* is

$$\sigma_B^2(k^*) = \max_{1 \le k < L} \sigma_B^2(k) \quad (19)$$

From the problem, the range of k over which the maximum is sought can be restricted to

$$S^* = \{k; \ \omega_0\omega_1 = \omega(k)[1-\omega(k)]>0$$
$$\text{or}$$
$$0< \omega(k)<1\}$$

We shall call it the effective range of the gray-level histogram. From the definition in (14), the criterion measure σ_B^2 (or η) takes a minimum value of zero for such k as $k \in S - S^* = \{ k; \ \omega(k) = 0 \text{ or } 1\}$ (i.e., making all pixels either C_1 or C_0, which is, of course, not our concern) and takes a positive and bounded value for $k \in S^*$. It is, therefore, obvious that the maximum always exists.

Threshold algorithm applied on the dataset which is in the DICOM images that has converted into JPEG images then apply the Threshold level of 75. The value of threshold will set depends on the contrast of an image. The figure1 shows segment the tumor alone from the MRI brain tumor image.

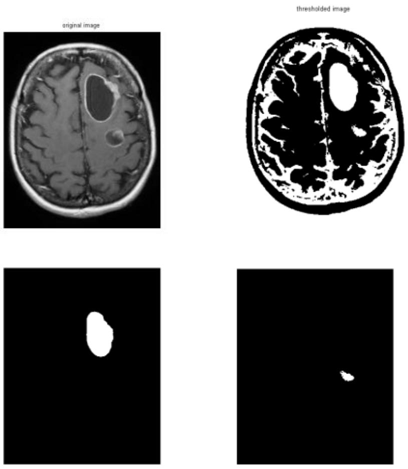


Figure 1 Segmented tumor area by Threshold technique

IV. THE K-MEANS ALGORITHM

Since the simplest form of K-means clustering algorithm, it is used by most of the researchers in the field of data mining. The process of k-Means follows eminent and simple way to classify a given data set by means of a certain number of clusters. The K-means clustering is a method used to divide an image as *n* patterns {x_1, x_2…, x_n} in *d* dimensional space into k clusters (assume k clusters). [10]The result is a set of *k* clusters based on *k* centres, each of which is located at the centroid of the separated dataset. This algorithm can be shortened in the following steps:

The steps involved in clustering the MRI brain images by k-Means algorithm are given below.
Step 1: Insert the original images as input.
Step 2: Convert the fetched MRI DICOM format file into .JPG
Step 3: Cluster dataset images.
Step 4: Find out the 'k' in image by algorithm itself.
Step 5: Get the clustered objects.

The process on clustered images are given below.[11]

Step 1: Give the number of cluster value as k.
Step 2: Randomly choose the k cluster centres.
Step 3: Calculate mean or centre of the cluster.
Step 4: Calculate the distance between each pixel to each cluster centre.
Step 5: If the distance is near to the centre then move to that cluster.
Step 6: Otherwise move to next cluster.
Step 7: Re-estimate the centre.
Step 8: Repeat the process until the centre doesn't move.
The k-means clustering algorithm fixes the k value as 5 in MR image.

In this research work, tumor detection by identifying the pixel values in MRI brain images are taken for analysis. The source code is written in MATLAB software. The k-Means algorithm is applied to find the clusters of MRI images by dividing the image into 5 groups. The various stages of images are given in the Figure 2.

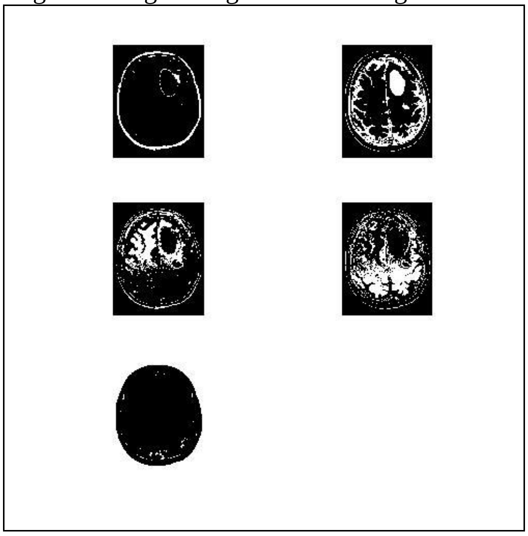


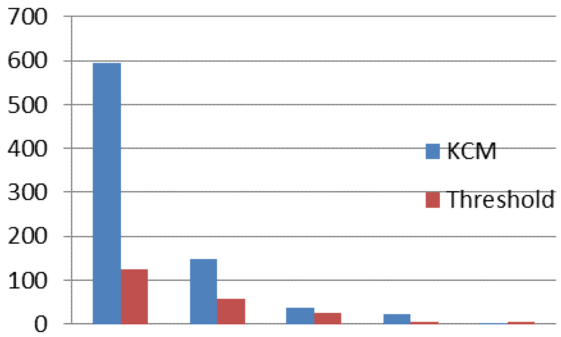
Figure 2. K means algorithm output when k=5

Result of segmented image by *K*-means clustering in Figure 3.

Figure 3 Segmented tumor area by K-means clustering

The figures from figure-1 and figure-3 were calculated and isolated brain tumor from its properties by two methods. The first stage is to determine the area of segmented image. The second stage is to determine the process time, perimeter, minimum axis and maximum axis of the segmented image. Segmented area is automatically calculated from MRI brain images. The segmented images are compare by the following properties.

Methods	*Area*	*Perimeter*	*Major Axis*	*Minor Axis*	*Pocess Time*
KCM	*595*	*147*	*37*	*21*	*3.3252*
Threshold	*124*	*57*	*24*	*6*	*5.5813*



V. CONCLUSION

In this paper, a fully automated tumor detection method based on Threshold and *K*-means clustering techniques are proposed. Threshold technique segments the tumor area by its intensity value in MRI brain images. The K-means clustering gives tumor area by cluster the object which is the methods. The K-means clustering has minimum processing time and also gives an accurate value for the infected area. In future, the entire tumor area identification approach is extendable to 3D to convert into volumetric data. The K-means clustering method is a suitable method to segment brain MR images. This method can also be applied to other medical images e.g., heart or liver MRI.

REFERENCES

[1] M. Sezgin and B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", Journal of Electronic Imaging, Vol. 13, pp. 146-165 , 2004.

[2] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram", IEEE Trans. Systems Man, and Cybernetics, Vol. 9, pp. 62-66, 1979.

[3] Matthew C. Clark "Segmenting MRI Volumes of the Brain With Knowledge- Based Clustering" MS Thesis, Department of Computer Science and Engineering, University of South Florida, 1994

[4] Dzung L. Pham, Chenyang Xu, Jerry L. Prince;"A Survey of Current Methods in Medical Medical Image Segmentation" Technical Report JHU / ECE 99-01, Department of Electrical and Computer Engineering. The Johns Hopkins University, Baltimore MD 21218, 1998.

[5] Osama Abu Abbas "Comparison of various clustering algorithms" The International Arab Journal of Information Technology, Vol. 5, No. 3, July 2008.

[6]. S.Xavierarockiaraj, K.Nithya and R.Maruni Devi, "Brain tumour Detection Using Modified Histogram thresholding-Quadrant Approach", Hournal of Computer Applications(JCA), Vol. 5, No.1, 2012, pp.21-25.

[7]. Anam Mustaqeem, ali Javed and Tehseen Fatima, "An Efficient Brai tumour Detection Algorithm using Watershed and Thresholding based segmentation", International Journal of Image, Graphicss and signal Processing (IJIGSP), Vol. 4, No.10, 2012, pp.34-39.

[8]. Kai Xie, Jie Yang, Z.G.Zhang and Y.M.Zhu, "Semi-sutomated brain tumour and edema segmentation using MRI:, European Journal of Radiology, Vol.56, Issue 1,October 2005, pp.12-19

[9]. Nikola K. Kasabov, Foundations of Neural Networks, Fuzzy systems and Knowledge Engineering, Massachusetts Institute of Technology, 1998, pp.167-473

[10] Jaehui Park, Sang-goo Lee "Probabilistic Ranking for Relational Databases based on Correlations" ACM 2010.

[11] Kanade BP, Gumaste PP. Brain tumordetection using MRI images. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering. 2015; 3(2):146–50.

IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Special Issue-8, Apr 2019 E-ISSN: 2347-2693

A Detailed Survey of Text Line Segmentation Methods in Handwritten Historical Documents and Palm Leaf Manuscripts

R. Spurgen Ratheash[1*], **M. Mohamed Sathik**[2]

[1, 2] Sadakathullah Appa College, Tirunelveli, Tamilnadu, India
Manonmaniam Sundaranar Univesity, Abishekapatti, Tirunelveli, India

Corresponding Author: spurgen@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7si8.99103 | Available online at: www.ijcseonline.org

Abstract— The revolution of document analysis provides the handwritten palm leaf manuscripts and historical documents, epigraphic into digital. Digital document is available as auto recognition of those historical documents which is the second revolution in document research. Many algorithms are available from the period of the 19th century to the 20th century with the researchers across the world. In order to achieve auto recognition, the line segmentation is a foremost process. Though automatic segmentation of text lines is still a burning research, many technical issues remain unsolved yet. The present survey has been carried out as the survey of newly proposed and modified methods of text line segmentation in palm leaf manuscripts and handwritten historical documents published during the period 2008 – 2018. It could benefit the researchers who do research in handwriting recognition.

Keywords— pre-processing, path finding approach, performance measure

I. INTRODUCTION

The digital world provides the potential way to convert handwritten documents into electronically usable data. This study is known as Document Image Analysis (DIA). The olden day documents such as palm leaf manuscripts and handwritten historical documents make DIA a Herculean task by means of strains, yellowing, low intensity variations, random noises, fading and degradation. In DIA, the text line segmentation is considered the most significant process step because the text line segmentation efficiency affects the accuracy of the whole recognition system. It contains line segmentation, word segmentation, character segmentation and text recognition modules. The line segmentation task is extracting and separating the text regions into individual lines. Many of the language scripts such as Japanese, Chinese and Latin have attained a consummate position. However, the Asian scripts pose many issues[1] from fluctuation in the base line, variability in skew between different as well as same line, flexible writing style of different writers, presence of touching components among two adjacent lines. Ceaseless survey is needed to give out many methods reported every decade in text line segmentation. It gives one the motivation to carry out a quality survey after the publishing concerned in 2008.The paper precedes pre-processing in section II, survey of line segmentation methods 2008-2018 given in table format in section III, performance measure in section IV, and conclusion in section V.

II. PREPROCESSING

The line segmentation requires noise free documents. The digitized palm leaf and historical handwritten document are practically impossible to provide noiseless documents. The pre-processing techniques are used to remove the noises and extract the text from their dominating backgrounds. Most of the papers provide the following pre-processing methods:

A. Binarization

Converting RGB or Gray scale valued image into binary image using various threshold techniques[2]. The thresholding provides Global and Local thresholding. Otsu's thresholding method produces best result for text data in document analysis.

B. Skew Correction

The skewed lines of binary images can changed into proper horizontal lines by skew correction. The skew has categorized three types as follow: Global, Multiple, and Non-Uniform Skew[3]. In skew correction the vertical projection profile method is used to rotate the image across different angles. The maximum value of standard deviation is calculated for an optimal rotation in the projection profile.

C. Extracting Connected Components

The edge of the text image calculated using edge detection methods and Stroke Width Transform is applied to identify the strokes in each pixel. The small dots, patches, and unwanted noise components of the images are removed from the text image[4].

III. SURVEY OF LINE SEGMENTATION

A. Adaptive Partial Projection (App)

In 2012, APP is derived from the partial projection method. It divides the image into vertical columns and the projection profile is applied on that column in order to achieve 'smoothing'. It removes peaks and valleys in the histogram of the image. The line extraction process is (i) find the number of lines, divide the image into vertical columns, calculate the horizontal projection profile, smoothen the histogram, find the base lines, (ii) find the valleys of smooth histogram, test all the valleys, check for incorrect top and bottom line, test the number of base lines, test for connected component. APP method is used on palm leaf manuscripts to segment the lines that applied on 264 text lines from 60 palm leaf manuscripts, 129 lines (48.86%) and 177 lines (67.05%) were correctly segmented. The method provides the error when the text is overlapping with the succeeding or preceding lines[5].

B. Matched Filtering

Matched Filtering method consists of three major stages such as (i) Foreground Pixel Density (FPD) – binarize the image using otsu's threshold, extract the skeleton using thinning, form a smallest rectangle contains foreground pixels, randomly select circle regions on that rectangle and compute the mean in all foreground pixels. (ii) Centers of Text Line extraction (CTL) – filter binary image using convolution of G and L to generate filtered image, binarize the filtered image and remove the connected components, perform filling, thinning, removing spurs to get coarse skeleton image(CSI), refine CSI to get centers of text line by using top-down grouping techniques. Overlapping Connected Components (OCC) separation – get the overlapping connected component according to CTL and perform morphology to reduce noises, extract the skeleton, detect all junction points and get the overlapping CC. It is applied on HIT-MW database and produces 99.91% of Detection Rate(DR), 100% of Recognition Accuracy (RA), 99.95% of Performance Measure (FM)[6].

C. Mid Point Detection

The binarized image stored into a matrix, height of the line is calculated, and divide into vertical strips equal to 100 pixels. Horizontal Profile Projection finds white spaces between the two adjacent lines and identifies the midpoint then calculates the difference between adjacent midpoints. The difference is greater than the height of the line that it can be identified as the touching and overlapping lines. The number of lines is calculated by the midpoints and marks the segmentation points. Subsequently, save the matrix and display the output from the matrix. This method is applied on different Gurmukhi script scanned document written by different writers for well spaced document, overlapping and touched lines and it achieves 94% of accuracy in overall performance [3].

D. Thinning

The scanned Bangle image is blurred and scans the vertical column from the top to identify the continuous change of their intensity values. This intensity provides the information of changing the text lines and it suggests the separation point between the two lines. The touched lines and overlapping lines are separated by fixing the junction point in the middle of the touched lines. The junction point fixes above the segmentation point that identifies as the upper line and below the point consider as lower line. The overlapped area is treated as suspected area that are specified by the rectangle and then verified further to fix the segmentation point. The method applied on ICDAR 2013 Bangla documents and it provides 93.16% of Detection Rate (DR), 92.32% of Recognition Accuracy (RA), 92.74% of Performance Measure (FM)[7]

E. Competitive Learning Algorithm

The binary image processed to extract the Connected Components (CC) using two-scan algorithm, common height (H_{cc}), and width (W_{cc}) is calculated using median width and height and center of mass of each CC are calculated. Adaptive Partial Projection technique is used to detect the number of lines in the document. The vertically divided binary image is set to $4W_{cc}$, then the columns are built by Y – Projection profile histogram and smooth twice is used to move the average filter method with the window size of H_{cc} and H_{cc}. The peaks of histogram are identified as the median lines in the column. The competitive learning algorithm is used to identify the midpoint of the line. The method is applied on EFEO database and it provides 99.528% precision, 99.534%recall and 99.531% Performance Measure (FM)[4].

F. Fully Convolutional Network

The dense layers are removed from the Convolutional Neural Network forms FCN to process the image in variable size. Instead of decision taking FCN work as encoder and decoder. This method provides the solution for semantic segmentation problem. In this approach, the pooling layer is used to reduce image resolution in order to decrease the computation count and increase the filter size. In dilated convolution "A trous" algorithm is used with wavelet transform. It provides two advantages; size of the receptive field can be controlled without reducing the resolution not increasing the number of parameters. Secondly, reduce the number of parameters and depth of the network. The method is applied on ICDAR 2017 with 1600 images contains 10000 lines, 1500 images for training and 100 for validation. It also applied on cBAD dataset with 755 pages, 216 for training and 539 in testing. It provides the result of 0.66 Precision, 0.86 Recall, 0.75 Performance Measure (FM)[8].

G. Second Order Derivative Analysis

The text lines are blurred and it appears as a blob by Gaussian function. It gives an advantage of steerable that means the response of the filter which can be calculated as base filters. After processing all the scales, one gets line orientation, scale and strength of each pixel. Finally, one considers the pixels which have the strongest response within their line. Compare the neighbors which belong to the same line with the small difference between their orientation and scales. After further filtering, apply non-maxima suppression approach in selected ridge. When it falls within the area of other local maxima, which has higher filter response, the pixel with mild-low value is marked as red. This method applies on IAM database, and many dataset and it provides 99.8% Precision, 93.1% Recall and 91.6% Performance Measure(FM)[9].

H. Path Finding Approach

To extract the text components from palm leaf manuscripts, the edges of the image are calculated by Canny Edge detection and Stroke Width Transform. X –projection profile is used to identify the starting and ending point of the lines, Y – projection profile for each column to smooth the histogram. Further, the boundary of the line is identified by A* algorithm method and two cost functions are described such as Intensity difference cost function and Vertical cost function. Then, the image is compared with Ground Truth value. This method is applied on 100 pages of Khmer palm leaf manuscripts called Sleuk Rith Set collected form EFEO database and it produces the results of 92.15% Detection Rate(DR), 93.70% Recognition Accuracy, 92.92% Performance Measure(FM)[4]

IV. ANALYSIS ON PERFOMANCE MEASURE

In the survey of line segmentation from historical handwritten document and palm leaf manuscripts has two categories of performance measures. First, the result has the parameter of

Precision, Recall and Performance Measure. The Precision and Recall are specifically used for Information Extraction. Precision gives the result of number of document retrieved that relevant and Recall is the number of relevant document that retrieved. Precision and Recall are inversely proportional to each other and understanding of this difference is much more important to build an efficient classification system. For example, among the 15.6 million results the relevant links to my question is 2 million. 6 million of results were relevant but not produce by the particular search engine. Here, Precision calculated by 2M/15.6M = 0.13 that means all the retrieved links were relevant and Recall is calculated by 2M/8M = 0.25 that means retrieve all the relevant links[10]. The Precision and Recall is explained by the following mathematical formulae;

Precision P = TP/(TP+FP)

Here, True Positive that gives the number of correctly retrieved documents divided by the Total number of document retrieved.

Recall R = TP/(TP+FN)

Here, True Positive that gives the number of correctly retrieved documents divided by the Total number of relevant document retrieved.

The combination of Precision and Recall produces the Performance Measure (F – measure)

*F=2 * ((Precision * Recall) / (Precision + Recall)*

Second, the result provides the parameter of Detection Rate (DR), Recognition Accuracy (RA) and F- Measure. The text lines detected by the way of one-to-one matches between text lines detected detection in historical handwritten documents and palm leaf manuscripts. The efficiency measures by the F- measure values.

approach and the text line in the ground truth. The matching score is computed as

Match Score (i,j) =

where T(s) is function that counts the number of points in set S; $G_j$is the set of all points inside the union of all selected regions of isolated characters in ground truth belonging to text line j; $R_i$the set of all points inside the region of result text line i; and I_{IC} the set of all points inside the union of all selected regions of ground truth isolated characters in the whole document page[4][11][12].

The region pair is considering by the match score is above to the acceptance threshold T . Assume that N is the number of text lines found in the ground truth, M to be the number if text lines by the approach, and o2o to be the number of one – to – one match pairs. The detection rate is defined by,

DR = o2o/N

and recognition accuracy is

RA = o2o/M

the evaluation metric F-measure score is calculated by combining the above two relation

FM = 2.DR.RA/DR+RA

In analysis, among the two methods the second method of result parameters; Detection Rate (DR), Recognition Accuracy (RA) is the correct way to measure the text line

V. CONCLUSION

Text line segmentation is a Herculean task. This is because the method of lines is under the condition of skewed, multi skewed, well spaced, touching lines and overlapping lines. The Viterbi algorithm concentrates on skewed lines with high variability, skeletonization algorithm methods that are highly supported for multi skewed text lines. The fringe map method supports for the Telugu document text lines and the adaptive partial projection method is used to segment the touching characters in the document. The projection profile methods are highly supportive of the skewed line in the historical handwritten documents. Fully Convolution Network method provides the segmentation method for connected characters and overlapping characters. Table-1 provides a detailed survey of line segmentation and its limitations. The present survey analyzes the various line segmentation methods presented during the last decade. In all the algorithms, the results for the touching, overlapping text lines, skewed lines in the handwritten text documents and palm leaf manuscripts are analyzed. The survey concludes that there is no successive algorithm to segment the touched and overlapped lines in the Tamil palm leaf manuscripts. There is a vast space to be filled up with an efficient method to segment the Tamil palm leaf manuscript's touched and overlapping lines in the future.

Table -1: Survey of Text line segmentation methods during 2008 - 2018

Sl.No	Year	Authors	Proposed Method	Database	Applied for	Drawbacks	Text Languages
1.	2008	Themos et al[13]	Viterbi Algorithm	ICDAR 2007	Skew with high variability, Non strict left and right margins	Minimum efficiency to detect high variability of writing styles, sizes	English, French, German and Greek
2.	2008	Sanchez et al[14]	Skeletonization algorithm	PROHIST project data base	Multi skewed text lines	Not efficient in overlapped text lines	English
3.	2009	Rodolfo et al[15]	Morphology and Histogram Projection	IAM handwritten Data base	Handwritten document	Minimum Efficiency in upper character in the text line	English
4.	2010	Rajiv kumar et al[16]	Variable size windowing	Handwritten Gurumukhi document	Handwritten Gurumukhi Script	Not much efficient when the characters are combined in nature	Gurumukhi
5.	2010	Naresh kumar Garg	Block covering	Handwritten document	Handwritten	Not efficient in overlapping,	Bangla,

		et al[17]	method	from 15 writers	Devanagari, Hindi	broken parts and thick parts in upper modifiers	Devanakari, Telugu
6.	2010	Jija Das Gupta et al [18]	Handwritten text line segmentation approach	IAM database	Offline English letter	Not much efficient in overlap and touching component	English
7.	2011	Vijaya Kumar et al[19]	Fringe Map based method	Handwritten Telugu Character	Printed Telugu Script document	Minimum efficiency in non constant space exists between lines	Telugu
8.	2011	Rajib et al[20]	Skew Detection Algorithm	Handwritten Bangla words	Skewed Bangle words	Not efficient in some part of the word is skewed and the rest is not skewed	Bangla
9	2012	Ines Ben et al[21]	Multi level framework	IAM – HistDB Database, ICFHR 2010	Touching elements as well as skewed text lines	Minimum Efficiency in overlapping characters	English
10.	2012	Rapeeporn et al[5]	Adaptive Partial Projection	Thai palm leaf manuscripts from Mahasarakham University, Northeaster Thailand	Touching character in consecutive lines	Minimum Efficiency in vowels which are too close to the upper or lower line, long prolonged components, and connected components of consecutive lines.	Thai
11.	2012	HandeAdiguzel et al[22]	Connected component and projection profile	Ottoman printed book	Well Spaced lines	Minimum efficiency in small sized components placed in wrong lines	Ottoman documents
12.	2014	Xi Zhang et al[23]	Seam carving	ICDRA2013 Handwritten segmentation contest dataset	Minimum touched and well space lines	Inefficiency in large components which touch multiple text lines	English and Greek
13.	2014	Rapeeporn et al[24]	Combined method for segmentation	Thai palm leaf manuscripts	Text line segmentation	Error in touching characters, noise surrounded characters, incorrect line separation	Thai
14.	2014	Youbao Tang et al[6]	Matched Filtering and Top down Grouping Method	ICFHR 2010, HITMW, Contest database	Text line extraction	Inefficiency in overlapping lines	English, French, Chinese, German, Greek
15.	2015	Payal Jindal et al[3]	Midpoint Detection Technique	Gurumukhi Handwritten document	Skewed lines, overlapped lines, Connected components	Not efficient for complex overlap text lines	English, Gurumukhi
16.	2015	Mullick et al[4]	Thinning	ICDAR2013	Text line segmentation	Not efficient in skew lines, overlap and low space	English, Greek, Bangla
17.	2016	Dona Valy et al[25]	Competitive Learning Algorithm	EFEO data base	Skewed and fluctuated lines	Not efficient in cursive characters and need improvement in touching components	Khmer
18.	2016	Banumathi et al[26]	Projection Profile Technique	Kannada handwritten document	Skewed lines and spaced lines	Not efficient in Low space lines, Overlapping lines	Kannada
19.	2016	Quang Nhat Vo et al[27]	Fully Convolutional Network (FCN) and Line Adjacency Graph (LAG)	ICDAR 2013 Handwritten Segmentation Contest dataset	Touching Characters	Not efficient in complex touching characters, Error in Indian documents where the characters vertically connect to the text lines.	English, Greek, Bangla
20.	2017	Himanshu Jain et al[2]	Bottom up procedure	ICDAR2009 Text segmentation Contest Dataset	Touching and overlapping character	Not efficient while using minimum artificial parameters on overlapping characters	Greek, French, German, English
21.	2017	Guillaume et al[8]	Fully Convolutional Network	ICDAR 2017 and cBAD , ICDAR 2015 and ANDAR, RIMES Dataset	Well spaced lines	Not efficient in touching and overlapping characters	English
22.	2017	Dona Valy et al[7]	Path Finding Techniques	EFEO database National Library Budhist Institute	Skew, fluctuated, discontinued text lines	Not efficient in touched and overlapped characters	Khmer
23.	2017	Kathirvalavakumar et al[28]	Projection Profile and Connected component	Tamil printed document	Skew slant lines	Not efficient in handwritten document, Touching and overlapping characters in above	Tamil

			Techniques			and below text lines	
24.	2018	David Et al[9]	Second order derivatives	IAM, GRPOLY-DB, ICDAR 2009, cBAD data set	Skewed Lines	Not efficient in Overlapping characters	English

REFERENCES

[1] Made Windu Antara Kesiman, Dona Voly., "Benchmarking of Document Image Analysis Tasks for Palm Leaf Manuscripts from Southeast Asia",*JImaging,4,43;doi:10.3390/jimaging 4020043,* 2018.

[2] Himanshu Jain, Archana Praveen Kumar, "A Bottom Up Procedure for Text Line Segmentation of Latin Script", IEEE, *pp 1182 – 1187,* 2017.

[3] Payal Jindal, Dr. Balkrishnan Jindal, "Line and Word Segmentation of Handwritten Text Documents written in Gurmukhi Script using Mid Point Detection Technique", IEEE, *Proceedings of 2015 RACES UIET Panjab University Chandigarh 21-22nd December2015,* 2015.

[4] Dona Valy, Michel Verleysen, Kimheng Sek, "Line Segmentation for Grayscale Text Images of Khmer Palm Leaf Manuscripts", IEEE, 2017.

[5] Rapeeporrn Chamchong, Chum Che Fung, "Text Line Extraction Using Adaptive Partial Projection for Palm Leaf Manuscripts from Thailand", IEEE, *2012 International Conference on Frontiers in Handwriting Recognition, IEEE, PP 586-5591,* 2012.

[6] Youbao Tang, Xiangqain Wu, Wei Bu, "Text Line Segmentation Based on Matched Filtering an Top-down Grouping for Handwritten Documents", IEEE, *2014 11th IAPR International Workshop on Document Analysis System, PP 365 – 369,* 2014.

[7] Mullick, Banerjee, Bhatacharya, "An Efficient Segmentation Approach for Handwritten Bangla Document Image, IEEE, 2015.

[8] Guillaume Renton, Clement Chatelian, "Handwritten text line segmentation using Fully Convolutional Network", IEEE, *2017 14th International Conference on Document Analysis and Recognition, PP 5 – 9*, 2017..

[9] David Aldavert, Marcal Rusinol, "Manuscript Text Line Detection and Segmentation using Second Order Derivatives", IEEE, *2018 13th IAPR International Workshop ON Document Analysis Systems, PP 293 – 298*, 2018.

[10] https://towardsdatascience.com/model evaluation-i-precision-and-recall-166ddb257c7b

[11] Vishal Chavan, Kapil Mehrotra, "Text Line Segmentation of Multilingual Handwritten Documents Using Fourier Approximation", IEEE, *2017 Fourth International Conference on Image Information Processing(ICIIP), PP 250 – 255,* 2017.

[12] Ayush Padhan, Sidharth Behra, Paushpalata Pujari, "Comparative Study on Recent Text Line Segmentation Methods of Unconstrained Handwritten Scripts", IEEE, *International Conference on Energy, Communication, Data Analysis and Soft Computing(ICECDS-2017), PP 3853 – 3858*, 2017.

[13] Themos Stafylakis, Vasilis Papavassiliou, "Robust Test-Line and word segmentation for Handwritten Documents Images", IEEE, *PP 3393 – 3396*, 2008.

[14] Sanchez, Suarez, Mello, Oliveria, Alves,"Text Line Segmentation Images of Handwritten Historical Documents", IEEE, *Image Processing Theory, Tools and Applications*, 2008.

[15] Rodolfo P dos Santos, Gabriela S Clemente, Tsang Ing Ren, "Text Line Segmentation Based on Morphology and Histogram Projection", IEEE, *2009 10th International Conference on Document Analysis and Recognition PP 651 – 655*, 2009..

[16] Rajiv Kumar, Amardeep Singh, "Detection and Segmentation of Lines and Words in Gurmukhi Handwritten Text", IEEE, *2010 IEEE 2nd International Advance Computing Conference, PP 353 – 356*, 2010.

[17] Naresh Kumar Garg, Lakswinder Kaur, Jindal, "A new method for Line Segmentation of Handwritten Hindi Text", IEEE, *2010 Seventh International Conference on Information Technology, PP 392 – 397,* 2010.

[18] Jija Das Gupt, Bhabatosh Chanda, "A model based Text Line Segmentation method for Off-Line Handwritten Documents", *IEEE, 2010 12th International Conference on Frontiers in Handwritting Recognition, pp 125-129*, 2010.

[19] Vijaya Kumar Koppula, Atul Negi, " Fringe Map based Text Line Segmentation of Printed Telugu Document Images", IEEE, *2011 International Conference on Document Analysis and Recognition, PP 1294 – 1298,* 2011.

[20] Rajib Ghosh, Debnath Bhattacharyya, Tai hoon kim, Gang soo Lee, " New Algorithm for Skewing Detection of Handwritten Bangla Words", *Springer – Verlag Berlin Heidelberg, PP 153 – 159,* 2011.

[21] Ines ben Messaoud, Hamid Amiri, "A Multilevel Text Line Segmentation Framework for Handwritten Historical Documents", *IEEE, 2012 International Conference on Frontiers in Handwriting Recognition, PP 515-520,* 2012.

[22] Hande Adiguzel, Emre Sahin, Pinar Duygulu, "A Hybrid Approach for Line Segmentation in Handwritten Documents", *IEEE, 2012 International Conference on Frontiers in Handwriting Recognition, PP 503-508,* 2012.

[23] Xi Zhang, Chew Lim Tan, " Text Line Segmentation for Handwritten Documents Using Constrained Seam Carving", *2014 14th International Conference on Frontiers in Handwriting Recognition, IEEE, PP 98-103,* 2014.

[24] Rapeeporrn Chamchong, Chum Che Fung, "A Combined Method of Segmentation for Connected Handwritten on Palm Leaf Manuscripts",IEEE, *2014 IEEE International Conference on Systems, Man and Cybernetics, PP 4158 – 4161*, 2014.

[25] Dona Valy, Michel Verleysen, Kimheng Sek, "Line Segmentation Approach for Ancient Palm Leaf Manuscripts using Competitive Algorithm", IEEE, *2016 15th International Conference on Frontiers in Handwriting Recognition, PP 108-113,* 2016.

[26] Banumathi, Jagadeesh Chandra, " Line and Word Segmentation of Kannada Handwritten Text documents using Projection Profile Technique", IEEE, *2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT) PP 196 – 201*, 2016.

[27] Quang Nhat Vo, GueeSang Lee, " Dense Prediction for Text Line segmentation in Handwritten Document Images", IEEE, *ICIP 2016, PP 3264 – 3268*, 2016.

[28] Kathirvalavakumar Thangairulappan, Karthigai selvi Mohan, " efficient Segmentation of Printed Tamil Script into Characters Using Projection and Structure", IEEE, *2017 Fourth International Conference on Image Information Processing(ICIIP), PP 484 – 489*, 2017.

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-9 Issue-1, November 2019

Line Segmentation Challenges in Tamil Language Palm Leaf Manuscripts

R. Spurgen Ratheash. M. Mohamed Sathik

Abstract: ***The process of an Optical Character Recognition (OCR) for ancient hand written documents or palm leaf manuscripts is done by means of four phases. The four phases are 'line segmentation', 'word segmentation', 'character segmentation', and 'character recognition'. The colour image of palm leaf manuscripts are changed into binary images by using various pre-processing methods. The first phase of an OCR might break through the hurdles of touching lines and overlapping lines. The character recognition becomes futile when the line segmentation is erroneous. In Tamil language palm leaf manuscript recognition, there are only a handful of line segmentation methods. Moreover, the available methods are not viable to meet the required standards. This article is proposed to fill the lacuna in terms of the methods necessary for line segmentation in Tamil language document analysis. The method proposed compares its efficiency with the line segmentation algorithms work on binary images such as the Adaptive Partial Projection (APP) and A* Path Planning (A*PP). The tools and criteria of evaluation metrics are measured from ICDAR 2013 Handwriting Segmentation Contest.***

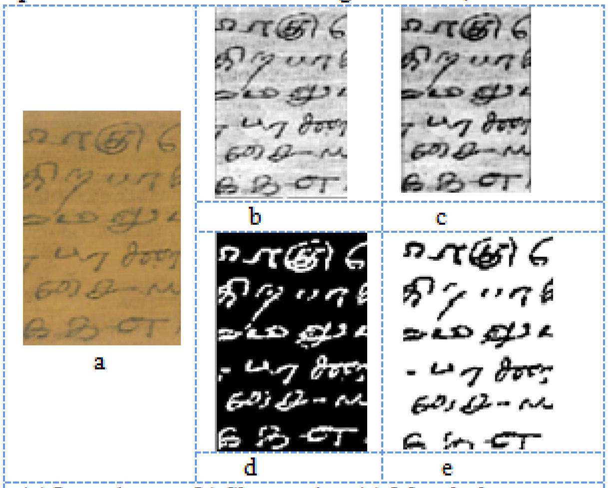
Keywords: line segmentation, Tamil palm leaf manuscripts, connected component, historical documents, Tamil character recognition.

I. INTRODUCTION

In digitizing palm leaf manuscripts, there are various challenges in terms of reading and understanding the scripts. Only scholars with knowledge in old scripts could read and understand the palm leaf manuscripts. However, reading this poses great challenge for the general people and researchers concerned.

Those eminent scholars who could read the palm leaf manuscripts have incorporated the information in the palm leaf manuscripts in printed form as books. In spite of the effort taken so far in preserving the palm leaf manuscripts and the information in it, there is still a long way to go in accumulating the entire essential information from the ancient archives. In order to grasp all the information from palm leaf manuscripts, it is of vital importance to recognize the information from the manuscripts in question.

Automatic recognition is a process that helps in reading the scripts with suggestions when the scripts are not recognizable. It is important to implement automatic recognition of the palm leaf manuscripts as the structure or form of the modern day letters is different from that of the ancient scripts. Optical Character Recognition is a process that recognizes the character in the ancient scripts automatically with high accuracy. Basically an OCR has five major phases such as 'preprocess', 'line segmentation', 'word segmentation', 'character segmentation' and 'character recognition'. The first and foremost phase Pre-process does the initial work such as noise removal, morphology, binarization to separate the dark background and text foreground as shown in Fig.1.



(a) Input image (b) Sharpening (c) Morphology
(d) Binarization (e) Output image

Fig. 1 Preprocessing phases

The second phase, line segmentation segment the text lines if it is touching and overlapping with the subsequent lines. The third phase, separate the words according to the language from the segmented lines. The character segmentation separates the characters from the words and leads to the final phase of character recognition. Amongst the umpteen of methods for line segmentation, this article concentrates on two methods such as Adaptive Partial Projection (APP) and A* Path Planning (A*PP) which are considered the best in Thai and Khmer scripts respectively. The process of those methods is implemented in Tamil palm leaf manuscripts and it compares the result with the proposed segmentation method.

In section II provide the details of widespread line segmentation methods used in various languages by Literature survey. The comparison of Thai and Khmer language manuscripts line segmentation methods with Tamil language manuscripts explains in Section III, and the evaluation results in section IV with the conclusion in section V.

Revised Manuscript Received on November 05, 2019.

R. Spurgen Ratheash, Assistant Professor of Information Technology, received his MCA degree in Computer Applications from Bishop Heber College, Bharathidasan University, Trichy, India in 2007.

M. Mohamed Sathik, PG and Research Department of Computer Science, Sadakathullah Appa College, Tirunelveli, Tamilnadu, India. Manonmaniam Sundaranar University, Tamilnadu, India.

II. LITERATURE SURVEY

The text line segmentation is performed in three ways: the connected component estimation method is used to group the component for top-down method, vertical projection provides bottom-up approach, nearest neighbour detection is used in the third method on historical documents [1]. Connected components are calculated and classified large components that identify the long ascenders or descenders. The later connect multiple lines and smooth the histogram with Gaussian Kernel with the mean of its height. Further, it derivates and removes the lower intensity components than threshold in historical documents [2][3]. The projection profile partitioning is used to segment the lines by separating the Vertical strips and Horizontal run for each strip of an image [4][5]. The Hidden Markov Model is built by dividing the text line image and apply the Viterbi algorithm to detect possible paths to segment the lines [6][7]. The competitive learning algorithm applied on center of mass of y-coordinates derived from 1-D vectors [8]. The connected components are labelled to form a bounding box around the text with the measure of height and width. The threshold height value is fixed by mean and deviation and is compared with the height of the text area and the greater value has broken into two lines [9]. The Fringe Map is generated on the binary images which lead to identify Peak Fringe Numbers (PFN). All the identified PFNs are put together to separate the text lines [10]

III.COMPARATIVE STUDY

Many of the algorithms are efficient to segment the lines in historical documents and palm leaf manuscripts of Thai, Khmer, and Hindi languages. However, umpteen of algorithms are available to segment the lines in Tamil language printed documents incompetent for Tamil palm leaf manuscripts. This chapter summarizes the comparison of two efficient line segmentation methods in Thai and Khmer manuscripts with the proposed method in Tamil manuscripts.

A. ADAPTIVE PARTIAL PROJECTION

In Adaptive Partial Projection line segmentation method for Thai language manuscripts, the line numbers, average line position, average height of the line is calculated from the horizontal projection profile that derived from the piece wise projection method [11]. The peaks are identified by applying the histogram on the image and defined the number of lines as against the peaks. The image divides into vertical column. In each column, the horizontal projection profile is calculated and the average filtering is applied more than one time to smooth the histogram as well as eliminate the spurious peaks and valleys in the projection. The lowest value between two peaks is known as valleys that define the base line of characters placed in the column. The segmentation line is formed by joining all the characters' base line. If a base line overlaps one or more connected components, the column is divided into two and processed further when the width of the column is less than the width of the character. [12].

B. A* PATH PLANNING

A* Path Planning has been applied in artificial intelligence to find the path for robotic system, route planners, and games. A* Path Planning uses heuristic function for quick computation to find out an optimal solution to reach the end state. A*PP algorithm cannot reach the end when all sides are covered by obstacles. In handwritten historical documents of Khmer language, the path identification is effective when two lines do not touched or overlap. An algorithm malfunctions when are touched or overlapped two lines. It is solved by cost functions in the path planning algorithm such as distance cost function, map obstacle cost function, vertical cost function, the path planning is successful in line segmentation only and neighbour cost function. In spite of A*PP has the above-mentioned cost functions the path planning is successful in line segmentation only when they are partially overlapped in the handwritten historical documents [13].

C. PROPOSED METHOD

Line segmentation in Tamil palm leaf manuscripts is a Herculean task. In order to simplify the process, the text lines are partitioned three zones according to their characteristics namely 'text zone', 'upward elongated zone', and 'downward elongated zones' as in Fig. 2.

Fig. 2. Text line characteristics

Text zone is an actual character existing zone [14]. The strokes of Tamil letters, in the course of writing, go beyond the text zones such as 'upward' and 'downward' elongated zones.

An extension from the text zone by upward or/and downward characters in Tamil palm leaf manuscript is known as an 'obstacle'. It is categorised by 'presence of' and 'absence of' an obstacle. The first category makes the line segmentation as challengeable because considering other language manuscripts, the presence of obstacles can be identified only in vowels or in consonants. In Tamil language manuscripts, the presence of an obstacle may exist in any characters. Prior to the line segmentation process, the presence of an obstacle in text lines are classified in two ways by their nature of extension. If it touches with the subsequent lines, they are called 'touched obstacle text lines' and when the obstacle reaches up to the middle of the following or succeeding lines, they are defined as 'overlapped obstacle text lines' as shown in Fig.3. The second category is known as 'independent text lines' that segments the lines automatically without any complication by an identification of the character using connected component.

Fig. 3.	Touched obstacle text line	Overlapped obstacle text line	Independent text line

The APP method described in section 3.1 is one of the best methods for Thai language palm leaf manuscripts. This provides accuracy in independent characters that are placed in the text line. If the lines are touched and there is less space between the two lines, the result of APP line segmentation provides a by and large result. The A*PP method in section 3.2 is one of the best heuristic ways of approach in line segmentation of handwritten documents.

The path planning algorithm is applied on Bali, Sunda, Khmer palm leaf manuscripts provide a better result in ‘touched characters’ with minimum time. Both the algorithms are applied on binary images provide notable results in touched lines. Considering the overlapping lines, the above-mentioned methods are not effective as such. When the above-said methods are applied in Tamil palm leaf manuscripts, the result of APP is fairly reported in the image of independent scripts as in Fig.4. The character ‘ளு/Lu/’, ‘ள/La/’, and ‘ள/La/’ are placed first, second and third lines of the image respectively. The APP method segments the first line as three line spaces such as S1, S2, and S3. It furthers the cutting edge that breaks and changes the structure of the character ‘ளு/Lu/’ as ‘ள/La/’. The ‘cutting edge’ is defined as the point where the lines sever from its subsequent lines. S4 denotes the line space between the second and the third text lines. An ‘absence of’ obstacle in APP provides fair results. Besides, the method is not much effective in touched and less line space images of Tamil palm leaf manuscripts. The A*PP method provides good results with the independent lines as shown in Fig.5.

Fig. 4. a. Original Image **b. APP Line segmentation**

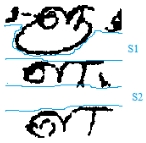


Fig. 5 A*PP Line segmentation for independent lines

However, in touched lines of Tamil palm leaf manuscripts, the path planning falls short in reaching the end state. The cutting point of an algorithm changes the characters structure. It also creates a continuous process without reaching the end state when more than two lines are touching each other in Fig.6. The character ‘தை/thy/’, ‘ந/na/’ and ‘த/tha/’, ‘யெ/ye/’ are in the first, second, third lines respectively. ES1 specified the ‘End State’ of the first line and ES2 for the second. S1 and S2 are the line space between the text lines. The A*PP segmentation proceeds the first line character ‘thy’ can segment without any difficulties. Considering the second line, the path planning creates a loop and it reaches the end state ES2 without segmenting the text line. If an APP proceeds, the cutting edge segments the line and changes the character ‘ந/na/’ into unpredictable structure and ‘த/tha/’ into wrongly predicted ‘க/ka/’.

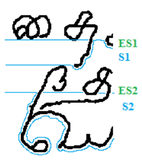


Fig.6 A*PP for Touched obstacle text lines

The proposed method provides one of the best results in three ways such as independent text lines, touched and overlapped obstacle text lines. The Connected Component (CC) algorithm analyses the quality of extracted group and the quantitative measures of grouping quality within certain constraints. It relates to a random image model, and set figure and background. The probability of the features of the image belongs to the figure is P_f and the background is P_{bg}. The process considers binary cues. Value 1 is assigned where P_f and P_{bg} are in the same group, when otherwise it is 0. The probability of finding the connected point in the background pixel is $P_c(L)$ with the distance L, that may be connected with the figure and the background. The density of such points implies with the error probability E_f [14].

$$P_{bg}^{(1)}(L) = P_{bg}\left[1-\sum(1-E_f)^m P_f^m (1-P_f)^{k(L)-m} C_{k(L)}^m\right]$$

The total number of connected features is

$$N_f^{(i)} = N_{bg} N_f^{(i-1)} \qquad N_{bg} < 1$$

The presence of character in the text lines are measured by $N_f^{(i)}$ through the binary cues of an image. In image matrix, when the value is drastically changed and extends with minimum value, it is defined as an obstacle of the character. If an extended minimum value ends with greater value that identified as touched obstacle text lines otherwise independent text lines. The process also predicts an overlapped obstacle text lines when the value is maximum than all other characters are present in the text lines. The proposed method step by step process algorithmic way in the following Algorithm.1

Algorithm.1

```
// Input Image
Input:
    InImg  :  Binary Image
//Output Image
Output:
    OutImg  : Line Segmented Image
//Variables
VT ← Vertical Space Track
HT ←  Horizontal Space Track
M  ←  InImg Width
N  ←  InImg Height
//Sum of column to calculate Vertical space
for ( k = 1; k<= N; k++) do begin
    Vs  ←  sum of zero from InImg(k, 1:20) Location
    if (Vs > Threshold)
            VT k = assign 1
    else
            VT k = assign 0
    end
end for
//Sum of row to calculate Horizontal space
for ( k  = 1; k<= N; k++) do begin
    if (VT k ==1)
            Hs  ←  sum of zero from InImg(k, 1:M)
    if (Hs ==0)
            HT k = assign 1 (Space Found)
    else if (Hs < Threshold)
            HT k = assign 2 (Obstacle Found)
    else
            HT k = assign 3 (space Not Found)

    end
    end
end for
OutImg ← HT k
```

35. Return (Out_{Img})

In image matrix, sum of the value k presents in the row to predict whether the background or character exists within the 20 positions against N and assign the sum value into Vs. The vertical space identifies by zero which is otherwise identified as the horizontal space by the same way against M and assign the sum value into Hs as shown in Fig. 7.

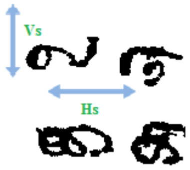


Fig. 7. Vertical and Horizontal space

The vertical space identifies by VT_k, horizontal space HT_k to predict three categories such as 'space between two lines', 'obstacle is present in the space' and 'no space'. An optimality of the performance for touched and overlapped obstacle text lines of Tamil palm leaf manuscripts are implemented by fixing the cutting edge in the place of minimum value exists in HT_k with the consideration of VT_k between the text lines.

IV. EXPERIMENTAL EVALUATION

The line segmentation algorithms specified in section 3 is applied on 750 lines of the Tamil palm leaf manuscripts of 225 images. This matches the resultant images with the ground truth images to evaluate the metrics of DR, RA, and FM based on MM, NN, and o2o. These evaluation criteria and tools are provided by ICDAR 2013 Handwritten Segmentation Contest.

An evaluation of this article by the metrics of one-to-one (oTo) match score, NN and MM [15]. oTo is computed for a region pair based on the evaluator's acceptance threshold. Let NN be the total number of ground truth elements and MM be the total number of result elements. An above said three metrics are calculated with the oTo score. The Detection Rate (DR) is defined by

$$DR = \frac{o2o}{NN}$$

Recognition Accuracy (RA) is

$$RA = \frac{o2o}{MM}$$

and F-measure (FM) is calculated by

$$FM = \frac{2.DR.RA}{DR+RA}$$

In Tamil palm leaf manuscript images the text lines have the challenges such as Low contrast (LC), Damaged Letters (DL), Overlapped Lines (OL), Low Space (LS), Cross Lines (CL), Touched Lines (TL), and Independent Lines (IL). The performance on various challenges of text lines are shown in Table.1 – Table. 3. The performance metrics of DR, RA, FM on Hundred various kinds of images shows in Fig.8 – Fig.10.

V. CONCLUSION AND FUTURE WORK

The present article sheds light on the line segmentation work of 'touched' and 'overlapped text lines' of Tamil palm leaf manuscripts. In this article, the comparison of line segmentation methods are working on binary images among umpteen of methods. The proposed method provides optimality about the ins and outs of line segmentation in the Tamil palm leaf manuscripts with 95% of Recognition Accuracy while APP and A*PP provides 51% and 89% from the same value of NN. The Table.4 provides an overall performance of evaluation metrics with the line segmentation methods such as APP, A*PP and Proposed method. The structure of the Tamil character may break when the proposed line segmentation is implemented on Tamil palm leaf manuscripts. All said and done, the proposed method is successful in bringing out the very existence of lines.

Table 1 Performance of DR in different kind of challenges

DETECTION RATE (DR)	LC	DL	OL	LS	CL	TL	IL	OVERALL
APP $_{DR}$	30.64	51.79	35.66	37.18	36.88	44.23	40.66	39.57
A*PP $_{DR}$	66.66	87.53	83.89	78.67	87.42	85.43	84.56	82.02
PROPOSED $_{DR}$	70.63	89.33	91.54	85.94	91.09	88.94	92.25	87.1

Table 2 Performance of RA in different kind of challenges

RECOGNITION ACCURACY(RA)	LC	DL	OL	LS	CL	TL	IL	OVERALL
APP $_{RA}$	45.7	65.16	49.69	51.02	51.97	60.12	56.98	54.38
A*PP $_{RA}$	79.17	86.15	89.04	85.28	87.45	88.9	90.34	86.62
PROPOSED $_{RA}$	98.86	90.69	94.64	94.73	91.72	93.68	96.46	94.4

Table 3 Performance of FM in different kind of challenges

F-MEASURE (FM)	LC	DL	OL	LS	CL	TL	IL	OVERALL

APP $_{FM}$	36.63	57.68	41.39	42.91	43.09	50.92	47.42	45.71
A*PP $_{FM}$	72.33	86.74	86.37	81.75	87.36	87.12	87.34	84.14
PROPOSED $_{FM}$	82.14	89.94	93.05	89.6	91.31	91.14	94.27	90.2

Table 4. Performance of evaluation metrics

	ALGORITHMS		
Metrics	**APP**	**A * PP**	**Proposed**
NN	27058	27058	27058
MM	18586	25650	25485
oTo	9762	22895	24347
DR	35.94	84.64	90.12
RA	50.68	89.05	95.46
FM	41.96	86.73	92.64

Fig. 9. Image wise performance of RA

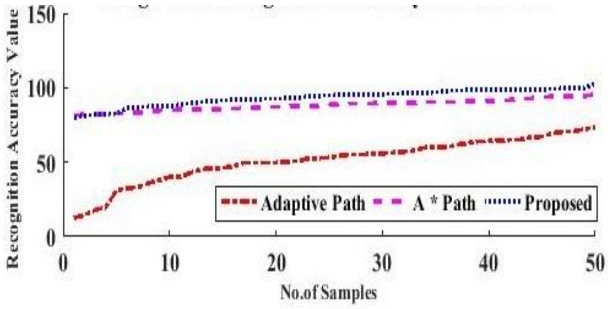


Fig. 8. Image wise performance of DR

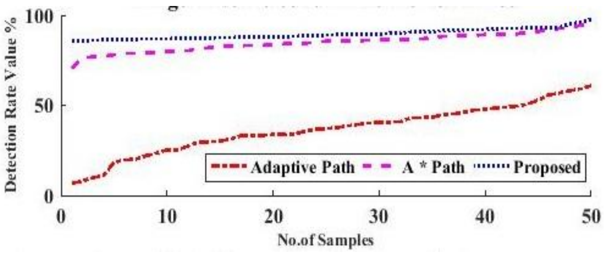
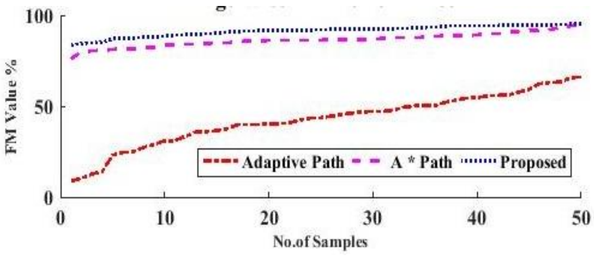


Fig. 10. Image wise performance of FM



REFERENCES

1. Ines Ben Messaoud, Hamid Amiri, Haikal El Abed, Volker Margner "A Multilevel Text line Segmentation Framework for Handwritten Historical Documents", ICFHR, IEEE, pp. 515-520, 2012.
2. Xi Zhang, Chew Lim Tan "Text Line Segmentation for Handwritten Documents Using Constrained Seam Carving", ICFHR, IEEE, pp. 98-103, 2014.
3. MadeWindu Antara Kesiman, Dona Valy, Jean-Christophe Burie, Erick Paulus, Mira Suryani, Setiawan Hadi, Michel Verleysen, Sophea Chhun and Jean-Marc Ogier, "Benchmarking of Document Image Analysis Tasks for Palm Leaf Manuscripts from Southeast Asia" J.Imaging, pp. 1- 27, 2018.
4. Papangkorn Inkeaw, Atcharin Klomsae, Sanparith Marukatat, "Lanna Dharma Handwritten Character Recognition on Palm Leaves Manuscript based on Wavelet Transform", ICSIPA, IEEE, pp. 253-258, 2015.
5. Nagendra Panini Challa, R.Vasanth Kumar Mehta, "Applications of Image Processing Techniques on Palm Leaf Manuscripts - A Survey", Helix, pp. 2013-2017, 2017.
6. Ge Peng, Pengfei Yu, Haiyan Li and Lesheng He, "Text Line Segmentation Using Viterbi Algorithm For The Palm Leaf Manuscripts of Dai", ICALIP, IEEE, pp. 336-340, 2016.
7. Ge Peng, PengFei Yu, HaiYan Li, HongSong Li, XuDong Zhu, "A Character Segmentation Algorithm for the Palm Leaf Manuscripts", ICCIA, IEEE, pp. 354-358, 2017.
8. Dona Valy, Michel Verleysen, and Kimheng Sok, "Line Segmentation Approach for Ancient Palm Leaf Manuscripts using Competitive Learning Algorithm", ICFHR, IEEE, pp. 108-113, 2016.
9. Jija Das Gupta, Bhabatosh Chanda, "A Model Based Text Line Segmentation Method for Off-line Handwritten Documents", ICFHR, IEEE, pp.125-129, 2010.
10. Vijaya Kumar Koppula, Atul Negi, "Fringe Map Based Text Line Segmentation of Printed Telugu Document Images", ICDAR, IEEE, pp. 1294-1298, 2011.
11. N. Tripathy and U. Pal, "Handwriting Segmentation of Unconstrained Oriya Text", Sadhana, Springer, pp.755-769, 2006.
12. Rapeeporn Chamchong, Chun Che Fung, "Text Line Extraction Using Adaptive Partial Projection for Palm Leaf Manuscripts from Thailand", ICFHR, IEEE, pp. 586-591, 2012.
13. Olarik Surinta, Michiel Holtkamp, Faik Karabaa, Jean-Paul van Oosten, Lambert Schomaker and Marco Wiering, "A* Path Planning for Line Segmentation of Handwritten Documents", ICFHR, IEEE, pp.175-180, 2014.
14. Alexander Berennolts, Michael Lindenbaum, "On The Performance of Connected Components Grouping", IEEE, pp.189-192, 2000.
15. R.Spurgen Ratheash, and M. Mohamed Sathik, “A Detailed Survey of Text Line Segmentation Methods in Handwritten Historical Documents and Palm Leaf Manuscripts”, IJCSE, pp 99-103 , 2019.

AUTHORS PROFILE

R. Spurgen Ratheash, an Assistant Professor of Information Technology, received his MCA degree in Computer Applications from Bishop Heber College, Bharathidasan University, Trichy, India in 2007. He received his MPhil degree in Computer Science and an M. Tech in Information Technology from Manonmaniam Sundaranar University, Tirunelveli, India in 2012 and 2014 respectively. He is a research scholar at Sadakathullah Appa College, affiliated to Manonmaniam Sundaranar University, Tirunelveli, India. His major research interests include Digital Image Processing, Document Image Analysis and Character Recognition of Tamil language palm leaf manuscripts.

Dr. M. Mohamed Sathik is the Principal of Sadakathullah Appa College, Tirunelveli, India. He received two Ph.Ds majored in Computer Science and Computer Science & Information Technology in Manonmaniam Sundaranar University, Tirunelveli, India. He has many more feathers in his cap by degrees such as M. Tech, MS (Psychology) and MBA. He is pursuing post Doctoral degree in Computer Science. Known for his active involvement in various academic activities, he has attended many national and international seminars, conferences and presented numerous research papers. With publications in many international journals, he has published two books besides having guided more than 40 research scholars. The prolific academician is a member of curriculum development committee of various universities and autonomous colleges in Tamil Nadu, India. His areas of specialization are Virtual Reality, Image Processing and Sensor Networks.

Optimal Character Segmentation for Touching Characters in Tamil Language Palm Leaf Manuscripts using Horver Method

M. Mohamed Sathik, R. Spurgen Ratheash

***Abstract:** An optimality of an automatic character recognition for Tamil palm leaf manuscripts can be achieved only by an efficient segmentation of touching characters. In this article, the touching characters are segmented as a single character to achieve an optimum solution by the recognizer in Optical Character Recognition (OCR). The proposed method provides a novelty in touching character segmentation of Tamil palm leaf manuscripts. An initial process of separation of background image and foreground characters is applied on the palm leaf images by filtering and removing unwanted pieces of characters by noise removal methods. The thickening process overcomes the difficulty of small breakages in the characters. The aspect ratio of the character image can be used to categorize the character such as single or multi touching. Single touching is divided by yet another ways such as horizontal or vertical touching. Finally, the proposed algorithm for* **Horizontal and Vertical character segmentation named as HorVer method** *is applied on the horizontally and vertically touching characters to segment as independent character. Experimental result produces 91% of an accuracy on segmenting the touching characters in Tamil palm leaf manuscript images collected from various resources and Tamil Heritage Foundation (THF). A novelty method can be achieved in Tamil touching character segmentation by the proposed algorithm.*

***Key words :** Character segmentation, Pre processing, Touching characters, Tamil character segmentation*

I. INTRODUCTION

Tamil language accommodates one of the oldest scripts of the world from 6th century BC. And the birth place of those characters is believed to be Keezhadi, Tamil Nadu, India. The shapes of Tamil characters have evolved from the Greek characters. The script was named as “Tamili” earlier and then called as “Tamil”. The writing style of the script is from left to the right. The script has been categorized into two such as vowels (Uyirezhuthu) and consonants (Meiyezhuthu). The vowel has 12 characters and the consonant has 18 characters. The combinations of those two (UyirMeiyezhuthu) categories make 247 characters with one special character ‘∴’ [1].

Ever since the evolution of the language, the characters have specific strokes as its forms. During the period of 17th Century AD, the Christian preacher Constantine Joseph Beschi named who was later known as ‘Veeramamunivar’ reshaped the forms by means of different strokes and shapes to differentiate the short (Kuril) and long (Nedil) vowels. In 19th Century AD, the great reformer from Tamil Nadu E.V. Ramasamy (Periyar) made changes on the shapes of the characters which continue to date. Prior to the invention of paper, an older civilization of Tamil people used to document their

medicinal hints, information about architecture, astrological ideas, and literatures on palm leaves because it was easily available then in their land. Palm leaf is generally 3.5 cm width and 35cm length as shown in Fig. 1. While writing on palm leaves, the writer holds the palm leaves on the left hand and writes his right hand. The issue here is the way stylus is held makes the difference in the forms of the scripts. Unlike holding a pen with fingers and thumb, a stylus is held while writing within the palm and the four fingers. A stylus is a needle like pen with no ink made of iron rods used to write on palm leaf. The writers write on the palm leaf with minimum pressure without any punctuation so as not to damage the palm leaf. While writing with a stylus, normally the letters or words are written without taking the stylus off the palm leaf and giving room neither for space nor for punctuation. This makes all the difference in the way the scripts are shown as touching characters [2]. OCR only recognizes single characters overlooking the touching characters. Hence, character segmentation is an important phase to provide an optimal solution in recognition phase. The touching characters are divided into two types such as single and multiple touching. The single touching characters are categorized further into two i.e. horizontal and vertical touching as shown in Fig.2. When the characters touch together with the same line characters that is known as horizontal touching as Fig.2a and when the characters touch with the subsequent line characters, they are classified as vertical touching as shown Fig.2b. The characters that touch both ways are called as multiple touching characters as shown in Fig.2c.

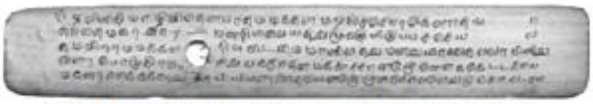


Fig.1 Image of Tamil palm leaf manucript

Revised Manuscript Received on March 30, 2020.

M. Mohamed Sathik, Principal and Research Supervisor
mmdsadiq@gmail.com PG and Research Department of Computer Science,
Sadakathullah Appa College, Tirunelveli, Tamilnadu, India. Manonmaniam Sundaranar University, Tamilnadu, India
R. Spurgen Ratheash, Research Scholar, Reg. No: 12334
spurgen@gmail.com Sadakathullah Appa College, Tirunelveli, Tamilnadu, India.

(a) (b) (c)

Fig. 2. Categories of Touching characters (a) Horizontal Touching (b) Vertical Touching (c) Multi Touching

II. LITERATURE REVIEW

The Tamil characters have loops, crossings, junctions and so on and so forth. These highly challenging character segmentation tasks are based on feature extraction using the global and local features. Gabor filter and co-occurrence matrix is used to identify the texture of the character and extracted the global feature. These character segmentation and feature extraction of writings are used by Support Vector Machine (SVM) [3], [4]. The Vertical projection is used to split the words and sub images into individual characters using Structure Based Character Segmentation (SBCS). The Character Threshold is an important part for the separation of the characters while taking an account of width of the character [5]. The characters are segmented horizontally by the template bank. The boundaries are located in edges that are identified by vertical segmentation [6]. The vertical projection profile is calculated for thin images and smoothed using Gaussian low pass filter with standard deviation. The segmentation path can be identified through the projection lines retained that controls an over segmentation of the characters [7]. The multi factorial analysis is used to identify cutting point for the touching characters by the measurement of dissimilarity and aspect ratio with the five fuzzy factors [8]. H. Fujisawa *et al*, introduces a pattern oriented segmentation method to achieve through boundary box to separate the overlapping strokes [9]. The image is known as decomposition by their general features such as dissection method using contextual knowledge by Richard G. Casey *et al* [10]. The word segmentation based on Dominant Overlap Criterion Segmentation performs to separate a stroke group. Besides it and provides the special attention to spatial and temporal features derived from the characteristics and detect the under segmented stroke groups. The SVM used for feedback the stroke group is based on the features to correct the wrong segmentation in the Attention Feedback Segmentation [11]. The touching and broken characters are processed to separate by Graph Partitioning-based Character Segmentation method and it is generalised for multi level writing style of Lamma Dhamma alphabet on palm leaf manuscripts [12]. The character segmentation is categorized by three ways such as ‘dissection’, ‘recognition-based’ and ‘holistic method’. The dissection methods have predefined rules to obtain segmentation points and it is specifically designed to use contour information [13].

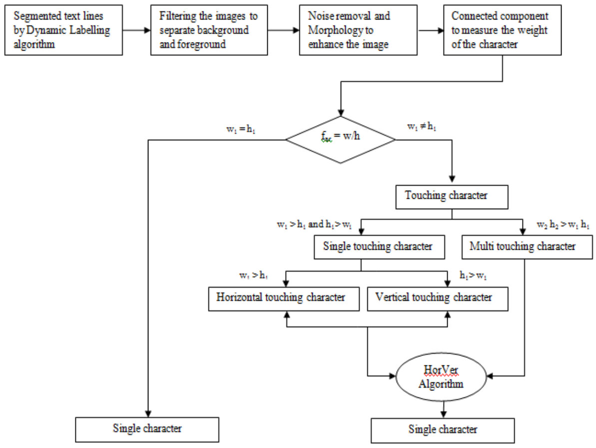


Fig.3 Architecture of HorVer algorithm

The remaining part of the paper has arranged as Section II provides the details of widespread character segmentation methods used in Tamil and various languages by means of a literature survey. The line segmentation method done by the researchers in section III and proposed method of touching character segmentation by HorVer method is explained in section IV. An evaluation results is dealt with in section V with the conclusion in section VI.

III. LINE SEGMENTATION

An OCR in palm leaf manuscripts starts with the primary phase of line segmentation process. An Adaptive Partial Projection (APP) method is used to identify the line numbers and space between the text lines by a piece wise projection method in Thai manuscripts. The second method is known as A* Path Planning (A*PP) is used to identify the touching and partially overlapping characters in text lines by heuristic way. The latter uses by means of various cost functions in Thai and Khmer language manuscripts. In Tamil manuscripts, both the above mentioned methods are ineffective and they change the structure of the character. The Dynamic Labelling Algorithm (DLA) is proposed by the researchers. This method resulted in segmentation of text lines with 96% of Recognition Accuracy (RA). This provides an optimality in Tamil palm leaf manuscripts line segmentation even the characters are deeply touching and overlapping with the proceeding or succeeding line characters [14].

IV. PROPOSED METHOD

The character segmentation is the second major phase in OCR of Tamil palm leaf manuscripts that proceeds after the successful line segmentation process by DLA. The pre-processing stage has image filtering, image sharpening, morphology that are used to remove unwanted particles other than the character present in the lines. Further, they make the character stroke clear. The flow diagram in Fig.3 provides the step by step process of proposed HorVer character segmentation method.

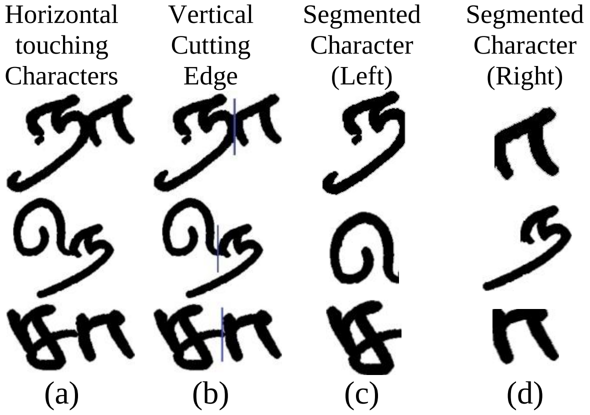


Fig. 4 Segmentation of horizontal touching character

Character Segmentation

In this article the character segmentation process has divided by the following two ways.

1. Identification of Touching Characters

The weight of the character is measured by connected component. A threshold value has fixed to evaluate the weight, when lower value gets than threshold can be considered as single characters and it is automatically separated without any complication. The greater value considered as touching characters using the following factor: *Factor 1: An aspect ratio(ar) of the touching character is larger than automatically separated single characters.*

The above shown factor (f_{ar}) is used to identify the touching characters and that need additional efforts to produce the single characters. The touching characters are defined by

$$\mathbf{f}_{ar} = e^{a}/(1 + e^{a}) \qquad (1)$$

where $a = w/h$, and w and h is the parameter of width and height of the character. After identification of the touching character, the latter is categorized as horizontal touching, vertical touching and multi touching characters as shown in Fig.2.

Table 1 Possible ways of Single touching horizontal characters

CATEGORY	TYPE	TOUCHIN G POINT	EXAMPLES
SINGLE TOUCHING (HORIZONTAL)	1		
	2		
	3		
	4		
	5		
	6		

The value of a=w_1>h_1 can be identified as a horizontal touching and a=h_1>w_1 fall under the category of vertical touching characters. The multi-touching characters have the observation of a=w_2h_2> w_1h_1. An analysis of Tamil palm leaf manuscripts gives an observation of maximum possibilities on touching point between two touching characters. That is consolidated in six ways of horizontal touching as in Table 1, four ways of vertical touching as in Table 2, and two ways of multi-touching as listed in Table 3.

Table 2 Possible ways of Single touching vertical characters

CATEGORY	TYPE	TOUCHING POINT	EXAMPLES
SINGLE TOUCHING (VERTICAL)	7		
	8		
	9		
	10		

Table 3 Possible ways of Multi Touching characters

CATEGORY	TYPE	TOUCHING POINT	EXAMPLES
MULTIPLE TOUCHING	11		
	12		

2. Segmentation of Single Touching Character

The segmentation of touching characters is possible when fixing the cutting edge between the characters. The cutting edge is defined as placing the segmentation point to separate two joined characters. Type 1 to 6 in Table 1 shows the possible ways of horizontal touching between two characters. The cutting edge is used for the above shown category characters by calculating their weight in columns. They are identified as the least weight that can predict as the touching point of the characters and also it is considered as cutting edge to segment the character. For this category of horizontal touching characters as shown in Fig. 4a, the cutting edge is vertical way because the weight is calculated by the column of the character as shown in Fig. 4b and the segmented characters are shown in Fig. 4c and Fig. 4d. The vertical touching character as discussed earlier in section 4.1 can be identified by height of the character. The type 7 to 10 in Table 2 shows the possible ways of touching point in vertical touching characters as shown in Fig. 5a. The cutting edge placed between the characters in least weight is calculated by rows. In vertical touching characters, the cutting edge is horizontal as shown in Fig. 5b and the segmented characters are shown in Fig. 5c and Fig. 5d. Type 11 and 12 in Table 3 shows the possibility of multiple touching characters that can be identified by horizontally touching with the same line characters and vertically touching with the subsequent line characters. The process of horizontal and vertical cutting edge provides the touching characters to segment single characters.

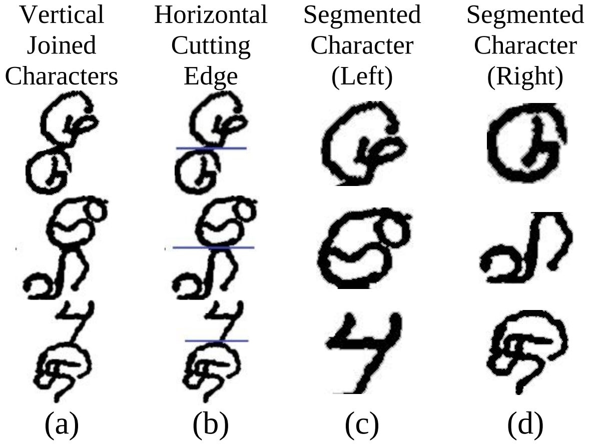


Fig. 5 Segmentation of vertical touching character

V. EXPERIMENTAL AND EVALUATION RESULTS

The HorVer algorithm is applied on line segmented Tamil palm leaf manuscript images by Dynamic Labelling, APP, A*PP algorithms. The proposed method identifies horizontal, vertical and multi touching characters successfully. Besides, it segments the characters as a single character. The experimental values are evaluated by the metrics of DR, RA, and FM based on MM, NN, and o2o. These evaluation criteria and tools are provided by ICDAR 2013 Handwritten Segmentation Contest.

An evaluation of this article by the metrics of one-to-one (oTo) match score, NN and MM [15]. oTo is computed for a region pair based on the evaluators acceptance threshold. Let NN be the total number of ground truth elements and MM be the total number of result elements. An above said three metrics are calculated with the oTo score.

The Detection Rate (DR) is defined by

$$DR = \frac{o2o}{NN} \quad (2)$$

Recognition Accuracy (RA) is

$$RA = \frac{o2o}{MM} \quad (3)$$

and F-measure (FM) is calculated by

$$FM = \frac{2.DR.RA}{DR+RA} \quad (4)$$

VI. CONCLUSION AND FUTURE WORK

The present article sheds light on the character segmentation of touching characters with the category of ‘horizontal touching’ and ‘vertical touching’ of Tamil palm leaf manuscript characters. In this article, the comparison of character segmentation method is applied on line segmented images by a novel approach of line segmentation algorithm named as Dynamic Labelling. The proposed DL is invented by the researchers. The latter have used the DL along with two other recent methods of line segmentation such as APP, A*PP algorithms. The HorVer character segmentation algorithm proposed in this article provides optimality about the ins and outs of character segmentation in the Tamil palm leaf manuscripts with 91% of Recognition Accuracy while APP and A*PP provide 87% and 89% respectively from the same value of NN. An overall performance of Recognition Accuracy in character segmentation on line segmented images by APP, A*PP, and Dynamic Labelling listed in Table 4. The Detection Rate of the image and F-Measure shows in Fig. 6 and Fig. 7. The image wise performance of DR and FM is in Fig. 8 and Fig. 9.When the segmented character structure is mismatched with the original structure of the Tamil character, it may solve the mismatch by fixing the least weight in various places on the characters and involve them for continuous iterations in Convolution Neural Network.

ACKNOWLEDGEMENTS

The authors whole heartedly thank Dr. V. Kattalai Kailasam, former Tamil Professor, Tirunelveli, Tamil Nadu, India for the generosity in permitting to access his palm leaf resources. The authors extend their gratitude to Dr. Subashini, Founder of Tamil Heritage Foundation who shared the images of various writing style Tamil Palm Leaf manuscripts

Table 4 Overall performance by HorVer method

ALGORITHMS	NN	MM	O2O	RA
APP	4254	4114	3506	86.15
A*PP	4254	4242	3726	88.78
HorVer	4254	4340	3884	90.63

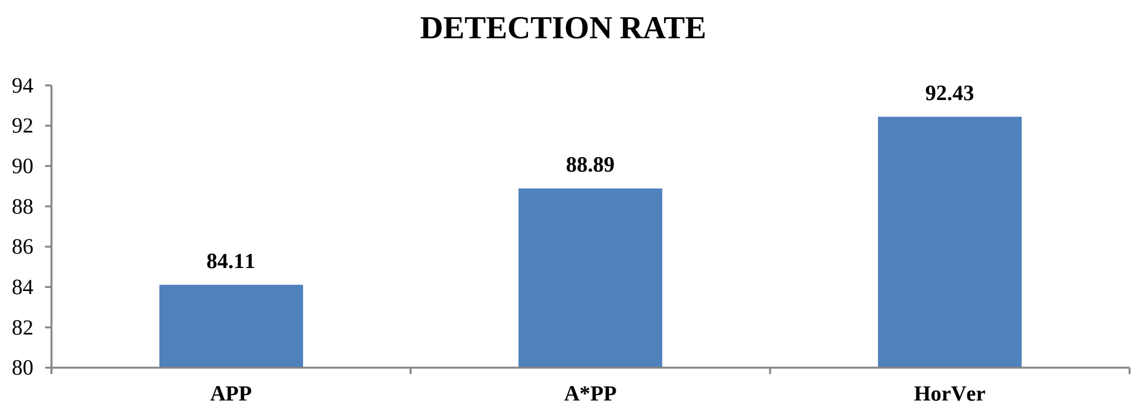


Fig. 6 Comparison of HorVer Detection Rate with APP and A*PP algorithms

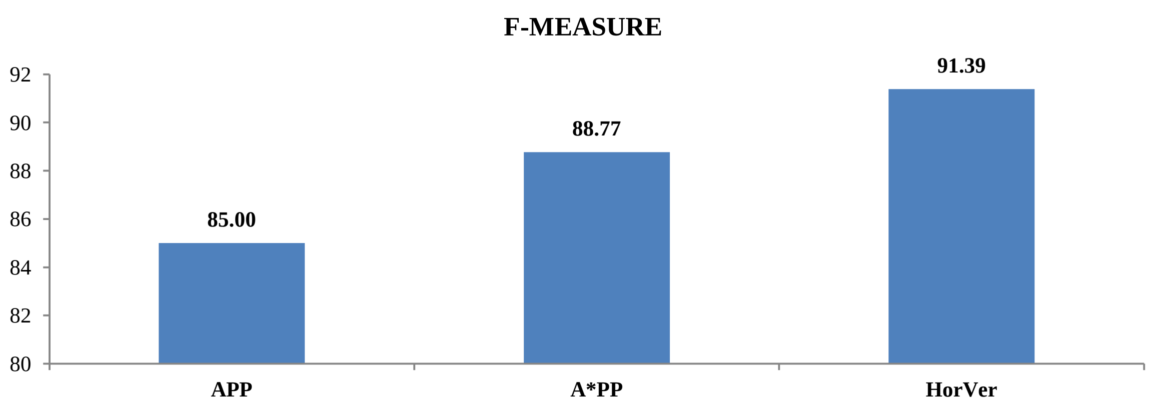


Fig. 7 Comparison of HorVer F-Measure with other two algorithms

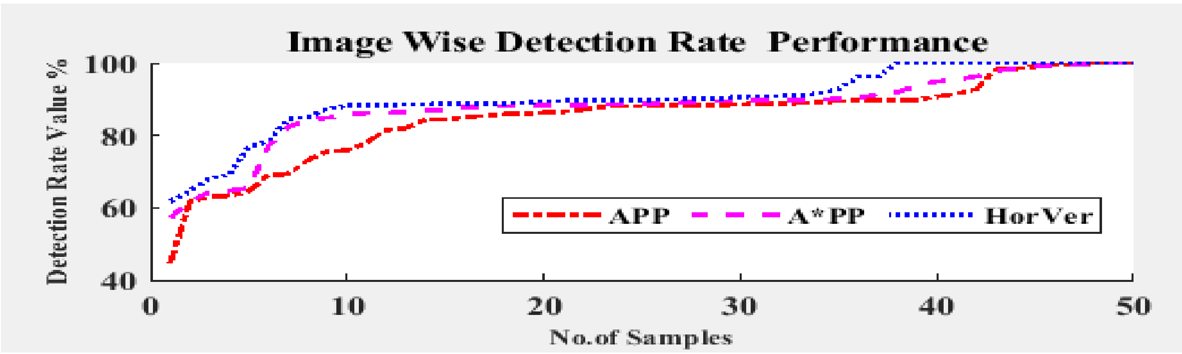


Fig. 8 DR image wise performance of HorVer and other two algorithms

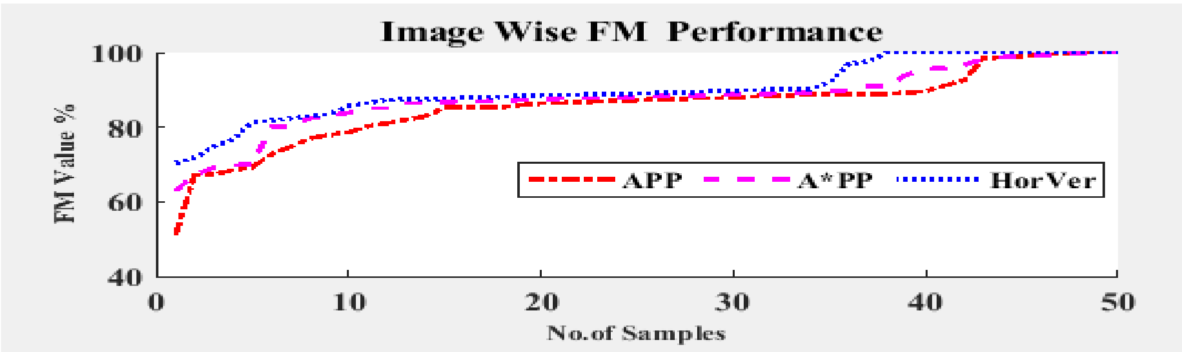


Fig. 9 FM image wise performance of HorVer and other two algorithms

REFERENCES

1. Ponniah.N, Ganesaiyar. C, "Tholkappiyam Eluthadhigara Moolamum" Sunnagam Thirumagal Azhuthagam, pp. 18-19, 1937.
2. D. Udaya Kumar, G.V.Sreekumar, U. A. Athvankar, "Traditional writing system in Southern India — Palm leaf manuscripts", Design Thoughts. July 2009.
3. Thendral T, Vijaya MS, Karpagavalli S, "Analysis of Tamil Character Writings and Identification of Writer Using Support Vector Machine",IEEE IEEE International Conference on Advanced Communications, Control and Computing Technologies, pp 1407 - 1411, 2014.
4. Manigandan T, Dr. V.Vidhya, Dr.Dhanalakshmi V, Nirmala B, "Tamil Character Recognition from Ancient Epigraphical Inscription using OCR and NLP",IEEE, International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp 1008 - 1011, 2017.
5. Kathirvalavakumar Thangairulappan, Karthigaiselvi Mohan, "Efficient Segmentation of Printed Tamil Script into Characters Using Projection and Structure", IEEE, Fourth International Conference on Image Information Processing (ICIIP), pp 484 - 489, 2017.
6. J. Tian, R. Wang, G. Wang, J. Liu, and Y. Xia, "A two-stage character segmentation method for Chinese license plate," Computers and Electrical Engineering, pp 539-553, 2015.
7. Parul Sahare and Sanjay B. Dhok, "Multilingual Character Segmentation and Recognition Schemes for Indian Document Images", IEEE, pp 1-9, 2018.
8. U. Garain, and B.B. Chaudhuri, "Segmentation of touching characters in printed Devnagari and Bangla scripts using fuzzy multifactorial analysis", IEEE Proceedings of Sixth International Conference on Document Analysis and Recognition, pp. 805-809, 2002.
9. H. Fujisawa, Y. Nakano, and K. Kurino, "Segmentation methods for character recognition: From segmentation to document structure analysis", Proceedings of the IEEE, pp 1079–1092, 1992.
10. R. G. Casey and E. Lecolinet, "A survey of methods and strategies in character segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp 690 - 706, 1996.
11. Suresh Sundaram, A.G.Ramakrishnan, "Attention Feedback based robust segmentation of online handwritten isolated Tamil words", ACM Trans. Asian Language Information Processing, pp. 1-25, 2013.
12. Papangkorn Inkeaw, Jakramate Bootkrajang, and Phasit Charoenkwan, "Recognition based character segmentation for multi-level writing style", Springer, pp. 21-39, Jun 2018.
13. Shi.Z, Govindaraju.V, "Segmentation and recognition 0f connected handwritten numeral strings", Pattern Recognition, 30(9), pp 1501-1504, Sep.1997.
14. Spurgen Ratheash.R, Mohamed Sathik. M, "Line Segmentation Challenges in Tamil Language Palm Leaf Manuscripts, IJITEE, pp.2363-2367, 2019.
15. R.Spurgen Ratheash, and M. Mohamed Sathik, "A Detailed Survey of Text Line Segmentation Methods in Handwritten Historical Documents and Palm Leaf Manuscripts", IJCSE, pp 99-103, 2019.

AUTHORS PROFILE

Dr. M. Mohamed Sathik is the Principal of Sadakathullah Appa College, Tirunelveli, India. He received two Ph.Ds majored in Computer Science and Computer Science & Information Technology in Manonmaniam Sundaranar University, Tirunelveli, India. He has many more feathers in his cap by degrees such as M. Tech, MS (Psychology) and MBA. He is pursuing post Doctoral degree in Computer Science. Known for his active involvement in various academic activities, he has attended many national and international seminars, conferences and presented numerous research papers. With publications in many international journals, he has published two books besides having guided more than 40 research scholars. The prolific academician is a member of curriculum development committee of various universities and autonomous colleges in Tamil Nadu, India. His areas of specialization are Virtual Reality, Image Processing and Sensor Networks.

R. Spurgen Ratheash, an Assistant Professor of Information Technology. He is a research scholar at Sadakathullah Appa College, affiliated to Manonmaniam Sundaranar University, Tirunelveli, India. His major research interests include Digital Image Processing, Document Image Analysis and Character Recognition of Tamil language palm leaf manuscripts. He received his MCA degree in Computer Applications from Bishop Heber College, Bharathidasan University, Trichy, India in 2007. He received his MPhil degree in Computer Science and an M.Tech in Information Technology from Manonmaniam Sundaranar University, Tirunelveli, India in 2012 and 2014 respectively.

Handwritten Tamil Character Recognition in Palm Leaf Manuscripts using BiLSTM Classifier

Dr. M. Mohamed Sathik [1], R. Spurgen Ratheash [2]

[1]Principal and Head, Research Department of Computer Science,
Sadakathullah Appa College, Tirunelveli, Tamil Nadu, India.
Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012.
TamilNadu, India.
[2]Research Scholar, Reg.No.12334, Sadakathullah Appa College, Tirunelveli, Tamil Nadu, India
Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012.
TamilNadu, India.

E-mail :[1]mmdsadiq@gmail.com, [2] spurgen@gmail.com

Abstract

Recognizing and reading the Tamil characters which were written many centuries ago in the palm leaf manuscripts is a tough task. As the character has taken a different form over the centuries, the shape of the characters has not been known to contemporary Tamil readers. Neural network classifiers are the tool to unlock the treasure of knowledge paralyzed for such reasons. Only those who are familiar with the shape and strokes of Tamil characters can read palm leaf manuscripts. The knowledge gained from the Tamil literatures have learned about the palm leaf manuscripts has been prepared through computer so that all the people can know the written ideas. This method works in three stages like text line segmentation, character segmentation, character recognition. The final stage of recognition has done by the Bidirectional Long Short Term Memory (BiLSTM) classifier that produces a better result than other conventional CNN methods in Tamil character recognition.

Keywords: *RNN, BiLSTM, LSTM, CNN, palm leaf manuscripts, Tamil character recognition, softmax layer, fully connected layer*

1. Introduction

Initially, ancient people have written the Tamil scripts on the pots and then started writing on palm leaves. The palm leaf manuscripts existed as a medium to write and record incidents, events, innovations in medicine, astrology and literatures. The collection of palm leaf manuscripts is known as *suvadigal* which contains approximately 40 to 50 preserved leaves with 15 to 30 cm length and 3 to 12 cm width written on both the sides. Each side of the palm leaf has 5 to 6 written lines with sequence of characters from left to right. When the leaves were in dilapidated condition, the scribes copied the contents of one leaf to another new leaf [1]. While copying, an inherited knowledge about shapes and writing style of locality influences to define the shapes of the character. Tamil language has 247 independent characters with the combination of 12 vowels and 18 consonants and one special character ' ⁄' as in *Figure 1*. Most of the recent Tamil characters were not written in palm leaf manuscripts and the written characters are also different in forms. As a result, very much vital information in siddha medicine, astrology, and literatures could not be recognized [2] because of the shapes of old characters have been forgotten by the new shapes. Even though digitization is the key solution to preserving and storing the palm leaf manuscripts, there has been no progress in publishing them as books as the characters are unrecognizable. This article discusses the highly productive way of Tamil character recognition by matching the early day character shapes with present-day

			VOWELS											
			அ	ஆ	இ	ஈ	உ	ஊ	ஐ	எ	ஏ	ஒ	ஓ	ஔ
CONSONANTS	VALLINAM	க்	க	கா	கி	கீ	கு	கூ	கை	கெ	கே	கொ	கோ	கௌ
		ச்	ச	சா	சி	சீ	சு	சூ	சை	செ	சே	சொ	சோ	சௌ
		ட்	ட	டா	டி	டீ	டு	டூ	டை	டெ	டே	டொ	டோ	டௌ
		த்	த	தா	தி	தீ	து	தூ	தை	தெ	தே	தொ	தோ	தௌ
		ப்	ப	பா	பி	பீ	பு	பூ	பை	பெ	பே	பொ	போ	பௌ
		ற்	ற	றா	றி	றீ	று	றூ	றை	றெ	றே	றொ	றோ	றௌ
	MELLINAM	ங்	ங	ஙா	ஙி	ஙீ	ஙு	ஙூ	ஙை	ஙெ	ஙே	ஙொ	ஙோ	ஙௌ
		ஞ்	ஞ	ஞா	ஞி	ஞீ	ஞு	ஞூ	ஞை	ஞெ	ஞே	ஞொ	ஞோ	ஞௌ
		ண்	ண	ணா	ணி	ணீ	ணு	ணூ	ணை	ணெ	ணே	ணொ	ணோ	ணௌ
		ந்	ந	நா	நி	நீ	நு	நூ	நை	நெ	நே	நொ	நோ	நௌ
		ம்	ம	மா	மி	மீ	மு	மூ	மை	மெ	மே	மொ	மோ	மௌ
		ன்	ன	னா	னி	னீ	னு	னூ	னை	னெ	னே	னொ	னோ	னௌ
	IDAYINAM	ய்	ய	யா	யி	யீ	யு	யூ	யை	யெ	யே	யொ	யோ	யௌ
		ர்	ர	ரா	ரி	ரீ	ரு	ரூ	ரை	ரெ	ரே	ரொ	ரோ	ரௌ
		ல்	ல	லா	லி	லீ	லு	லூ	லை	லெ	லே	லொ	லோ	லௌ
		வ்	வ	வா	வி	வீ	வு	வூ	வை	வெ	வே	வொ	வோ	வௌ
		ழ்	ழ	ழா	ழி	ழீ	ழு	ழூ	ழை	ழெ	ழே	ழொ	ழோ	ழௌ
		ள்	ள	ளா	ளி	ளீ	ளு	ளூ	ளை	ளெ	ளே	ளொ	ளோ	ளௌ

Figure 1. Tamil Characters

characters. The result produces better performance on Tamil character recognition in palm leaf manuscripts than other conventional methods.

The remaining part of the article has been arranged as follows. Section 2 discusses the related work in Tamil character recognition. Section 3 explores the general BiLSTM layer framework. Section 4 discusses the process of Tamil character recognition. Section 5 describes dataset and training as an experimental setup. Section 6 presents the result and discussions. The last section gives the conclusion of the paper.

2. Related work

The character recognition process offered in many languages such as Thai, Khmer, Arabic, English, and Tamil are in different techniques. The Tamil character recognition has been taken into consideration for the analysis. The Kohonen Self-Organizing Map (SOM) tuned by global feature technique in the type of Artificial Neural Network used to classify handwritten Tamil characters [3]. The symbols, numerals and Tamil characters are recognized by the techniques of Gabor Filter and Support Vector Machines (SVM) [4]. Hilditch’s Algorithm is used in Neural Network to recognize typed Tamil characters by passing the Horizontal histogram, Vertical histogram, radial, input and output features in minimum number of classes [5]. The features of character height, width, number of vertical and horizontal lines, curves, circles, slope lines, dots are extracted and processed by SVM and Kohenen SOM in Artificial Neural Network to recognize offline handwritten Tamil characters for eight classes[6]. A method to extract the feature of the characters is Hu’s invariant and Zernike movements and classifies the characters using Feed Forward Neural Network [7]. A survey in Tamil character recognition explained deep belief network method to extract the features, Restricted Boltzmann Machines model to train the character using deep learning in large data [8]. An ideal edge identification method used in palm leaf Tamil characters using Canny Edge Detection by three ways such as great discovery, great confinement, and negligible reaction in Artificial Neural Network. Finally, their method enhances the character with the binarization technique to provide a remarkable result in recognition [9]. A survey provides the method about character recognition in palm leaf manuscripts of Southeast Asia languages like Balinese, Khmer, Sundanese using CNN [10]. CNN is used to recognize the character of Tamil palm leaf manuscripts by five layers such as convolution, pooling, activation, fully connected layers and softmax classifiers [11]. Nine layers including five convolution layers and each two of max pooling and fully connected layers are used to recognize handwritten Tamil characters. The ReLU activation function is used in each convolution layers [12].

3. Background

The digitized palm leaf manuscript image is passing through the process of pre-process, text line segmentation, character segmentation and finally reaches the character recognition as in *Figure 2*. The digitized image contains background in yellow colour and foreground text in black. The stains of leaf become as unwanted picture information known as noise. The noise removal is used to remove the objects other than characters. Binarization converts colour images into binary and separate the text foreground from dominating leaf colour background. The binary image has two major processes such as line segmentation and character segmentation prior to character recognition. In the first process, text line segmentation using Dynamic Labelling Algorithm (DLA) segments the text lines from the binary images of palm leaf manuscript [13]. In the second process, character segmentation separates the sequence of characters into individual character using HorVer method from the segmented text line images [14]. The two DLA and HorVer methods developed by the researchers have provided novelty in text line and character segmentation respectively. Both the methods are giving major participation to improve the result in the process of character recognition using Recurrent Neural Networks (RNN).

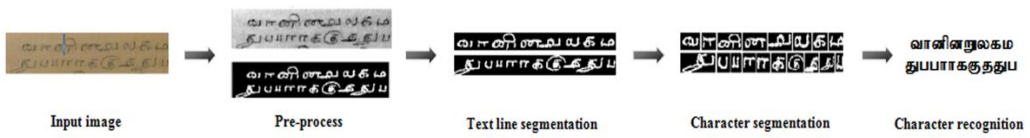


Figure 2. Stages of Tamil Character Recognition

3.1 Long Short Tem Memory (LSTM) Block

LSTM is a classifier in Recurrent Neural Networks (RNN). LSTMs are used to process the sequential data for classification such as character recognition, sentiment analysis, speech modelling, and language modelling. LSTM overcomes vanishing gradient and exploding gradient problem by memory which has the remembrance of past data. The gates are used to retain the relevant information in memory or fail to remember unrelated information. The three gates such as forget gate, input gate, and output gate are used respectively in the architecture of LSTM block as in Figure 3. The first gate is used to decide whether to discard or keep the information identified by 0 and 1respectively. The input (x_t) and previous hidden state (h_{t-1}) values are combined and passed to sigmoid function that decides which should retain or discard from the gate. In the second gate, two

ISSN: 2233-7857 IJFGCN

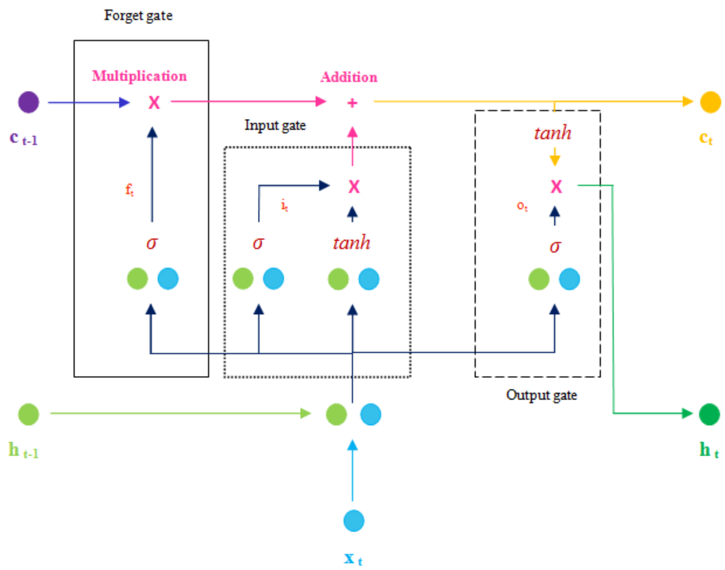


Figure 3. LSTM block

functions are applied on input such as sigmoid and tanh. The first decides which input value gets 0 or 1. The tanh adds weight to the value from input ranges from -1 to 1. The results of both functions are multiplied and update the previous cell state which has multiplied by the forget gate. The third gate is output gates that apply sigmoid function on the combined value of hidden state and input values to know which gets 0 or 1. The tanh function applied on the cell state updated by previous gates. The multiplied results of both produce an input to the next LSTM block [15].

3.2 Recurrent Neural Network (RNN)

In RNN, the image vectors are taken as input for sequence input layer to extract the features and passed to the Bidirectional Long Short Term Memory (BiLSTM) layer. The sum and multiplying options are applied on features in forward and backward LSTM cells. The output of LSTM blocks passed to multiply the weight matrix and to add the bias vector by fully connected layer [16]. The activation function is added to the weighted features in softmax layer followed by classification layer used to compute loss and finally the output produced.

3.2.1 Sequence Input Layer: Sequence Input Layer is the first layer in PLTCR architecture. The two dimensional vector sequence input image considers input size as a scalar for the count of features. The vector has three elements such as height (h) and width (w), and number of channels (c) of an image.

3.2.2 BiLSTM Layer: BiLSTM is a wrap that conjoins two parallel LSTM layer. One of the two with input processed forward and other one with output processed backwards. Merger mode of the bidirectional layer combined the forward and backward outputs and passed to the subsequent layer. The merge being with the options of sum, multiplication to add and multiply the output together, concatenation and average are used to produce output for the next layer, the default option is concatenation. The LSTM is also known as memory blocks when they are connected recurrently. The memory blocks have three multiplicative units such as input gate (i_t), output gate (o_t) and forget gate (f_t). The memory cells update by hidden layer content (h_{t-1}), input (x) with the current time step (t) and add bias (b) value [17]. The sigmoid (σ) function makes the decision to retain the values in each gates as in the following mathematical representations.

$$i_t = \sigma(C_i[h_{t-1}, X_t] + b_i) \tag{1}$$

$f_t = \sigma(C_f[h_{t-1}, X_t] + b_f)$ (2)

$o_t = \sigma(C_o[h_{t-1}, X_t] + b_o)$ (3)

The tanh function supports to distribute the gradient longer by vector cell state (c_t) to memory cell to evade the vanishing or exploding gradient problem. The tanh function adds weight to the input with bias value and updates the previous cell state as in relationship 4.

$c_t = tanh(C_c[h_{t-1}, X_t] + b_c)$ (4)

The benefit of sequence modelling to access both the past and future contents in BiLSTM can be achieved by forward and backward LSTM layers (Alex Graves et al., 2005). In character recognition, the output of two LSTM blocks for the character (h_i) is sum of features in forward and backward block output as in the following representation.

$h_i = [\overrightarrow{h_t} \oplus \overleftarrow{h_t}]$ (5)

3.2.3 Fully Connected Layer: The weight and bias add to all neurons in this layer produced by the previous BiLSTM Layer. The patterns can be identified by combining all the features that learned from the succeeding layer. The process is used to classify the image. The layer is independent at each time step when the sequence input [18]. The bias adds with the weight of an input (*x*) with current time stamp (t).

3.2.4 Softmax Layer: The softmax is activation function to compute the probability distribution for the list of classes (x_i) in the range between 0 and 1 with the sum of probability is equivalent to 1. The softmax can be calculated by the following representation [19].

$S(x_i) = \frac{exp^{x_i}}{\sum_j exp^{x_j}}$ (6)

The layer is unlike sigmoid rather performs multi class classification task. In loads of architecture, the layer more or less exists at the end so it also known as output layer of deep learning architecture.

3.2.5 Classification Layer: The layer calculates the cross entropy by assign mutually exclusive classes to each input values taken from the softmax function by the following mathematical relation. The number of samples (*s*), number of classes (*c*), an output (*o*), and an indicator (*t*) of i^{th} sample fit in with j^{th} class as in the following representation [20].

$C = -\frac{1}{n}\sum_{i=1}^{s}\sum_{j=1}^{c} t_{ij} \ln o_{ij}$ (7)

4. Proposed method

The Tamil character recognition in palm leaf manuscript using RNN has two phases and each has two processes such as normalization and labelling for first phase, training and testing for second phase as in *Figure 4*. In the first phase, the character segmentation by HorVer method produces different size of images. Each image must be normalized in equal aspect ratio 30 x 30 by the centre point of the strokes. The measurement of above and below pixels form the centre decides to acquire the complete shape of the character. In the second phase, the normalized images are labelled individually to train the character. The background of an input image is black in colour and foreground character is in white colour with the value of 0 and 1 respectively.

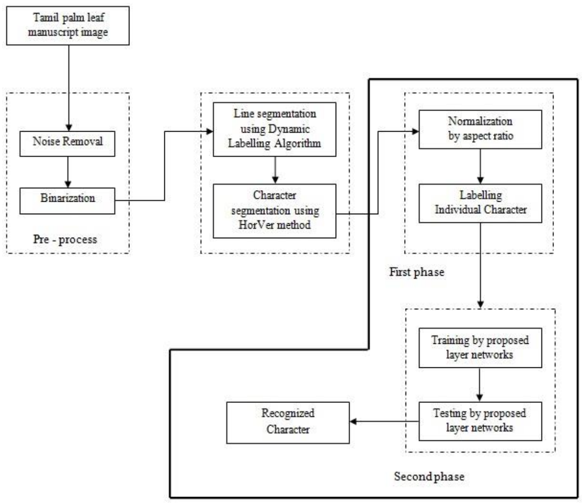


Figure 4. Process of Tamil Character Recognition

4.1 Architecture

The Tamil character recognition in palm leaf manuscripts have the hierarchical order of sequence input layer, BiLSTM layer, Fully connected layer, Softmax layer and Classification layer with suitable functionalities to implement RNN as in *Figure 5*.

4.2 BiLSTM Layer in Tamil Character Recognition

The RNN model has created as specified in the previous section for Tamil character recognition in palm leaf manuscripts. BiLSTM layer update the features with 'sigmoid' gate activation function, 'tanh' state activation function, and 200 x 1 hidden states. 100 epochs of training cycle and maximum 97 iterations per epoch with the learning rate of 0.001. The third layer, fully connected layer update the weights by total number of characters x 200. The loss calculated by crossentropyex loss function in softmax layer as fourth layer end with the allocation of labels for classes in fifth layer as classification layer. The second phase of testing, a new data set contains 2,640 Tamil characters in palm leaf manuscripts used to recognize 88 characters which was not trained. In training the layer used several hyper parameters to recognize Tamil characters in palm leaf manuscripts as in *Table 1*.

Table 1. Hyper Parameters for Tamil Character Recognition

Hyper Parameters	Values
Initialization	Glorot
Batch Size	27
Interpreter	Adam
Epochs	100
Learning rate	0.001

5. Experimental setup

5.1 Dataset

The palm leaf Tamil character recognition has a unique dataset created by the researchers. The different styles of vowel and consonant character images are in white strokes in black background. The dataset has IWFHR2010Tamil vowel characters collected from HP Labs India available at free of cost. The consonant character images are handwritten characters collected from 270 members and digitalized by the scanner.

Table 2. Dataset of Palm Leaf Tamil Characters

Contents	Counts
Total no of palm leaves	950
No of text lines per leaf	5
Total no of lines	4750
No of characters in each line	50 or 45
Total no characters	213750

The scanned image has been segmented by the researchers and classified as single. The detail of the dataset has shown in *Table 2*. In the collection of data, 10457 individual character mages have been selected for training and 2640 for testing in neural network.

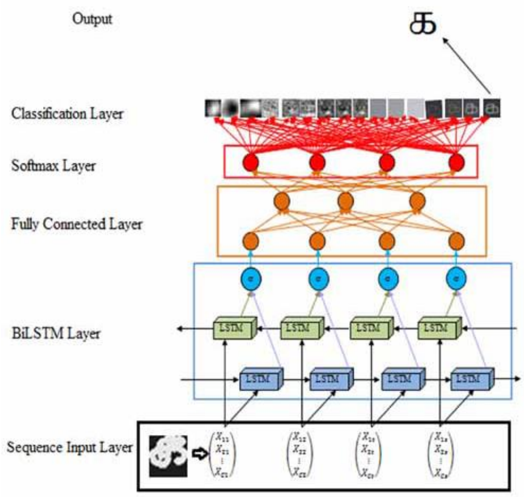


Figure 5. Layer Architecture for Tamil Character Recognition

5.2 Training and Testing

Different batch size were trailed and fixed the batch size as 27. After the layer architecture is compiled, the specific processed character dataset is loaded to train the model. The training is done by 100 epochs as in *Figure 6*. The tuning is made in hidden states and activation function to achieve an optimum in accuracy.

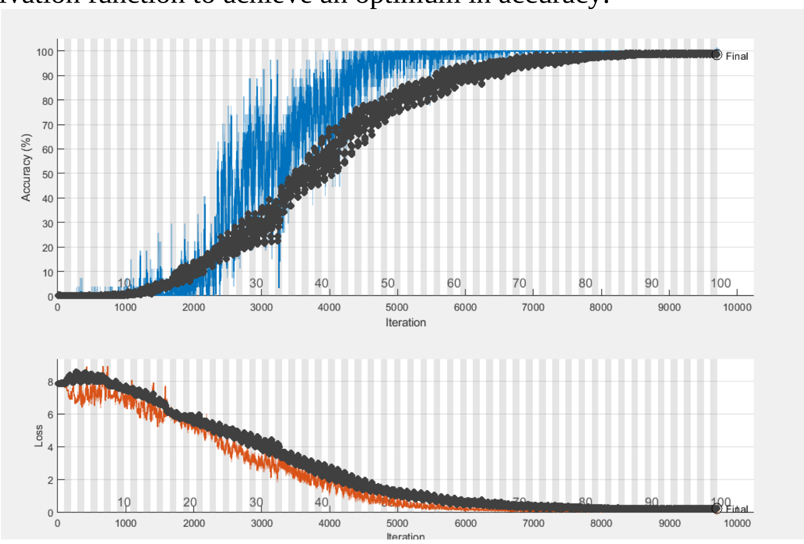


Figure 6. BiLSTM Training and Validation

6. Results and discussions

The RNN layer architecture recognizes the character written by different scribers. This work provides 1 % of wrong prediction rate in testing with the Tamil character in palm leaf manuscripts. The dataset is trained separately for CNN and LSTM classifiers and compared the Recognition Accuracy with proposed layer architecture using BiLSTM classifier as in *Figure 7*. The BiLSTM classifier has the benefit to defeat over-fitting

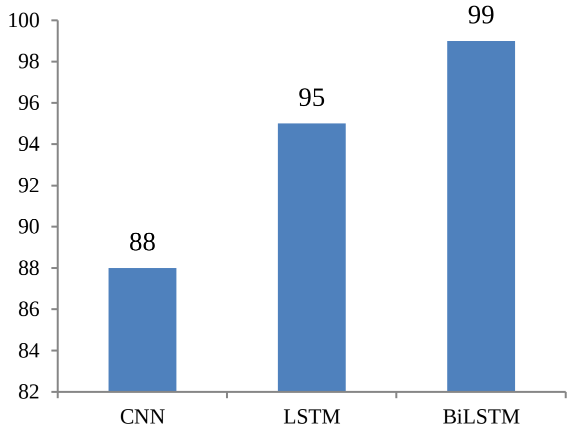


Figure 7. Performance of Palm Leaf Tamil Character Recognition

problem without using dropout techniques in CNN. The LSTM does not have backward propagation that reduces the performance in palm leaf Tamil character recognition. The recognition accuracy of proposed work is compared with other previous works as in *Table 3.*

Table 3. Comparison of Tamil Character Recognition

Years	Dataset	Language	Document	No. of classes	Method	Training Accuracy	Recognition Accuracy
2016	HP Labs	Tamil	Handwritten	35	CNN	99%	94.40%
2018	HP Labs	Tamil	Handwritten	146	CNN	-	88.86%
2019	Own	Tamil	Palm Leaf	60	CNN	-	96.21%
2019	HP Labs	Tamil	Handwritten	156	CNN	95.16%	97.70%
Proposed	Palm leaf	Tamil	Palm Leaf	88	BiLSTM -RNN	92.31%	99.57%

The CNN classifier combined with principal component analysis and trained 50 epochs to get maximum accuracy [21]. The over-fitting problem between training and validation and different kind of regularization methods were applied to solve that produced 89.3% of test accuracy initially in CNN. Stochastic pooling, probabilistic weighting and dropout techniques were used to get marginal changes in result [22]. The single scriber character set 3.79% of erroneous prediction and 0.64sec time has taken to predict one character [11]. The dropout regularization technique were used in every convolution layer with an initial probability 0.1 and increased by the same to overcome the over-fitting [12].

ISSN: 2233-7857 IJFGCN

Figure 8. Two Characters in Single Image

7. Conclusion

The research work used RNN classifier to recognize Tamil characters in palm leaf manuscripts. The work has recognized all characters used in Tamil palm leaf manuscripts long before 300 to 400 years. The RNN classifier also predicts exact character when two characters have joined together in a single image *Figure 8*. The recognition accuracy in this work provides much better result than the previous conventional handwritten Tamil character recognition methods. This work is benchmarking for Tamil character recognition in palm leaf manuscripts that will extend to recognize the stone epigraphs in future.

References

[1] D. Udaya Kumar, G.V Sreekumar and U.A Athvankar, "Traditional Writing System in Southern India Palm Leaf Manuscripts", Design thoughts, IDCIITB, (2009), pp. 1–7.

[2] S. Rajkumar, N. Srinivasan, T. Thirunarayanan and R. Sangeetha, "Survey of Tamil Siddha manuscripts in possession of Traditional Healers in Northern Tamil Nadu", International Journal of Pharmacology and Clinical Sciences, vol. 1, no. 3, (2012), pp. 68-73.

[3] K. Sarveswaran and D. Ratnaweera, "An Adaptive Technique for Handwritten Tamil Character Recognition", Proceedings of International Conference on Intelligent and Advanced Systems, Kuala Lumpur, Malaysia, (2007) November 25 – 28.

[4] P. Soman, R. Ramanathan, S. Ponmathavan, N. Valliappan, L. Thaneshwaran and S. Arun.S.Nair, "Optical Character Recognition for English and Tamil Using Support Vector Machines", Proceedings of International Conference on Advances in Computing, Control, and TelecommunicationTechnologies, Trivandrum, India, (2009) December 28 -29.

[5] V. Karthikeyan, "Hilditch's Algorithm Based Tamil Character Recognition", International Journal of Computer Science & Engineering Technology, vol. 4, no. 3, (2013), pp. 268 - 273.

[6] P. Banumathi and M. Nasira, "Handwritten Tamil Character Recognition using Artificial Neural Networks", Proceedings of International Conference on Process Automation, Control and Computing, Tamilnadu, India, (2011) November 20 - 22.

[7] Amitabh Wahi, S. Sundaramurthy and P. Poovizhi, "Handwritten Tamil Character Recognition" Proceedings of Fifth International Conference on Advanced Computing, Chennai, India, (2013) December 18 - 20.

[8] R. Jagadeesh Kannan and S. Subramanian, "An Adaptive Approach of Tamil Character Recognition Using Deep Learning with Big Data-A Survey", Proceedings of the 49th Annual Convention of the Computer Society of India, Telangana, India, (2014) December 12 - 14.

[9] P. Selvakumar and S. Hari Ganesh, "Tamil Character Recognition using Canny Edge Detection Algorithm", Proceedings of World Congress on Computing and Communication Technologies, Tamil Nadu, India, (2017) February 2 – 4.

[10] MadeWinduAntaraKesiman, Dona Valy, Jean-Christophe Burie, Erick Paulus, Mira Suryani, SetiawanHadi, Michel Verleysen, SopheaChhun and Jean-Marc Ogier, "Benchmarking of Document Image Analysis Tasks for Palm Leaf Manuscripts from Southeast Asia", J. Imaging, vol. 4, no. 2, (2018), pp 1- 27.

[11] R. S. Sabeenian, M. E. Paramasivam and P. M. Dinesh, "Palm-Leaf Manuscript Character Recognition and Classification Using Convolutional Neural Networks", Computing and Network Sustainability, Springer, (2019), pp. 397-404.

[12] B. R. Kavitha and C. Srimathi, “Benchmarking on offline Handwritten Tamil Character Recognition using Convolutional Neural Networks”, Journal of King Saud University – Computer and Information Sciences, (2019), pp. 1–8.

[13] R. Spurgen Ratheash and M. Mohamed Sathik, "Line Segmentation Challenges in Tamil Language Palm Leaf Manuscripts", International Journal of Innovative Technology and Exploring Engineering, Vol. 9, No.1, (2019), pp 2363 - 2367.

[14] M. Mohamed Sathik and R. Spurgen Ratheash, "Optimal Character Segmentation for Touching Characters in Tamil Language Palm Leaf Manuscripts using Horver Method", International Journal of Innovative Technology and Exploring Engineering, Vol.9, no. 6, (2020), pp 1010 - 1015.

[15] SeppHochreiter and JurgenSchmidhuber, “Long Short-Term Memory”, Neural Computation, Vol. 9, No. 8, (1997), pp 1735 – 1780.

[16] Alex Graves, Santiago Fernandez and Jurgen Schmidhuber, "Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition", Proceedings of International Conference on Artificial Neural Networks, Hamburg, Germany, (2014) September 15-19.

[17] Rania Maalej and Monji Kherallah, "Convolutional Neural Network and BLSTM for offline Arabic handwriting recognition", Proceedings of International Arab Conference on Information Technology, Lebanon, (2018) November 28-30.

[18] J. Wu, "Compression of fully-connected layer in neural network by Kronecker product," Proceedings of Eighth International Conference on Advanced Computational Intelligence, Chiang Mai, (2016) February 14 - 16.

[19] R. Hu, B. Tian, S. Yin and S. Wei, "Efficient Hardware Architecture of Softmax Layer in Deep Neural Network," Proceedings of International Conference on Digital Signal Processing, Shanghai, China, (2018) November 19 - 21.

[20] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network,"Proceedings of International Conference on Engineering and Technology, Antalya, Turkey,(2017) August 21 - 23.

[21] Prashanth Vijayaraghavan and Misha Sra, "Handwritten Tamil Recognition using a Convolutional Neural Network", Massachusetts Institute of Technology, (2016), pp 1- 7.

[22] M. Sornam and C. Vishnu Priya, "Deep Convolutional Neural Network for Handwritten Tamil Character Recognition Using Principal Component Analysis", Springer Nature, (2018), pp. 778–787.

Authors

Dr. M. Mohamed Sathik is the Principal of Sadakathullah Appa College, Tirunelveli, India. He received two Ph.Ds majored in Computer Science and Computer Science & Information Technology in Manonmaniam Sundaranar University, Tirunelveli, India. He has many more feathers in his cap by degrees such as M. Tech, MS (Psychology) and MBA. He is pursuing post Doctoral degree in Computer Science. Known for his active involvement in various academic activities, he has attended many national and international seminars, conferences and presented numerous research papers. With publications in many international journals, he has published two books besides having guided more than 40 research scholars. The prolific academician is a member of curriculum development committee of various universities and autonomous colleges in Tamil Nadu, India. His areas of specialization are Virtual Reality, Image Processing and Sensor Networks.

R. Spurgen Ratheash is a research scholar of Sadakathullah Appa College, affiliated to Manonmaniam Sundaranar University, Tirunelveli, India. He is an Assistant Professor of Information Technology, received his MCA degree in Computer Applications from Bishop Heber College, Bharathidasan University, Trichy, India in 2007. He received his

MPhil degree in Computer Science and an M. Tech in Information Technology from Manonmaniam Sundaranar University, Tirunelveli, India in 2012 and 2014 respectively. His major research interests include Digital Image Processing, Document Image Analysis and Character Recognition of Tamil language palm leaf manuscripts.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/8DWSQ

TEXT LINE SEGMENTATION IN TAMIL LANGUAGE PALM LEAF MANUSCRIPTS – A NOVEL APPROACH

Dr. M. Mohamed Sathik [1], **R. Spurgen Ratheash** [2]

[1] Principal, Sadakathullah Appa College, Tirunelveli, Tamil Nadu, India. Manonmaniam Sundaranar University, Tamil Nadu, India.
[2] Research Scholar, Reg. No: 12334, Sadakathullah Appa College, Tirunelveli, Tamil Nadu, India. Manonmaniam Sundaranar University, Tamil Nadu, India.

Abstract. Segmentation of text lines from palm leaf manuscripts is an essential prior activity for character recognition. The scribers writing style creates intricacy in text line segmentation by low space between text lines and elongated characters placed in the text lines. Inefficient text line segmentation makes unproductive when promoting to character segmentation and character recognition process. The researchers have proposed a new way of text line segmentation algorithm named as Text Line Slicing algorithm for Tamil palm leaf manuscripts. This article explores text line segmentation from the scratch of preprocessing. The identification, segmentation of touching and overlapping text lines by an elongation of the character proves uniqueness of an algorithm. Text Line Slicing provides successful result in Tamil text line segmentation amidst several challenges. This outcome is an evidence of novelty among aplenty of text line segmentation methods in Tamil and other language palm leaf manuscripts.

Keywords: binarization, line segmentation, obstacle, palm leaf, preprocessing, Tamil manuscripts, text line slicing, touching line, overlapping lines.

Introduction

Tamil, one of the most ancient classical languages, has its inscriptions that date back to 600 BC. During the ancient period, many of the literatures, medicinal hints, astrology and much more essential information are present in palm leaves. Lifespan of preserved palm leaf manuscripts is minimum years. The reasons for the dilapidated condition of the manuscripts are weather, fungal and termite. The information of palm leaf manuscripts can be preserved when they are copied into new leaf by the scribers. The palm leaf writing is unique skill that needs patience, practice, and training to the writers. Generally, the Tamil palm leaf manuscripts are written by a pointed needle metal named as stylus [1]. Many of the text lines are not in exact straight line as typed letters. Writing the Tamil characters with stylus creates extension in shapes of the character and makes to touch with the succeeding text lines (Fig. 1). The stylus writing produces the challenges of low space, cross line, touching and overlapping text lines in the process of character recognition from the text images [2]. The successful text line segmentation can lead accuracy when character images are recognizable. In Tamil language, elongation categorizes the strokes as upper part and lower part of the text lines. Tamil character strokes are elongated by nature or in the course of writing by the writers. The letters such as து /thu/, ற/ra/ are the examples of downward and ணீ/nee/, ரீ /ree/ are upward elongated characters respectively [3]. The impediment in line segmentation starts from the text lines during the segmentation when they have elongated characters. The proposed Text Line Slicing (TLS) algorithm identifies elongation of the character as an obstacle. The existence or prolongation of an obstacle categorizes the space between the text lines as space without obstacle and space with obstacle. An obstacle touches with the succeeding text lines are considered touching text lines and it pervades the text zones of lines in the next text line is known as overlapping text lines. TLS shows an excellence in segmenting text lines in both the ways.

The rest of the paper is organized in section 2 for related works of various text line segmentation methods. In Section 3 discusses about the proposed text line segmentation algorithm. Section 4 presents the results and discussions are presented and the paper is concluded in section 5.

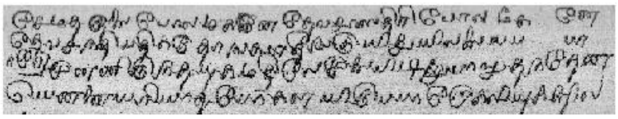


Fig. 1 Tamil palm leaf manuscript

Related works

The touching lines were formed as blocks that can be separated by fixing the bounding box around the components that were present in the lines. If the component spreads over k lines, divide them vertically and fix the cutting locations by using clustering algorithms [4]. In [5], the horizontal projection profile was used to identify the white pixels in the touching text lines and the subsequent histogram was constructed. However, this method detects text line with moderate accuracy. The modified A* Path Planning algorithm has two functions such as intensity difference cost function and vertical distance cost function to construct the segmented line between the touching and overlapping text lines in Khmer palm leaf manuscripts with good recognition accuracy [6]. The minimum horizontal projection values are calculated for each row to identify the touching or overlapping text lines at the end portion of a line and the starting portion of the next line. The minimum values fix the indexing point to segment the overlapping text lines in the printed Tamil scripts [7]. In Adaptive Partial Projection method, the image is divided by vertical columns and the histogram is applied by smoothing to segment the lines [8]. The component break procedure is used to identify the average of the right and left regressions in the connected component to segment the line on handwritten and printed documents [9, 12]. The line adjacency graphs are used to reduce the text components as small called as segments when a set of vertical block runs by Run Length Encoding. The overlapping lines are identified from the connection of vertical runs [10]. The fringe map generated on the binary text-image and hen the peak fringe numbers are located. The filters are applied on the map images and clustering the peak fringe numbers to create the segmentation path between the text lines [11]. An energy map is used to extract the seams from horizontal and vertical orientations and the lower value pixels are removed. The higher value pixel information can be regarded as residing in the text line. The Signed Distance Transform is used to indicate the nearer points and identify the space between the text lines [13].The midpoint detection based method is used to segment the text lines, words, skewed lines, overlapped lines and connected components in handwritten Gurmukhi scripts [14]. An improved piece wise projection based method for handwritten documents to improve an execution speed. The signal approximation using Fourier series and statistical approaches are applied on text lines for segmentation [15]. The Gaussian filter is used to blur the images and discard the pixels which have no local maximum when compare the neighbors. The second order derivatives are used to detect and segment the text lines [16]. A labeling is used to denote the position of an object in a set of features. This tracking method used to segment the text lines in handwritten documents. The method has given failure result in connected components from different lines are close [17]. The text line segmentation based on peaks and valleys in projection profile method yields better result in good spacing text lines. Considering touching lines the result is low in text line segmentation [18]. The starting and ending point of the space can be identified by smoothed horizontal ink density histogram in A* path planning algorithm with cost functions to segment the text lines [19].

Proposed method

The text line segmentation starts from the initial process of preprocessing. The preprocessing makes the colour images of palm leaf manuscript in to binary images as it is easier to execute an algorithm to identify an obstacle using image enhancement methods. After preprocessing, the TLS algorithm for text line segmentation is applied on the binary images to analyze the space between the text lines and also categorize the space as space with obstacle and space without obstacle. In the case of space without obstacle, considered as standard images because don't have difficulty in text line segmentation. Space with obstacle categorized as touching and overlapping text lines by an obstacle as in (Fig.2).

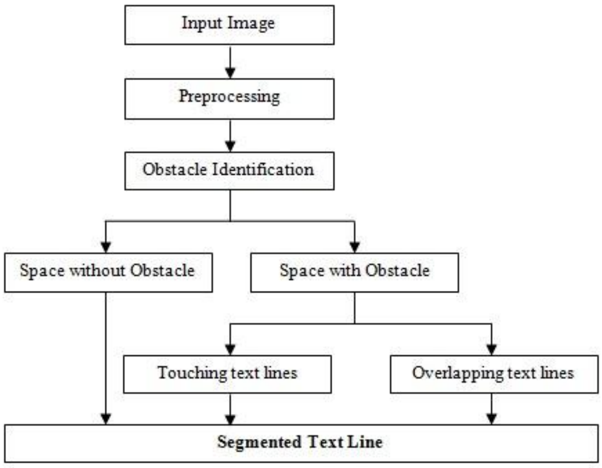


Fig. 2 Architecture of the proposed work

Preprocessing

Digitized palm leaf manuscripts image have dark leaf colour as background and text as foreground. The presence of pickup noise in palm leaf manuscript images when scanned or photo taken by digital camera can be reduced by sharpening and morphology methods. The noise removal is a necessary step to obtain useful information from digital text images. In pre-processing, the background has converted as black and foreground text as white in colors from the range of 0 to 255, into 0 and 1 only then the characters can be clear to process. The background removal and morphology are the methods to promote the images for text line segmentation.

Background removal

Binarization is a process of assigning 0s and 1s using fixed threshold value. The fundamental idea of the fixed binarization method [21] is in the following relation. The background of palm leaf manuscripts can be taken as black by 0 and the foreground text as white by 1. T shows global threshold value 50. The pre-processed images are shows in (Fig. 3).

$$g(x, y) = \begin{cases} 1 \; if \; f(x,y) \geq T \\ 0 \; otherwise \end{cases} \quad (1)$$

Morphology

In morphological operations, dilation and erosion are the fundamental operations. Addition of pixels with the boundaries of text objects in an image is known as dilation. The reversal of this operation, i.e., extricating the pixels from the text object boundaries is termed as erosion

[20]. In order to process the text-image in palm leaf manuscripts, the pixels may be added or removed depending on the size and shape of the text. In grayscale morphology, the images are mapped into the Euclidean space or grid E S∪{*r,-r*}, the grayscale erosion of the palm leaf image *i* by text boundaries *b* as in the following relation.

$$(i\ominus b)(x)= \underset{y\in B}{\propto}\ [i(x+y)-b(y)] \qquad (2)$$

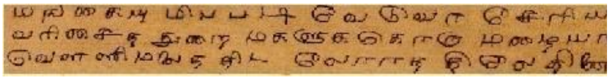
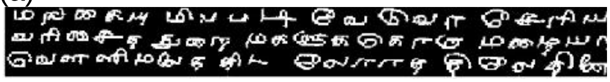
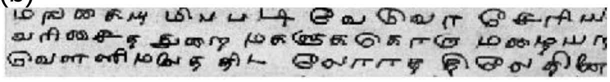

(a)

(b)

(c)

Fig. 3 Preprocessed image (a) Input image
(b) Background removal (c) Morphology

Text Line Slicing (TLS)

The text line segmentation in Tamil palm leaf manuscripts is a Herculean task and it influences till an end of the character recognition process. An absence of text line segmentation process is not possible for the successful character segmentation and character recognition in Tamil palm leaf manuscripts. The TLS applied on the preprocessed binary palm leaf manuscript images to segment the text lines. The new way of approach in text line segmentation of Tamil palm leaf manuscript images by an obstacle presence between the text lines. Whenever the strokes of the character exceed from the text zone and extend in the space between the lines are known as an obstacle (Fig. 4).

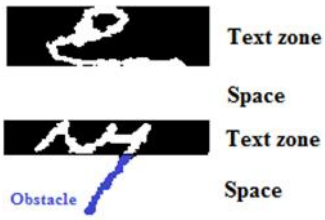


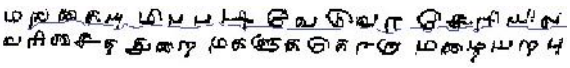
Fig. 4 Obstacle

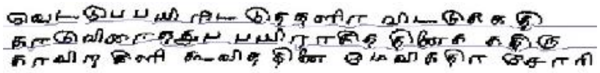
Space without Obstacle

The TLS, the text lines of Tamil palm leaf manuscripts were the text line has enough space to the subsequent text lines or an elongation of character does not reach the below text line as in (Fig 5) are considered as space without obstacle or standard category. TLS can segment these text lines without any complication.

Space with Obstacle

The presence of an obstacle in the space between the text lines can be categorized as two by the length of an obstacle that helps to decide whether touching or overlapping text line. In Tamil character, an


(a)



(b)

Fig.5 Standard text lines (a) no obstacle (b) obstacle not reached to the below text line.

obstacle is an important part to decide the character. The length of an obstacle extends and reached to the subsequent text line is known as touching text lines as in (Fig. 6). The first line character "ய /you/" is touching with the next line character "த/tha/". In Tamil, the text line segmentation is complicated because if we ignore an obstacle of the character "ய " becomes "ய" and if we cut an obstacle in a fixed length the second line character "த/tha/" becomes "தி/thi/". The overlapping text lines also have the same wrong prediction problem as touching text lines when we precede by existing text line segmentation algorithms. The proposed TLS solves the problem of touching and overlapping text lines by fixing the cutting edge at the end of an obstacle and also prevent wrong predictions of the character.

(a) (b) (c)

Fig.6 Obstacle defined in the textine (a) touching characters (b) ignoring obstacle (c) cut in fixed length

The purpose of text line segmentation is to precede the character segmentation. The touching and overlapping text lines make complicate the text line segmentation and also makes the further process unproductive. An overlapping text line builds complication in text line segmentation. An obstacle pervades the text zone of subsequent lines and mixed up with the character strokes that may precede wrong prediction of the character or different than expected character. The first text line character "ந/na/" extends its elongation up to second text line character "ம/ma/" as in (Fig.7). The proposed line segmentation algorithm TLS segments the character "ந/na/" by fixing the cutting edge in vertically minimum value on the obstacle.

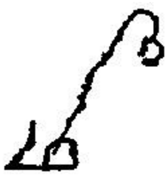
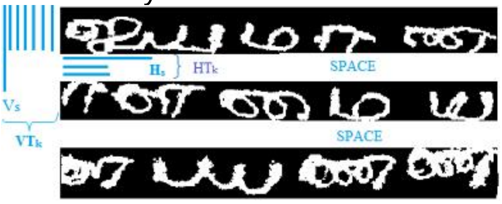


Fig. 7 Overlapping text line

3.2.3 Text Line Slicing algorithm for Tamil text lines segmentation

The proposed TLS line segmentation algorithm identifies an extension of character strokes using four variables such as Vertical space (V_s), Horizontal space (H_s), Vertical Track (*VT*) and Horizontal Track (*HT*). The variable V_s used to count zeros vertically to know the stroke of a character exists in text line of binarized Tamil palm leaf manuscript image. The total columns count of zeros assigned to the variable *VT* and compared with the threshold by the value of 1 denotes that the space has no obstacle and 0for an obstacle. The variable H_s is used to count the zeros in horizontally and the total value assigned to *HT* that compare with threshold value. Three values are used to decide whether an obstacle present or not such as zero defines the space between the character in text line; one defines the character has an obstacle and two defines the space not found that means character exists concurrently.



(a)

(b)

Fig. 8 Text Line Slicing algorithm (a) implementation on Tamil palm leaf manuscript (b) cutting edge

The obstacle creates touching text lines can be defined by Connected Component (CC). The connectivity of the character is calculated by the weight of the character using CC. The touching and overlapping text line characters are considered as single character when they are connected to each other as in (Fig. 8). (a) An algorithm implementation proves obstacle identification in the space between the text lines and defines the category of connected characters by connected component and calculates the weight for the character. (b) When the minimum weight of the character identifies the TLS algorithm implements cutting edge to segment the connected text lines. Cutting edge is a breaking point of touching characters in text lines. The CC provides continuation of the character strokes and also vertical stroke values. The minimum value of the character stroke is known as end of an obstacle that has to be fixed as a cutting edge for the text lines.

Results and Discussions

The proposed text line segmentation algorithm TLS proves novelty in text line segmentation in Tamil palm leaf manuscripts as in (Table 1). The touching text lines can be segmented without changing the original shape and preserve all strokes without any loss of information in the characters. In this research, the Tamil language palm leaf manuscripts text lines are categorized by the writing methods of the writer. They are considered as challenges that described by the length of an obstacle and space between the text lines. The segmentation accuracy defines the novelty of an algorithm. Although many systems have been found to recognize the Tamil alphabet, this method has introduced an innovative method of recognizing the Tamil alphabet in palm leaf manuscripts. This has created the process of segmenting the text lines and then segmenting the characters and then recognizing the Tamil characters more accurately. For this research, Tamil palm leaf manuscripts have been taken as 2200 x 300 pixel dimensions for width and height respectively with 300 pixels of resolution. In this section, for the clear vision of challenges, the researchers show the image in a same size of 190 x 280 pixel dimensions for width and height with 100 pixels of resolution. The text line segmentation algorithms are applied on the Tamil palm leaf manuscript images which have all challenges and the results are compared. TLS produced notable results on those challenges than that of other two algorithms.

Table 1 Performance of TLS on Tamil palm leaf manuscripts text lines

	Detection Rate	Segmentation Accuracy	Performance Accuracy
TLS for Standard Images	94.32	98.91	95.96
TLS for Touching Images	92.45	96.58	94.53
TLS for Overlapping Images	90.56	95.84	92.47

CONCLUSION

The text line segmentation is the most important and an initial major process in Optical Character Recognition. The researchers provide a novel approach for Tamil language text line segmentation in palm leaf manuscripts through the proposed algorithm. The touching and overlapping text lines are the major challenges in Tamil palm leaf manuscripts that can be successfully resolved by TLS and it provides an error-free way for other challenges as well. The TLS has advantages from fairly simple to implement, quite fast, and robust for Tamil language palm leaf manuscripts. In future, the TLS can be extended to apply on other language palm leaf manuscripts and Tamil language epigraphs to recognize the characters.

REFERENCES

1. Udaya Kumar, D, Sreekumar, G, V, Athvankar, U, A., “Traditional writing system in Southern India — Palm leaf manuscripts”, IDC.IITB, pp. 1–7, 2009.
2. Ayush Pradhan, Sidharth Behera, and Pushpalata Pujari, "Comparative Study on Recent Text Line Segmentation Methods of Unconstrained Handwritten Scripts", ICECDS, IEEE, pp 3853 - 3858, 2017.
3. Kathirvalavakumar, Thangairulappan, and Karthigaiselvi Mohan, “Efficient Segmentation of printed Tamil Script into characters Using Projection and Structure”, ICIIP, IEEE, pp 484 – 489, 2017.
4. Dona Valy, Michel Verleysen, and Kimheng Sok, “Line Segmentation Approach for Ancient Palm Leaf Manuscripts using Competitive Learning Algorithm”, ICFHR, IEEE, pp.108 – 113, 2016.
5. B. Kiruba, A. Nivethitha and Dr. M. Vimaladevi, “Segmentation of Handwritten Tamil Character from Palm Script using Histogram Approach”, IJIFR, pp.6418-6424, 2017.
6. Dona Valy, Michel Verleysen, and Kimheng Sok, “Line Segmentation for Grayscale Text-images of Khmer Palm Leaf Manuscripts”, IEEE, 2017.
7. Thangairulappan Kathirvalavakumar, and Karthigai Selvi, “Efficient Touching Text Line Segmentation in Tamil Script Using Horizontal Projection”, Springer, pp.279–288, 2013.
8. Rapeeporn Chamchong, Chun Che Fung, “Text Line Extraction Using Adaptive Partial Projection for Palm Leaf Manuscripts from Thailand”, IEEE, pp. 586 – 591, 2012.
9. Himanshu Jain, Archana Praveen Kumar, "A Bottom up Procedure for Text Line Segmentation of Latin Script", IEEE, pp.1182-1187, 2017.
10. QuangNhatVo, GueeSang Lee, "Dense Prediction For Text Line Segmentation in Handwritten Document Images", IEEE, pp. 3264-3268, 2016.
11. Vijaya Kumar Koppula, AtulNegi, "Fringe Map Based Text Line Segmentation of Printed Telugu Document Images", IEEE, pp. 1294-1298, 2011.

12. Ines Ben Messaoud, Hamid Amiri, Haikal El Abed, Volker M¨argner, " A Multilevel Text line Segmentation Framework for Handwritten Historical Documents", IEEE, pp.515-520, 2012.
13. Xi Zhang, Chew Lim Tan, "Text Line Segmentation for Handwritten Documents Using Constrained Seam Carving", IEEE, pp. 98-103, 2014.
14. Rodolfo P. dos Santos, Gabriela S. Clemente, Tsang Ing Ren, and George D.C. Calvalcanti, "Text line segmentation based on morphology and histogram projection", ICDAR, IEEE, pp. 651 – 655, 2009.
15. Vishal Chavan, Kapil Mehrotra, "Text Line Segmentation of Multilingual HandwrittenDocuments Using Fourier Approximation", ICIIP, IEEE, pp. 250-255, 2017.
16. David Aldavert and Marc¸al Rusinol, "Text Line Detection and Segmentation using Second-OrderDerivatives", IAPR International Workshop on Document Analysis Systems Manuscript, IEEE, pp. 293-298, 2018.
17. Setitra, I., Hadjadj, Z., and Meziane, A., "A Tracking Approach for Text Line Segmentation in HandwrittenDocuments", ICPRAM, SCITEPRESS, pp.193-198, 2017.
18. Banumathi. K., Jagadeesh Chandra A. P., "Line and word Segmentation of Kannada HandwrittenText documents using Projection Profile Technique", ICEECCOT, IEEE, pp. 196-201, 2016.
19. Olarik Surinta, Michiel Holtkamp, Faik Karabaa, Jean-Paul van Oosten, Lambert Schomaker and Marco Wiering, "A* Path Planning for Line Segmentation ofHandwritten Documents", ICFHR, IEEE, pp.175-180, 2014.
20. Rodolfo P. dos Santos, Gabriela S. Clemente, Tsang IngRen and George D.C. Calvalcanti, "Text Line Segmentation Based on Morphology and Histogram Projection", ICDAR, IEEE, pp. 651-655, 2009.
21. Nobuyuki Otsu, “A Threshold Selection Method from Gray-Level Histograms”, IEEE, pp. 62 - 66, 1979.

HIGHLIGHTED ARTICLE

DOI: 10.34256/irjt21210

INTERNATIONAL RESEARCH JOURNAL OF TAMIL
சர்வதேசத் தமிழ் ஆய்விதழ்

OPEN ACCESS

பனை ஓலைச்சுவடிகளில் எழுதப்பட்டுள்ள தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்துதல்

இரா. ஸ்பர்ஜன் ரத்தீஷ் [அ, *], மு.முகம்மது சாதிக் [அ]

[அ] சதக்கத்துல்லாஹ் அப்பா கல்லூரி, (மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகத்துடன் இணைவு பெற்றது), திருநெல்வேலி-627011, தமிழ்நாடு, இந்தியா.

Tamil Character Recognition in Palm Leaf Manuscripts

R. Spurgen Ratheash [a,*], M. Mohamed Sathik [a]

[a] Sadakathullah Appa College, (Affiliated to Manonmaniam Sundaranar University), Tirunelveli-627011, TamilNadu, India.

* Corresponding Author: spurgen@gmail.com

Received : 31-01-2021
Revised : 15-02-2021
Accepted : 21-02-2021
Published : 26-04-2021

Check for updates

ABSTRACT

Tamil characters are of historical significance. The shapes of the character and its writing continued to be changed by several reformers in each century until the period of the nineteenth century. The Tamil characters in the palm leaf manuscripts are based on the author's writing style and strokes in use at the time. Reading palm leaf manuscripts is a challenge for the modern generation who are unaware of the character strokes and writing patterns written in earlier times, and the younger generation neglects to read and understand the contents written in palm leaf manuscripts. To read Tamil palm leaf manuscripts it is necessary to remember the shapes of the character that has changed over time and to compare and recognize the characters. This research paper explains how to recognize the Tamil characters written in palm leaf manuscripts by computer. By this research, the Tamil characters can be compared with different strokes and shapes and the exact character can be recognized accurately and quickly.

Keywords: Tamil manuscripts, Tamil character, Text line segmentation, Character segmentation, Character recognition

ஆசிரியர் குறிப்பு

இரா. ஸ்பர்ஜன் ரத்தீஷ், சதக்கத்துல்லாஹ் அப்பா கல்லூரியின் ஆராய்ச்சி மாணவர் ஆவார். இது இந்தியாவின் தமிழ்நாட்டில் உள்ள திருநெல்வேலி மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகத்துடன் இணைக்கப்பட்டுள்ளது. எண்மின் மயமாக்கல் பட செயலாக்கம், ஆவண பட பகுப்பாய்வு மற்றும் தமிழ் மொழி பனை ஓலை கையெழுத்துப் பிரதிகளின் எழுத்து அங்கீகாரம் ஆகியவை அவரது முக்கிய ஆராய்ச்சி ஆர்வங்களில் அடங்கும்.

முனைவர் மு. முகமது சாதிக், இந்தியாவின் திருநெல்வேலியில் உள்ள சதக்கதுல்லாஹ் அப்பா கல்லூரியின் முதல்வராக உள்ளார். இந்தியாவின் திருநெல்வேலியில் உள்ள மனோன்மனியம் சுந்தரனார் பல்கலைக்கழகத்தில் கணினி அறிவியல் மற்றும் கணினி அறிவியல் மற்றும் தகவல் தொழில்நுட்பத்தில் தேர்ச்சி பெற்ற இரண்டு முனைவர் பட்டங்களைப் பெற்றார். அவர் பல தேசிய மற்றும் சர்வதேச கருத்தரங்குகள், மாநாடுகளில் கலந்து கொண்டு ஏராளமான ஆய்வுக் கட்டுரைகளை வழங்கியுள்ளார். பல சர்வதேச பத்திரிகைகளில் வெளியீடுகளுடன், 40 க்கும் மேற்பட்ட ஆராய்ச்சி

அறிஞர்களுக்கு வழிகாட்டியதைத் தவிர இரண்டு புத்தகங்களை வெளியிட்டுள்ளார். இந்தியாவின் தமிழ்நாட்டில் உள்ள பல்வேறு பல்கலைக்கழகங்கள் மற்றும் தன்னாட்சி கல்லூரிகளின் பாடத்திட்ட மேம்பாட்டுக் குழுவில் உறுப்பினராக உள்ளார். மெய்நிகர் உண்மை, எண்மின் மயமாக்கல் பட செயலாக்கம் மற்றும் உணரி பிணையம் ஆகியவை அவரது ஆராய்ச்சியின் சிறப்புப் பகுதிகள்.

முன்னுரை

உலகின் செம்மொழிகளில் தமிழ் மொழியும் ஒன்றாகும். தொன்மை வாய்ந்த இம்மொழி ஆசிய கண்டத்தில் குறிப்பாக இந்தியா, சிங்கப்பூர், இலங்கை ஆகிய நாடுகளில் ஆட்சி மொழியாகவும். மேலும், மொரீசியஸ் போன்ற பிற கண்டங்களில் உள்ள நாடுகளில் கலாச்சார மொழியாகவும் ஏற்றுக்கொள்ளப்பட்டுள்ளது. முதன்முதலில் முன்னோர்களால் ஒலி வடிவில் தொடங்கி பின் குறியீடுகளால் எழுதப்பட்டு அதன்பின் ஓலைச் சுவடிகளிலும், கற்களிலும் பதிக்கப்பட்ட தமிழ் எழுத்துக்கள் தற்போது உருவத்தில் அநேக மாற்றம் பெற்று உயிரெழுத்துக்கள் பன்னிரெண்டும், மெய்யெழுத்துக்கள் பதினெட்டும் ஆயுத எழுத்து ஒன்றும் சேர்ந்து மொத்தம் 247 ஆக எழுத்து, எண், பேச்சு மற்றும் எழுத்து வழக்கிலும், கணிப்பொறிப்பதிப்பிலும் பயன்படுத்தப்பட்டு வருகிறது. பழங்காலத்தில் எழுதப்பட்ட பனை ஓலைச் சுவடிகள் அறிவியல், மருத்துவம், புவியியல், சோதிடம், இலக்கியம், நாடகம் என அனைத்து வகையிலும் முன்னோர்களின் அறிவாற்றலால் எழுதப்பட்ட அறிவுப் பொக்கிசமாக திகழ்கிறது. பனை ஓலைச்சுவடிகள் என்பது பனை மரத்தின் ஓலைகளை நறுக்கி பல நிலைகளில் பக்குவப்படுத்தி அதில் 5 முதல் 6 எழுத்து வரிகளை இரும்பினால் செய்யப்பட்ட ஊசியால் எழுதி 40 முதல் 50 ஓலைகளை ஒன்றாக கட்டி வைத்துள்ள தொகுப்பு ஆகும். அவை படம் 1-இல் காட்டப்பட்டுள்ளன. இந்த அறிவுப் பொக்கிசத்தை பாதுகாப்பது மிகவும் கடினமானதாகவும், அதிக செலவினங்களை ஏற்படுத்தக்கூடியதாகவும் மற்றும் இடத்தை அதிகம் ஆக்கிரமிப்பதாகவும் உள்ளது. எண்மின்மயமாக்கல் (Digitization) என்ற தொழில் நுட்ப முறைப்படி அநேக ஓலைச்சுவடிகள் வருடிநகலி (Scanner) உதவியுடன் படமாக்கப்பட்டு கணினியில் சேமித்து வைப்பதன் வழியாக எளிதாக கையாளவும், அனுப்பவும், பலமுறை பார்த்தாலும் பழுதாகமலும், பார்ப்பதற்கு ஏற்ப பெரிதாக்கவும், பராமரிப்பு செலவில்லாததாகவும், குறைந்த அளவு நினைவக இடத்தில் அதிக அளவு படங்களை சேமிக்கவும் முடிகிறது. ஆனாலும் அதில் எழுதப்பட்டுள்ள ஒரு சில எழுத்துக்களின் தோற்றம், எழுத்து முறை தற்போது உள்ள தமிழ் எழுத்துக்களை விட வேறுபட்டிருப்பதாலும், எழுத்துக்கள் ஒன்றையொன்று ஒட்டிக் கொண்டும், வார்த்தைகளுக்கு இடையே இடைவெளி இல்லாததாலும், அதனை வாசித்து அர்த்தம் விளங்கிக் கொள்ளவதில் சிரமம் ஏற்படுவதாலும் எண்மின்மயமாக்கல் பயனற்றதாய் போகும் அபாய நிலைக்குத் தள்ளப்பட்டுள்ளது. ஓலைச்சுவடிகளின் எழுத்துக்களை வாசிப்பது அதற்கென பிரத்யேக பயிற்சி பெற்றவர்களால் மட்டுமே சாத்தியமாகிறது.

படம் 1 ஓலைச்சுவடிகள்

இதனால் அநேக ஓலைச்சுவடிகள் புத்தகங்களாகப் பதிப்பிக்கப்படாமல் அழியும் நிலை நிலவி வருகிறது. இந்த ஆராய்ச்சி பனை ஓலைச் சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை கணினியை கொண்டு அடையாளப்படுத்தி அதனை கணினி தட்டச்சு செய்யப்பட்ட எழுத்துக்களாக காட்சிப்படுத்துவதைப் பற்றி விளக்குகிறது. இதனால் ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை

வேகமாக வாசிக்கவும், மிகச் சுலபமாக புரிந்து கொள்ளவும் முடியும் என்பதை இந்த ஆராய்ச்சி கட்டுரை விளக்குகிறது.

உரையாசிரியர்களின் விளக்கங்கள்

தமிழ் எழுத்துக்கள் எழுதப்பட்ட முறையும், பனை ஓலைச்சுவடிகளில் எழுதப்பட்டுள்ள தமிழ் எழுத்துக்கள் கொண்டுள்ள பல்வேறு வகையான சவால்களும், ஓலைச்சுவடிகளில் பயன்படுத்தப்பட்டுள்ள எண்களின் முறைமையையும் விளக்குகிறது (Kattalai Kailasam, 2019). கையால் எழுதப்பட்ட தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்துவதில் சுருளல் நரம்பணுப் பிணையம் (Convolutional Neural Networks) முறை பிரயோகப்படுத்தப்படுவது பற்றி குறிப்பிடுகிறது (Deepa et al., 2019). ஒரு நபரால் எழுதப்பட்ட பனை ஓலைச் சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை சுருளல் நரம்பணுப்; பிணையம் வழியாக அடையாளப்படுத்தப்பட்டு உள்ளது (Sabeenian et al., 2019). பூலியன் அணிக்கோவை (Boolean Matrix) மற்றும் மூச்சி முதல் தேடல் (Breadth First Search) நுட்பத்தைப் பயன்படுத்தி ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்தி உள்ளனர். அதில் சாய்ந்த எழுத்துக்கள் போன்ற அடிப்படை பிரச்சனைகளை கையாளுவதில் உள்ள சிக்கல்கள் விளக்கப்பட்டு உள்ளது (Vellingiriraj, & Balasubramanie, 2014). கையால் எழுதப்பட்ட தமிழ் எழுத்துக்களை அடையாளப்படுத்துவதில் செயற்கை நரம்பணுப்பிணையம் (Artificial Neural Networks) முறை பயன்படுத்தப்பட்டு உள்ளது (Banumathi, & Nasira, 2011). ஆழ்கற்றல் (Deep Learning) மற்றும் பெரிய தரவு (Big Data) முறையில் கற்கும் திறன் புகுத்தி தமிழ் எழுத்துக்களை அடையாளப்படுத்துவதில் கருத்தாய்வு செய்யப்பட்டு உள்ளது (Kannan, & Subramanian, 2015). பிம்பவியல் (Image processing) வழியாக பிம்பம் மேம்படுத்துதல் (Image Enhancement), பிம்பம் பிரித்தல் (Image Segmentation) மற்றும் பிம்பம் மறுசீரமைப்பு (Image Restoration) ஆகிய நுட்பத்தைப் பயன்படுத்தி தமிழ் எழுத்துக்களை கணினி அடையாளப்படுத்துவதில் கருத்தாய்வு செய்யப்பட்டு உள்ளது (Challa, & Mehta, 2017). தென்கிழக்கு ஆசிய நாடுகளில் உள்ள பனை ஓலைச்சுவடி எழுத்துக்களை ஆவண பட பகுப்பாய்வு (Document Image Analysis)நுட்பத்தின் வழியாக அடையாளப்படுத்துவது பற்றி விவாதிக்கப்பட்டு உள்ளது (Kesiman et al., 2018). முதன்மை உபகரண பகுப்பாய்வு (Principal Component Analysis) நுட்பத்தை ஆழ்சுருளல் நரம்பணுப் பிணையம் (Deep Convolutional Neural Networks) நுடபத்துடன் இணைத்து கையால் எழுதப்பட்ட தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்தப்பட்டு உள்ளது (Sornam, & Vishnu Priya, 2018).

முன்மொழியப்பட்ட ஆராய்ச்சி முறை

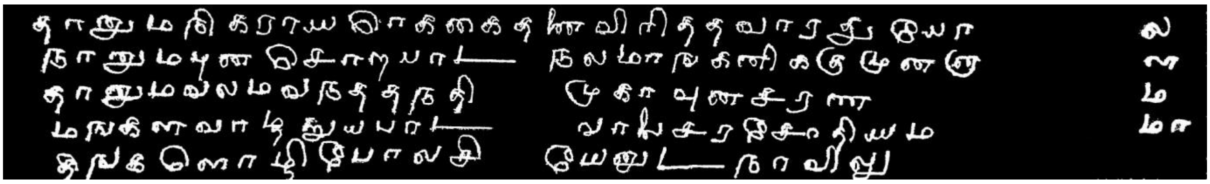
பனை ஓலைச் சுவடிகளில் தமிழ் எழுத்துக்கள் இரும்பாலான ஊசியைக் கொண்டு எழுதப்படுவதாலும், அகலம் குறைந்த ஓலைகளை மடியில் வைத்து எழுதுவதாலும் எழுத்து வரிசைகள் தட்டச்சு செய்தது போன்று நேர்கோட்டில் அமைந்து இருப்பதில்லை மற்றும் எழுத்துக்கள் ஒன்றை ஒன்று ஒட்டிக் கொண்டும், ஒன்றின் மேல் ஒன்று பிண்ணியும் காணப்படுகிறது. ஊசி கொண்டு எழுதப்பட்ட எழுத்துக்கள் பார்ப்பதற்கு தெளிவாக தெரிவதில்லை அதனால் அதன் மீது விளக்கு புகை, கரி ஆகியவை பூசப்படுகிறது. அவை படம் 2-இல் காட்டப்பட்டுள்ளது. இதனால் சரியாக பூசப்படாத இடங்களில் எழுத்துக்களின் தொடர்ச்சி விடுபடுவதால் கணினியில் அடையாளப்படுத்தும் போது எழுத்துக்களை உடைந்தாகக் கருதுகிறது. இதனை சரி செய்வதற்காக இந்த ஆராய்ச்சியில் பிம்பம் மேம்படுத்துதல் (Image Enhancement) நுட்பம் உபயோகப்படுத்தப்பட்டு உள்ளது.

படம் 2. அ. எழுத்தாணியால் எழுதப்பட்ட தமிழ் எழுத்துக்கள், ஆ. கரிபூசப்பட்ட எழுத்துக்கள்

சரியான எழுத்துக்களை ஒப்பிட்டு கணினி வழி அடையாளப்படுத்த ஒட்டிய தமிழ் எழுத்துக்களை பிரித்து தனி எழுத்து பிம்பங்களாக மாற்றுவது இன்றியமையாதது. இதற்காக எழுத்து வரிசை பிரித்தல் (Text Line Segmentation), எழுத்து பிரித்தல் (Character Segmentation), எழுத்து அடையாளப்படுத்துதல் (Character Recognition) என்ற படி நிலைகள் உபயோகப்படுத்தப்பட்டுள்ளது.

பிம்பம் மேம்படுத்துதல்

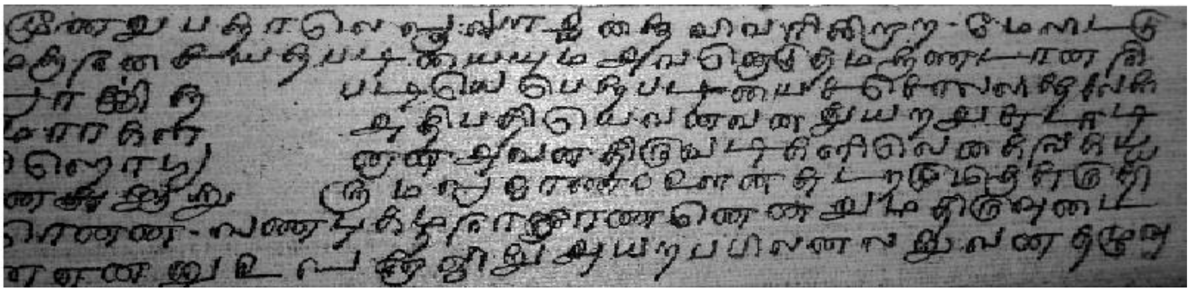
பனை ஓலைச்சுவடிகளின் தமிழ் எழுத்துக்களை அடையாளப்படுத்த சில முன் செயலாக்க முறைகள் பின்பற்றப்படுகின்றன. ஓலைச்சுவடிகள் வருடி நகலி மூலமாகவோ அல்லது எண்ணியல் நிழற்படக் கருவி (Digital Camera) மூலமாகவோ எண்மின்மயமாக்கப்படுகிறது. ஓலைச்சுவடிகளில் ஓலையின் நிறம் மற்றும் நாள்படிந்த கறைகளும் இருப்பதால் பின்னணி நீக்கத்தின் (background removal) போது எழுத்துக்களை தெளிவாக தெரியவிடுவதில்லை. இதனால் சுவடிகளின் மின்பிம்பங்கள் பிம்பம் மேம்படுத்தலுக்கு உட்படுத்தப்படுகிறது. அதன்படி எழுத்துக்கள் தெளிவைப் பெறுகிறது. அதன்பின் பின்னணி நீக்க முறையின்படி எழுத்துக்கள் வெண்மை நிறமாகவும், பின்னணியம் கறுப்பு நிறமாகவும் மாற்றப்படுகிறது அவை படம் 3-இல் காட்டப்பட்டுள்ளது. கடைசியாக உருவவியல் (Morphology) நுட்பத்தைப் பயன்படுத்தி எழுத்துக்களின் ஓரங்கள் மற்றும் அதன் கோடுகள் வலுவேற்றப்படுகின்றன. இத்தகைய நுட்பத்தினால் வலுப்பெற்ற எழுத்துக்களுடன் பிம்பம் அடுத்த கட்ட செயலாக்கத்திற்கு தயாராகிறது.



படம் 3 கருப்பு, வெள்ளை என இரு நிறங்களாக மாற்றப்பட்ட தமிழ் ஓலைச்சுவடி

எழுத்து வரிசையைப் பிரித்தல்

பனை ஓலைச்சுவடிகளில் எழுத்துக்கள் தட்டச்சு செய்வது போன்று சரியான வரிசையில் இல்லாமல் மேலும் கீழுமாக வளைந்து காணப்படும். ஒருசில நெடில் எழுத்துக்கள் நீண்டு கீழ் உள்ள எழுத்து வரிசையுடன் ஒட்டி காணப்படுகிறது. அவை படம் 4-இல் காட்டப்பட்டுள்ளது. தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்த ஒவ்வொரு எழுத்தும் தனி படமாக இருப்பது அவசியமாகிறது. அதன்படி மேல் உள்ள எழுத்து வரிசையின் எழுத்துக்கள் கீழ் வரிசையுடன் ஒட்டியிருந்தால் அதனை பிரிப்பது இன்றியமையாததாகிறது.



படம் 4 எழுத்து வரிசை ஒன்றுடன் ஒன்று ஒட்டியிருத்தல்

பிரிக்காமல் இருக்கும் பட்சத்தில் அசலான எழுத்துக்கள் வேறொரு எழுத்துக்கள் போன்று தோன்றிவிடும். இதற்காக எழுத்து வரிசை பிரிப்பதற்கென இக்கட்டுரையின் ஆராய்ச்சியாளர்கள் பிரத்தியேகமாக எழுத்து வரி வெட்டுதல் வழிமுறையை (Text Line Slicing Algorithm) கண்டுபிடித்துள்ளனர். அதன்படி எழுத்துக்கள் நீண்டு வருமாயின் அது பிற எழுத்துக்களுடன் சேருமிடத்தில் பிரித்து விடுகிறது. இதனால் எழுத்துக்களின் வடிவங்கள் மாறுபடாமல் அசலான எழுத்துக்களுடன் எழுத்து வரிசை பிரித்தெடுக்கப்படுகிறது, அவை படம் 5-இல் காட்டப்பட்டுள்ளது. எழுத்து வரி வெட்டுதல் வழிமுறை மற்ற ஆராய்ச்சியாளர்கள் கண்டறிந்த முறைகளை விட நல்ல விளைவைத் தருகின்றன. எழுத்து வரிசை வளைந்திருந்தாலும் அதனை சரியான எழுத்துக்களுடன் எழுத்துக்களின் வடிவம் மாறாமல் பிரிப்பது இதன் தனிச்சிறப்பாகும். எழுத்து வரி வெட்டுதல் முறையில் பிரிக்கப்பட்ட எழுத்து வரிசைகள் அடுத்த கட்ட செயலாக்கத்திற்கு உதவியாக அமைகிறது (Spurgen Ratheash, & Mohamed Sathik 2019).

படம் 5 எழுத்து வரி வெட்டுதல் முறையில் தனித்தனியாக பிரிக்கப்பட்ட எழுத்து வரிசைகள்

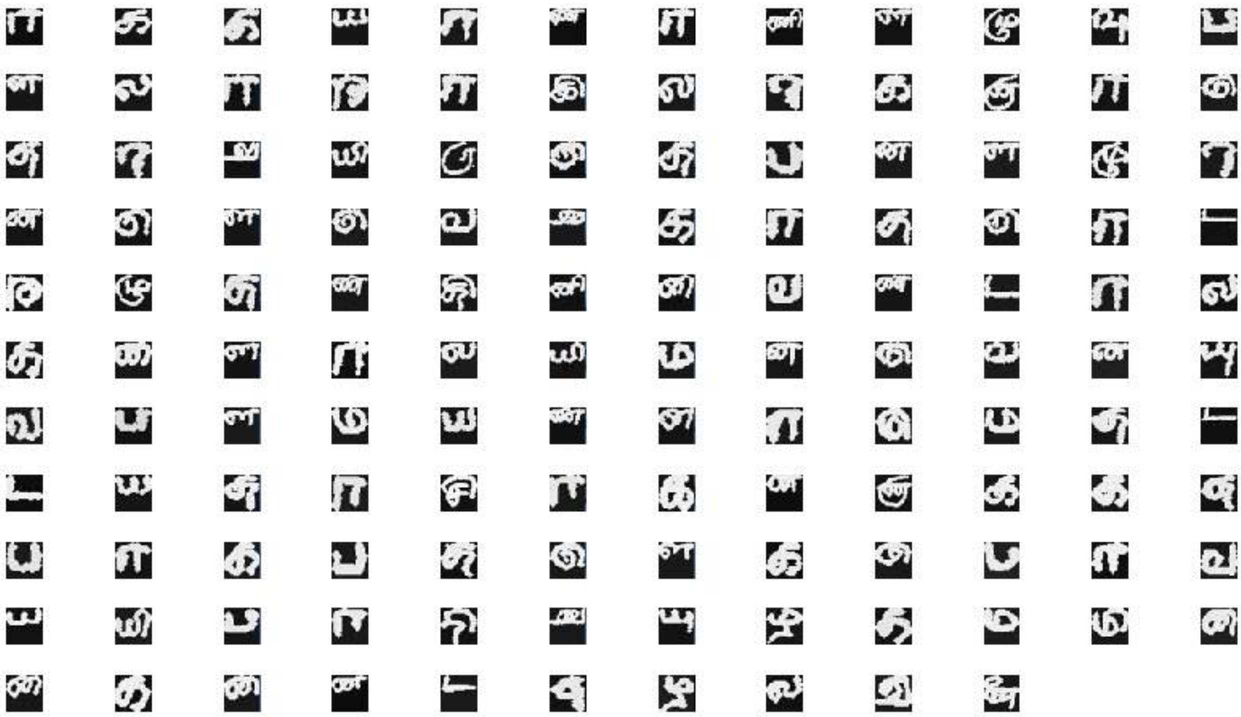
எழுத்துக்களைப் பிரித்தல்

பனை ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்கள் நீண்டு ஒரே எழுத்து வரிசையின் எழுத்துக்களுடன் ஒட்டி காணப்படும். இதனால் எழுத்துக்களை கணினி வழி அடையாளப் படுத்துவதற்கான தனி படங்கள் உருவாக்குதல் சிக்கலான நிலைக்குத் தள்ளப்படுகிறது. இத்தகைய சிக்கல்களை தீர்ப்பதற்காக ஆராய்ச்சியாளர்கள் கிடைமட்ட செங்குத்து வெட்டுதல் (Horizontal Vertical Slicing) வழிமுறையை பனை ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களுக்காக பிரத்தியேகமாக உருவாக்கியுள்ளனர் (Mohamed Sathik, & Spurgen Ratheash, 2020). அதன்படி எழுத்து வரி வெட்டுதல் முறையில் தனியாக பிரிக்கப்பட்ட எழுத்து வரிசையில் உள்ள எழுத்துக்கள் ஒன்றையொன்று ஒட்டிக் கொண்டும், பிண்ணிக் கொண்டும் இருந்தால் அதனை எழுத்துக்களின் வடிவம் மாறாமல் பிரித்து தனி தனி படங்களாக மாற்றப்படுகிறது அவை படம் 6-இல் காட்டப்பட்டுள்ளது.

படம் 6 கிடைமட்ட செங்குத்து வெட்டுதல் முறையில் தனித்தனியாக பிரிக்கப்பட்ட எழுத்துக்கள்

எழுத்துக்களை அடையாளப்படுத்துதல்

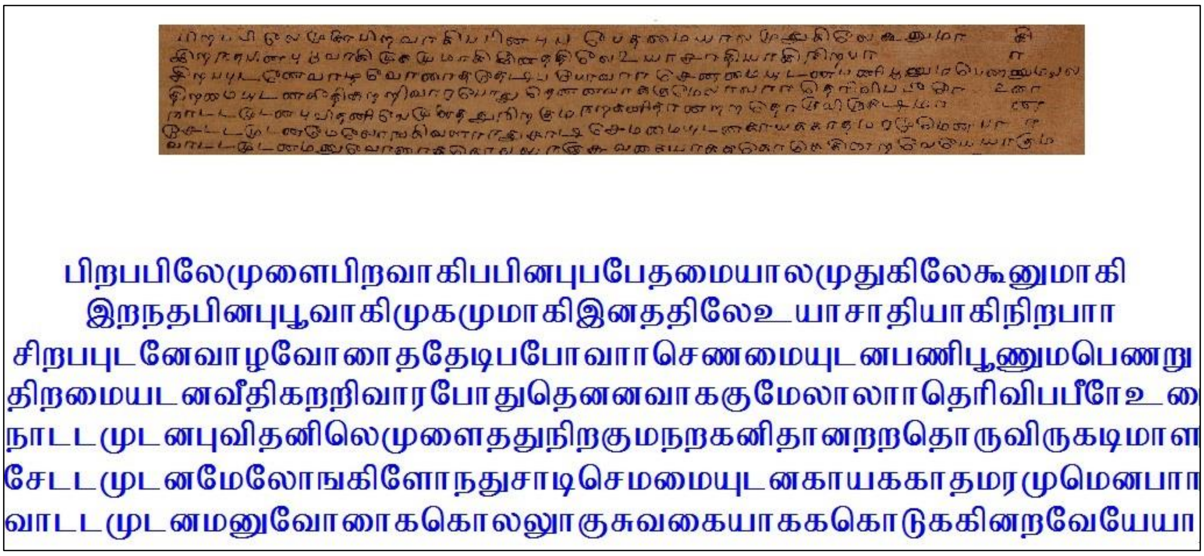
பனை ஓலைச்சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை கணினி வழி அடையாளப்படுத்துதலில் தொடர்ச்சியான நரம்பணுப் பிணைய (Recurrent Neural Networks) நுட்பம் பயன்படுத்தப்படுகிறது. தமிழ் எழுத்துக்களை அடையாளப்படுத்துவதில் பல்வேறு நுட்பங்களை பல்வேறு ஆராய்ச்சியாளர்கள் பயன்படுத்தியிருந்தாலும் பனை ஓலைச் சுவடிகளின் தமிழ் எழுத்துக்களை அடையாளப்படுத்துவதில் பயன்படுத்தப்பட்டுள்ள தொடர்ச்சியான நரம்பணுப் பிணைய நுட்பம் அதிகபட்ச எழுத்துக்களை அடையாளப்படுத்துகிறது. எழுத்து வரிசை வெட்டுதல் முறைப்படி பிரிக்கப்பட்ட எழுத்து வரிசைகள் கிடைமட்ட செங்குத்து வெட்டுதல் முறைப்படி தனி தனி எழுத்து படங்களாக பிரித்த பிறகு தொடர்ச்சியான நரம்பணுப் பிணைய நுட்ப முறை பயன்படுத்தப்படுகிறது. இம்முறையை பயன்படுத்துவதற்கு ஒருசில முன் செயலாக்க முறைகள் பின்பற்றப்படுகிறது. அதன்படி கிடைமட்ட செங்குத்து வெட்டுதல் முறைப்படி பிரிக்கப்பட்ட எழுத்துக்கள் வெவ்வேறு அளவுகளைக் கொண்டுள்ளதாக இருக்குமாயின் முதலில் அவற்றை எல்லாம் இயல்பாக்குதல் (Normalization) முறை மூலம் ஒரே அளவாக மாற்றப்படுகிறது, அவை படம் 7-இல் காட்டப்பட்டுள்ளது.



படம் 7. பிரிக்கப்பட்ட எழுத்துக்கள் ஒரே அளவாக மாற்றப்படுதல்.

அதன்பின் நான்கில் இரண்டு பங்கு படங்கள் பயிற்றுவிப்பதற்கும் (Training) மற்றவை சோதனை (Testing) செய்வதற்கும் பிரிக்கப்படுகிறது. பயிற்றுவித்தல் என்பது எழுத்துக்களின் உருவங்களின் அம்சங்களை (features) கணினியை அதற்கேற்றவாறு புரிந்துகொள்ள வைப்பது. சோதனை என்பது கணினி புரிந்து கொண்ட எழுத்துக்களின் அம்சங்களை ஒத்துப்பார்த்து சரியான எழுத்தை கணிக்கச் செய்வது. இந்த ஆராய்ச்சியில் தொடர்ச்சியான நரம்பணுப் பிணைய நுட்பத்தின் இருதரப்பு நீண்ட கால

நினைவக (Bidirectional Long Short Term (Bi-LSTM)) வகைப்படுத்தி உபயோகப்படுத்தப்படுகிறது. இதில் நினைவகம் (memory cell) இருப்பதால் எழுத்துக்களின் அம்சங்களை நினைவில் வைத்து சோதனை எழுத்துக்களுடன் ஒப்பிடுவது எளிதாக்கப்படுகிறது (Mohamed Sathik, & Spurgen Ratheash 2020; Mohamed Sathik, & Spurgen Ratheash 2020A). இதனால் பல்வேறு காலக்கட்டத்தில் எழுதப்பட்ட தமிழ் எழுத்துக்கள் கணினி வழியாக மிக நேர்த்தியாக அடையாளப்படுத்தப்படுகிறது, அவை படம் 8-இல் காட்டப்பட்டுள்ளது.



படம் 8. கணினியால் அடையாளப்படுத்தப்பட்ட ஓலைச்சுவடியில் உள்ள தமிழ் எழுத்துக்கள்.

முடிவுரை

தமிழ் எழுத்துக்களை கணினி வழியாக அடையாளப்படுத்துவதில் இது வரை பேனாவால் எழுதப்பட்ட தமிழ் எழுத்துக்களே அதிகமாக செயல்முறைப்படுத்தப்பட்டுள்ளது. பனை ஓலைச் சுவடிகளில் உள்ள தமிழ் எழுத்துக்களை கணினி வழியாக அடையாளப்படுத்துவது மிக குறைந்த அளவிலேயே செயல்முறைப்படுத்தப்பட்டுள்ளது. அவ்வாறு செய்திருந்தாலும் மிக குறைந்த எண்ணிக்கையிலான எழுத்துக்களை அடையாளப்படுத்தியுள்ளது. இந்த ஆராய்ச்சியில் பனை ஓலைச் சுவடிகளில் உள்ள அனைத்து தமிழ் எழுத்துக்களையும் கணினிஅடையாளப்படுத்துவது இதன் சிறப்பாக கருதப்படுகிறது. இது 99.5% சரியான எழுத்துக்களை அடையாளப்படுத்துகிறது. இது மற்ற ஆராய்ச்சியை விட அதிக வெற்றி விகிதத்தில் நல்ல விளைவைக் கொடுக்கிறது. தமிழ் எழுத்துக்கள் பொறிக்கப்பட்ட கல்வெட்டுகளில் இந்த முறையை பயன்படுத்தி எழுத்துக்களை அடையாளப்படுத்துவதே இந்த ஆராய்ச்சியின் எதிர் கால திட்டம் ஆகும். தமிழ் எழுதுமுறையில் இரண்டு எழுத்துக்களை ஒரு எழுத்தில் மறைத்து வைத்திருக்கும் சிறப்பம்சத்தை எந்த சிக்கலுமின்றி துல்லியமாக கண்டறிந்து வெளிக்கொணர்வதே இந்த ஆராய்ச்சியின் மிக முக்கிய அம்சம் ஆகும்.

References

Banumathi, P., Nasira, G.M., (2011) Handwritten Tamil Character Recognition Using Artificial Neural Networks, 2011 International Conference on Process Automation, Control and Computing, 1-5. https://doi.org/10.1109/PACC.2011.5978989

Challa, N.P., & Mehta, R.V.K. (2017) Applications of Image Processing Techniques on Palm Leaf Manuscripts-A Survey. *Helix: The Scientific Explorer*, *7*(5), 2013-2017.

Deepa, M., Deepa, R., Meena, R., Nandhini, R., (2019) Tamil handwritten text recognition using convolutional neural networks, International Journal of Engineering Science and Computing, 9(3) 20986-20988.

Kannan R.J., Subramanian S., (2015) An Adaptive Approach of Tamil Character Recognition Using Deep Learning with Big Data-A Survey. In: Satapathy S., Govardhan A., Raju K., Mandal J. (eds) Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) Volume 1. Advances in Intelligent Systems and Computing, vol 337. Springer, Cham. https://doi.org/10.1007/978-3-319-13728-5_63

Kattalai Kailasam, V., (2018) Yedum Eluthum, First Edition, Malli Padhipagam, Tirunelvlei, TamilNadu, India.

Kesiman, M.W.A., Valy, D., Burie, J.-C., Paulus, E., Suryani, M., Hadi, S., Verleysen, M., Chhun, S., Ogier, J.-M., (2018) Benchmarking of Document Image Analysis Tasks for Palm Leaf Manuscripts from Southeast Asia, Journal of Imaging, 4(2) 43. https://doi.org/10.3390/jimaging4020043

Mohamed Sathik, M., Spurgen Ratheash, R., (2020) Optimal Character Segmentation for Touching Characters in Tamil Language Palm Leaf Manuscripts using Horver Method, International Journal of Innovative Technology and Exploring Engineering, 9(6) 1010-1015. https://doi.org/10.35940/ijitee.E3126.049620

Mohamed Sathik, M., Spurgen Ratheash, R., (2020A) Handwritten Tamil Character Recognition in Palm Leaf Manuscripts using BiLSTM Classifier, International Journal of Future Generation Communication and Networking, 13(4) 4440-4450.

Sabeenian R.S., Paramasivam M.E., Anand R., Dinesh P.M., (2019) Palm-Leaf Manuscript Character Recognition and Classification Using Convolutional Neural Networks. In: Peng SL., Dey N., Bundele M. (eds) Computing and Network Sustainability. Lecture Notes in Networks and Systems, vol 75. Springer, Singapore. https://doi.org/10.1007/978-981-13-7150-9_42

Sornam M., Vishnu Priya C., (2018) Deep Convolutional Neural Network for Handwritten Tamil Character Recognition Using Principal Component Analysis. In: Bhattacharyya P., Sastry H., Marriboyina V., Sharma R. (eds) Smart and Innovative Trends in Next Generation Computing Technologies. NGCT 2017. Communications in Computer and Information Science, vol 827. Springer, Singapore. https://doi.org/10.1007/978-981-10-8657-1_61

Spurgen Ratheash, R., Mohamed Sathik, M., (2019) Line Segmentation Challenges in Tamil Language Palm Leaf Manuscripts, International Journal of Innovative Technology and Exploring Engineering, 9(1) 2363-2367. https://doi.org/10.35940/ijitee.L3159.119119

Vellingiriraj, E.K., Balasubramanie, P., (2014) Recognition of Ancient Tamil Handwritten Characters in Historical Documents by Boolean Matrix and BFS Graph, International Journal of Computer Science and Technology, 5(SPL – 1) 65-68.

Funding: NIL
Acknowledgement: NIL
Conflict of Interest: NIL
About the License:

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

AN EFFICIENT APPROACH FOR 2D TO 3D IMAGE CONVERSION USING FUZZY C-MEANS SEGMENTATION

M.SHEIK MANSOOR,
M.Phil Scholar,
Sadakathullah Appa College,
Tirunelveli.
sheik_haya@yahoo.co.in

Dr.M.MOHAMED SATHIK,
Principal,
Sadakathullah Appa College,
Tirunelveli.
mmdsadiq@gmail.com.

Abstract— **The main objective of this paper is to design an efficient 2D-to-3D conversion method based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. To analyse the performance of the proposed method several performance metrics are used. This paper uses PSNR,SSIM, MSE and RMSE to analyses the performance. From the experimental results it is shown that the proposed method performs better than the other existing methods.**

Keywords— ***2D-to-3D conversion, depth boundaries, depthmap,***

nonlocal neighbors, nonlocal edge information.

Introduction

3DTV is widely anticipated as the next revolution of television technology. '3D' (three dimensional) here means 'stereoscopic', which provides viewers with feeling of immersion. The promotion of 3DTV depends not only on technological advances in 3D displays, but also on availability of large amount of 3D content. However, it's both costly and time-consuming to make 3D content of high quality directly with stereoscopic cameras, so the shortage of 3D content becomes one of severe bottlenecks for 3D industry. Converting 2D images or videos to 3D is one way to alleviate the problem in the early stage of 3DTV development, because it not only can create 3D content with lower cost and less time, but also makes full use of large amount of existing 2D materials.

Generally, the existing 2D to 3D conversion approaches can be classified as two categories: human assisted conversion and automatic conversion. The human-assisted approach is to convert 2D images or videos to 3D with some corrections made "manually" by an operator [1]. Even though this approach has relatively better performance, it's still impractical in many scenarios. To convert the vast collection of available 2D material into 3D in an economic manner, an automatic approach is desired [2]. The automatic approach utilizes the depth cues in a single monocular image to generate another or more virtual views without any human assistance.

There is several research works are progressing on 2D to 3D conversion of images which shall be used in the motion pictures [4] and [5]. 3D imaging system has been incorporated in the televisions, cameras etc. In the health system the 3D body scanners help surgeons to determine the accurate status of various diseases. The 3D hardware is expensive compared with 2D hardware system. Therefore, it is necessary to develop a fast and accurate algorithm for converting 2D images to 3D images. In this research article, a new simple algorithm is proposed for converting 2D image to 3D image using image fusion. Xiaoyang Mao, lbsiyasu L. Kunii, Hierarchical was proposed G-octree as an extension of G-quadtree to 3D grey-scale images. They did the program in C on VAX 11/750. Application to the color coding of macro-auto-radiography images of rat brains demonstrated the advantages of the approach [6]. Chin-Tung Lin, Chiun-Li Chin, Kan-Wei Fan, and Chun-Yeon Lin was presented a 2D to 3D effect image conversion architecture integrated image segmentation system and depth estimation. They tested many 640*480 RGB format color images. They generated left view and right view image and displayed the 3D stereo image [7]. H. Murata, X Mori, S. Yamashita, A. Maenaka, S. Okada, K. Oyamada, and S. Kishimoto, proposed a system for converting all kinds of 2D images into 3D images. The method is used adaptively by computing the depth of each separated area of the 2D images with their contrast, sharpness, and chrominance [8]. Wa James Tam and Liang Zhang provided an overview of the fundamental principle of 2D to 3D conversion techniques, short note on approaches for depth extraction using a single image, and depth image based rendering [9]. Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen presented an automatic system for converting 2D videos to 3D videos. They grouped the regions into blocks using the edge information and applied bilateral filter to generate depth map [10]. Zhebin Zhang, Yizhou Wang, Tingting Jiang, and Wen Gao described an approach which estimated a 2.5D depth map by leveraging motion cues and photometric cues in video frames [11]. Ching-Lung Su, Kang-Ning Pang, Tse-Min Chen, Guo-Syuan Wu, Chia-Ling Chiang, Hang-Rnei Wen, Lung-Sheng Huang, Ya-Hsin Hsueh, and Shau-Yin Tseng, presented an algorithm for conversion of 2D to 3D in real time. The 2D video accompanied with a depth image was stored to create 3D video [12]. Yeong-Kang Lai, Yu-Fan Lai, and Ying-Chang Chen proposed a hybrid algorithm for 2D to 3D conversion. They used motion information, linear perspective, and texture characteristic for depth estimation. They used bilateral filter for depth map smoothing and noise removal [13].

Two approaches to 2D to 3D conversion can be loosely defined: quality semiautomatic conversion for cinema and high quality 3DTV, and low-quality automatic conversion for cheap 3DTV, VOD and similar applications. [14] In semiautomatic conversion a skilled operator assigns depth to various parts of an image or video. Based on this sparse depth assignment, a computer algorithm estimates dense depth over the entire image or video sequence. In the case of automatic methods, no operator intervention is needed and a computer algorithm automatically estimates the depth for a single image or video. Automatic methods estimates shape from shading, structure from motion or depth from defocus. Electronics manufacturers use stronger assumptions to develop real-time 2D-to-3D converters. Such methods may work well in specific scenarios. But generally it is very difficult to construct heuristic assumptions that cover all possible background and foreground combinations. An important step in any 3D system is the 3D content generation. Several specialcameras have been designed to generate 3D model directly. For example, a stereoscopic dual-camera makes use of a co-planar configuration of two separate, monoscopiccameras, each capturing one eye's view, and depth information is computed usingbinocular disparity. A depth-range camera is another example. It is a conventional videocamera enhanced with an add-on laser element, which captures a normal two-dimensionalRGB image and a corresponding depth map. A depth map is a 2D function that gives thedepth (with respect to the viewpoint) of an object point as a function of the imagecoordinates. Usually, it is represented as a gray level image with the intensity of eachpixel registering its depth. The laser element emits a light wall towards the real worldscene, which hits the objects in the scene and reflected back. This is subsequentlyregistered and used for the construction of a depth map.All the techniques described above are used to directly generate 3D content, which certainly contribute to the prevalence of 3D-TV. However, the tremendous amount of current and past media data is in 2D format and should be possible to be viewed with a stereoscopic effect. This is where the 2D to 3D conversion method comes to rescue. This method recovers the depth information by analyzing and processing the 2D image structures.

In this paper an efficient 2D-to-3D conversion method is developed based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display.

The remainder of the paper is organized as follows: In Section II, the overview of proposed method is presented. In Section III, the proposed method is specifically depicted, including its design idea and practical implementation approach. In Section IV, the performance of the proposed method is evaluated. Finally, conclusions are made in Section V.

2D to 3D image conversion using edge information

The overall block diagram of this approach is shwon in Fig.1. This work uses an efficient 2D-to-3D conversion method based on the use of edge information. Importantly, the edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Figure 1 schematically depicts the proposed conversion system. Initially, the block-based image is considered to segment it into multiple groups. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. The further details of these modules are discussed below:

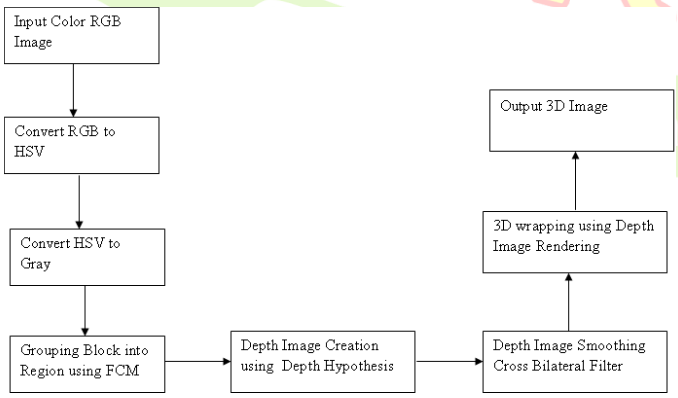


Fig. 1. Overall Block Diagram of Method1

Proposed approach

A. RGB to HSV Conversion

In image processing there are various color models: RGB model, CMY model, HSV model and YIQ model. RGB color model is used for color monitors and color video cameras, CMY model for printers, YIQ model is used in color TV broadcasting. Here RGB color model is converted into HSV color model. RGB (R-Red, G-Green and B-Blue) is a color model appears in primary colors and based on a Cartesian coordinate system. HSV is one of the frequently used color model. In HSV, H stands for hue which specifies the position of pure color on color wheel. Thus hue is related with dominant wavelength in a mixture of light wave. S is saturation, gives measure of the degree to which pure color is diluted with white. V is value called as lightness of color. Sometimes it is represented as I (intensity) or B (brightness). HSV model is having advantages over RGB model. HSV is strong model than RGB because it offers a more intuitive representation of the relationship between colors; it selects more specific color. RGB is costly in terms of computation time. The way in which human beings perceive color; hue and saturation components from HSV relates same way.

B. HSV to Gray Conversion

Convert HSV color image to grayscale image by averaging the H, S and V color space pixels.

C. Block-Based Region Grouping

Computational complexity is reduced mainly by block-based algorithm. This implies each pixel in the same block has the same depth value. A 4-by-4 block is used as an example. Each node is a 4-by-4 pixel block, and is four-connected. The value of each link is calculated by considering the absolute difference of the mean of neighboring blocks. A smaller value obtained implies a higher similarity between the two blocks. Following calculation of the absolute difference of the mean of the neighboring blocks, the blocks are then segmented into multiple groups by applying the fuzzy c-means (FCM) segmentation.

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

D. Depth from Prior Hypothesis

The extraction of depth is the crucial one in the conversion process. The greatest difference between 2D and 3D image is the depth information. The object can jump out of the screen and look like a real life due to the depth information. If we extract these depth signals and integrate them together, we will build a strong foundation to make 3D images of better and higher quality. The depth generation algorithms are roughly classified into three categories which utilize different kinds of depth cues: the binocular, monocular and pictorial depth cues. Each signal represents different depth information. In this conversion process, following the generation of the block groups, the corresponding depth for each block is then assigned by the depth gradient hypothesis. The process includes the generation of gradient planes, depth gradient assignment, consistency verification of the detected region, and finally the depth map generation. When each scene change is detected, the linear perspective of the scene can be analyzed with the help of line detection algorithm using Hough transform.

E. Bilateral Filtering

The bilateral filter is non-iterative and also achieves satisfying results with only a single pass. This makes the filter"s parameters relatively intuitive as their effects are not cumulated over several iterations. The bilateral filter has proven to be much useful although it is slow. It is nonlinear and also its evaluation is computationally expensive because the traditional accelerations like performing convolution after an FFT, are not applicable. Nonetheless, solutions have been proposed later in order to speed up the evaluation of the bilateral filter. Unfortunately, these methods seem to rely on approximations that are not grounded on firm theoretical foundations. Among the variants of the bilateral filter, this conversion method has selected the cross bilateral filtering. In some applications like computational photography, it is often useful to decouple the data to be smoothed to define the edges to be preserved. The chosen cross bilateral filter is a variant of the classical bilateral filter. This filter is used to smoothen the image to locate the edges to preserve. The depth map generated by block-based region grouping contains blocky artifacts.

F. Depth Image Based Rendering

The filtered depth map has a comfortable visual quality because the cross bilateral filter generates a smooth depth map inside the smooth region with similar pixel values and preserves sharp depth discontinuity on the object boundary. Following filtering by the cross bilateral filter, the depth map is then used for the generation of the left/right or multi-view images using depth image-based rendering (DIBR) for 3D visualization. DIBR for advanced 3D TV System can be illustrated by the following block diagram. This system includes three parts, pre-processing of depth map, 3D image Warping and Hole- Filling. Smoothing filter is first stage applied to smooth the depth image. Then the 3D image warping generates the left and right view according to the smoothed depth map and also intermediate view. If there are still holes in the image, hole-filling is then applied to fill color into these holes.

G. 3D Image Warping

3D image warping maps the intermediate view pixel by pixel to left or right view according to the pixel depth value. In the other words, 3D image warping transforms the location of pixels according to depth value.

1. Expiremental Images

Experiments were conducted on a group of color images to verify the effectiveness of the proposed scheme. For the experimental purpose several standard, 512 × 512 cover images are taken. Some of these images , i.e., Lena, Barbara, Babbon, Peppers, Sailboat, and Tiffany, are shown in Figure 2.

(a) (b)

Fig. 2. Expiremental Images (a) Building (b) Boat

B. Performance Analysis

To evaluate the performance of the steganography techniques several performance metrics are available. This paper uses the PSNR,SSIM, MSE and RMSE to analyses the performance.

1. Peak Signal-to-Noise-Ratio

The peak signal-to-noise ratio (PSNR) is used to evaluate the quality between the 3D image and the original 2D image. The PSNR formula is defined as follows:

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

PSNR

$$= 10 \times \log 10 \frac{255 \times 255}{\frac{1}{H \times W} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} [f(x,y) - g(x,y)]^2} \text{dB}$$

where H and W are the height and width of the image, respectively; and f(x,y) and g(x,y) are the grey levels located at coordinate (x,y) of the original image and attacked image, respectively.

2. Structural Similarity Index

The structural similarity index is a method for measuring the similarity between the 3D image and the original 2D image.

$$SSIM\ (y, \hat{y}) = \frac{\left(2_{\mu_y \mu_{\hat{y}}} + c_1\right)(2\sigma_{y\hat{y}} + c_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + c_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + c_2)}$$

where, $\hat{Y}$ is the 3D image, the Y is the original 2D image, μ is the mean and the is the variance.

3. Mean Square Error

The mean square error (MSE) is used to evaluate the difference between a 3D image and the original 2D image. The MSE can be calculated by,

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$

where, $\hat{Y}$ is the 3D image and the Y is the original 2D image.

4. Root Mean Square Error

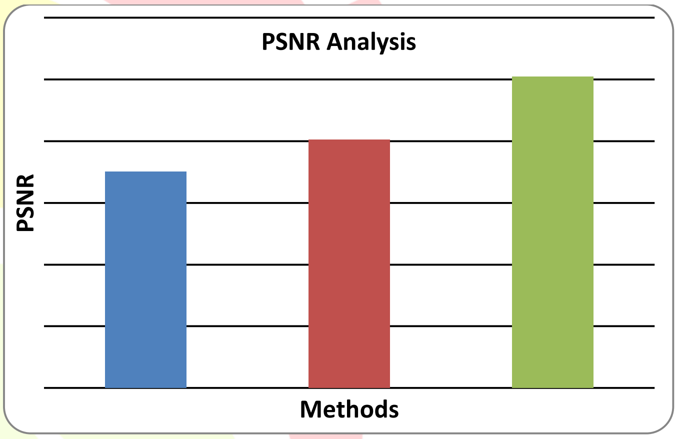
The Root Mean Square Error (RMSE) is a frequently used measure of the difference between 3D image values and the original 2D image values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2}{n}}$$

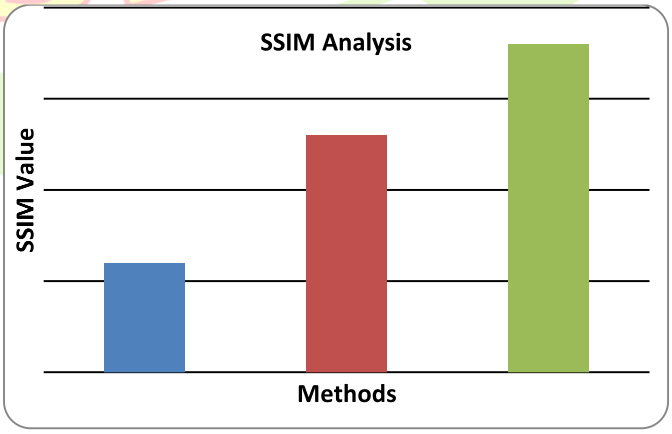
where $\hat{Y}$ is 3D image and Y is original 2D image.

To analysis the performance of the three methods by using the performance metrics which are mentioned above. This is shown in the below tables and graphs

Methods	PSNR
Depth with RGB	17.5
Depth with Fusion	20.1
Proposed Method	25.2

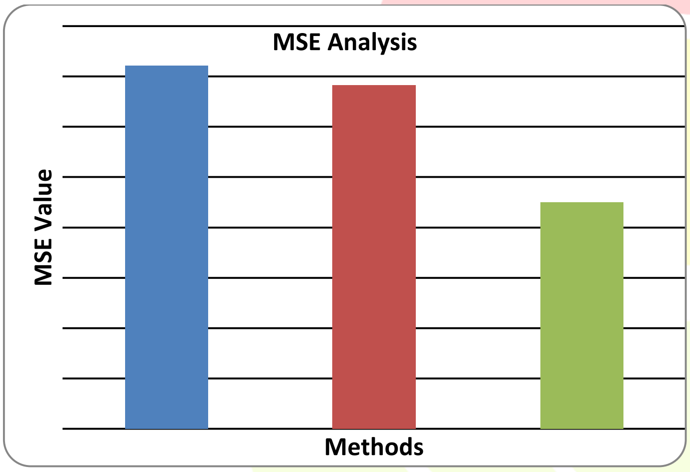


Methods	SSIM
Depth with RGB	0.61
Depth with Fusion	0.68
Proposed Method	0.73

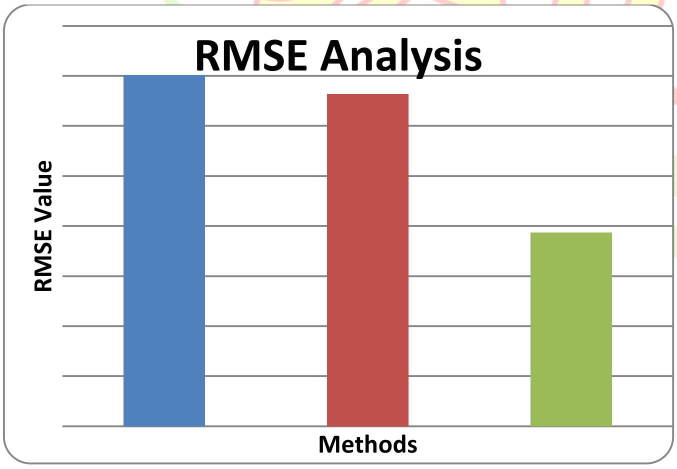


ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

Methods	MSE
Depth with RGB	7.21
Depth with Fusion	6.82
Proposed Method	4.50



Methods	RMSE
Depth with RGB	3.50651252
Depth with Fusion	3.32036033
Proposed Method	1.9362638



Conclusion:

This paper developed an efficient 2D-to-3D conversion method based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. To analyse the performance of these method several performance metrics are used. This paper uses PSNR,SSIM, MSE and RMSE to analyses the performance. From the experimental results it is shown that the proposed method performs better than the other two methods..

References

[1] Wikipedia the Free Encyclopedia: Stereoscopy [Online]. Available: http://en.wikipedia.org/wiki/Stereoscopy
[2] Getting Started in Stereo Photography. [Online]. Available: http://home.comcast.net/~dssweb/intro_to_stereo.htm
[3] Wikipedia the Free Encyclopedia: 3D film. [Online]. Available: http://en.wikipedia.org/wiki/3D_film
[4] J. Caviedes and J. Villegas, "Real time 2D to 3D conversion: Technical and visual quality requirements," *IEEE International Conference on Consumer Electronics*, pp. 897-898, 2011.
[5] T. N. Lu and S. Z. Tan, "An algorithm based on bilinear interpolation of converting 2D video to 3D video," *IEEE*, pp. 4813-4815, 2011.
[6] X. Y. Mao and L. K. lbsiyasu, "Hierarchical representations of 2D/3D Gray-Scale Images and their 2D/3D two way conversion," *IEEE*, pp. 37-44, 1987.
[7] T. L. Chin, C. L. Chin, K. W. Fan, and C.Y. Lin, "A novel architecture for converting single 2D image into 3D effect image," *IEEE*, pp. 52-55.
[8] H. Murata, X Mori, S. Yamashita, A. Maenaka, S. Okada, K. Oyamada, and S. Kishimoto, "A real-time 2-D to 3-D image conversion technique using computed

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

image depth," *SID Symposium Digest of Technical Papers*, vol. 29, no. 1, pp. 919-923, 1998.
[9] W. J. Tam and L. Zhang, "3D-TV content generation: 2D-TO-3D Conversion," *ICME*, pp.1869-1872, 2006.
[10] C. C. Cheng, C. T. Li, and L. G. Chen, "A 2D-to-3D conversion system using edge information," in *Proc. Digest of Technical Papers International Conference on Consumer Electronics*, 2010, pp. 377-378.
[11] Z. B. Zhang, Y. Z. Wang, T. T. Jiang, and G. Wen, "Visual pertinent 2D-TO-3D video conversion by multi-cue fusion," in *Proc. 18th IEEE International Conference on Image Processing*, 2011, pp. 909-912.
[12] C. L. Su, K. N. Pang, T. M. Chen, G. S. Wu, *et al.*, "A real-time Full-HD 2D-to-3D conversion system using multicore technology," in *Proc. fifth FTRA International Conference on Multimedia and Ubiquitous Engineering, IEEE*, 2011, pp. 273-276.
[13] Y. K. Lai, Y. F. Lai, and Y. C. Chen, "An effective hybrid depth-generation algorithm for 2D-to-3D conversion in 3D displays," *Journal of Display Technology*, vol. 9 no. 3, pp. 154-161, March 2013.

[14] Jean Maria Dominic," Recent Trends in 2D to 3D Conversion: A Survey", MG University, IJRASET Vol.2 IV, April 2014.
[15] GUAN, S., KLETTE, R. Belief Propagation on edge image for stereo analysis of image sequences. In Proceedings Robot Vision. LNCS 4931, 2006,
[16] BENCO, M., HUDEC, R. The advances image segmentation techniques for broadly useful retrieval in large image database. In NSSS IX. TatranskeZruby (Slovakia), 2006,
[177] HE, R., ZHU, Y. A hybrid image segmentation approach based on Mean Shift and fuzzy C – Means. In Asia – Pacific Conference on Information Processing., 2009.
[18] WANG, G., JU, H. A disparity map extraction algorithm for lunar rover BH2. In IEEE International Conference on Intelligent Computing and Intelligent Systems ICIS 2009. Shanghai, 2009, vol. 4.
[19] McKINNON, B., BALTES, J. Practical region – based matching for stereo vision. In IWCIA. 2005, vol. 3322 of Lecture Notes in Computer Science, Springer,
[20] KUHL, A. Comparison of stereo matching algorithms for mobile robots. Centre for Intelligent Information Processing System. 2005, University of Western Australia,

Early Exposure of Lung Cancer by Combining ANN and SVM Algorithms

M. Sheik Mansoor, M. Mohamed Sathik

Abstract*: Lung cancer is a lethal type of cancers as its rate of spreading is very high compared to the other cancers. Patient who have been affected from Small Cell Lung Cancer (SCLC) has fast outspread rate. Even at initial stage, around 67-75% of cancer victims with SCLC will have fast outspreads and serious damages to the nearby physical parts. Moreover, World Health Organization (WHO) has predicted the count of lung cancer deaths will reach 9.6 million in 2020. Identifying such a lethal type of cancer early can be lifesaving one. Because, cancer cells in lungs are capable of traveling to other body parts even before the doctor detects them in lungs. In this research work, we have designed a combined approach to prognosticate lung cancer and its type using Artificial Neural Networks (ANN) and Support Vector Machine (SVM). To train both the ML algorithm, an open access patient health dataset published by cancer imaging archives is used. The dataset has the information like pretreatment CT scans, 3D image details of tumor and clinical outcomes. The results produced by ANN and SVM algorithm are compared to predict the type of the lung cancer accurately. The result holds good for a real time implementation.

Keywords: Neural Networks, Lung Cancer Prediction, Cancer diagnosis, Support Vector Machine.

I. INTRODUCTION

Lung cancer is a dangerous kind of cancer which originates from the lungs and outspreads to other nearest physical parts in a short time. Lung cancer is the second most cancer which affects men and fifth most cancer which affects women. As stated by Global Cancer Observatory (GCO), every 5.4 person has lung cancer among one million peoples in India. The alarming issue in the raise of lung cancer is, it has very low survival rate compare to any other cancer diseases. In India, 25% of cancer victims loses their life every year. Due to late stage diagnosis and fast outspread, deaths rate of lung cancer is too high compared to other prostate, colorectal, skin, kidney and breast cancers.

In general, lung cancers can be majorly categorized into Non-Small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC). Here, NSCL cancer is further categorized into three major types, such as adenocarcinoma, squamous and large cell carcinoma. The second type that is SCLC is the dangerous cancer, in which cancer cells spreads to different body parts in short time period through lymphatic nodes. The recent report from National Institute of Cancer Prevention and Research (NICPR) says, around 85-90% are of NSCL cancer and 10-15% of cancer are of SCLC.

Identifying lung cancer cells and its forceful exertion in primitive stage is the only possible way to meliorate the patient's chance of survival. But, in the primitive stage, performing traditional way of histopathology using microscopes will not produce the clear cut results.

Sometimes, lesions in small size can't be effectively detected by CT scan device. To overcome these issues and to identify the cancer type in early stage, automatic ML techniques are used on the patient data. It helps the physicians to acquire a clear cut knowledge about condition of the patients. Moreover, it also helps physicians in identifying the type and vigorous of the cancer cells.

In this research work, we have designed a combined approach to prognosticate the lung cancer and its type using Artificial Neural Networks (ANN) and Support Vector Machine (SVM).

ANN maintains an interconnected nodes, called as neurons to gather information by identifying relationships and new pattern between the data. It has three layer such as, input neuron layer, hidden neuron layer and output neuron layer. Neurons in each layers will receive the input data, performs operations and forwards the data to the nearby connected neurons. Each neurons and the edge which connects the neurons has a particular weight. The weight will change on the neurons based on the learnings. ANN allows both forward and backward propagation for learning.

The final result of ANN are produced based on the maximum probability of neurons present in output layer. Even there exist several algorithms to predict the early state lung cancer, using ANN will produce an accurate result as it allows both forward and backward propagation of learning.

SVM is a discriminative classifier, which draws a hyper plane to differentiate the classes that are derived as outputs. The hyper planes are the decision boundaries. The maximum accuracy can be attained only if the SVM draws the hyper plane separating all the objects to its classes correctly.

This paper is organized in a manner such that, Section 2, describes the related works that are performed in predict lung cancer. Section 3, explains the proposed combined ANN and SVM lung cancer prediction approach. Section 4, evaluates the performance of the proposed approach and Section 5 concludes and discusses about the future work of the proposed combined approach.

Revised Manuscript Received on November 22, 2019

M Sheik Mansoor, Research Scholar (Reg. No. 17221192161007), Sadakathullah Appa College, Tirunelveli, affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India. Email: sheikmansoormsm@gmail.com

Dr. M Mohamed Sathik, Principal and Research Supervisor, Sadakathullah Appa College, Tirunelveli, affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India Email: mmdsadiq@gmail.com

II. RELATED WORKS

Machine learning algorithms are the subgroup of Artificial Intelligence (AI), which learns from data samples and generate insights to identify the pattern in given dataset. Every machine learning has two major learning process such as, identifying the dependencies between data and use the identified dependencies to prefigure the output result.

ML algorithms are been used in number of biomedical applications. The important application of ML in medical field is cancer prediction. Right after the evolvement of Personal Health Information (PHI) and increased computation power, ML algorithms are implemented to identify the cancer type. This section shows an elaborate study of related works that are carried out on the ML to predict cancers, especially on lungs.

The popular ML algorithms used in predicting cancers are Support Vector Machine (SVM) Bayesian Networks (BN) and Neural networks (NN).

A. Cancer prediction based on SVM

SVM is a supervised learning technique which performs classification and regression to identify the associations between the data in the given dataset. Kancherla [1], proposed a cancer prediction method using Support Vector Machine (SVM) based Recursive Feature Elimination (RFE) method. Here, image processing algorithms are executed on the dataset to draw out the basic features that differentiates the cancer cells from the normal cells. After extracting the basic attributes from the dataset, Tetrakis Carboxy Phenyl Porphine (TCPP) is added to the image file to identify the defected cells. For differentiation, SVM is executed. This SVM-RFE method attains accuracy up to 87.5%.

Trivedi [3], proposed a hybrid ML algorithm to identify the cancer defected cells in the lungs using SVM. This particular research combines three major categories of SVMs, such as, linear SVM, Radical Basic Function (RBF) SVM and polynomial SVM to categorize the cancerous and non-cancerous cells. The final optimized results are derived from all the three SVMs. Combination of three SVM are mentioned as hybrid classification. Hybrid SVM attains accuracy upto 77.27%.

Nadira [4], proposed a Global Artificial Bee Colony (GABC) SVM based technique for classifying the cancer type. GABC - SVM predicts important feature from dataset, improves run-time of the ML algorithm. The accuracy of the GABC based SVC classifiers gives accuracy rate as 96.42% (calculated from confusion matrix). The GABC SVM system uses k-folded cross validation method to ensure the maximum rate of testing and training.

B. Lung cancer prediction based on BN

Bayesian networks is probabilistic model which is formed by creating a Direct-Acyclic Graph (DAG). In here, nodes represent the variable and edge represent the dependency between the variables. Depending upon the relations between the nodes, results are derived. BN models can effectively handle uncertainty and missing values compared to the SVM.

Arora [5], presented a detailed survey on knowledge representation and ML tool for risk estimation using Bayesian Network (BN) models. Later, Sesen [6], proposed an idea by implementing Bayesian Networks for predicting the survival days and selecting the perfect treatment plan to cancer patients. The treatments are selected based on different treatment plan option on this estimated survival. It also uses a CAMML hybrid casual discovery algorithm to improve Receiver Operation Curve (ROC) rate while classification. Though the system has several advantages like data encoding, handling large patient data. The proposed method does not provide a significant improvement in prediction models with respect to SVM.

Jayasurya [7] proposed a framework to foresee the survival of the patients who are affected from the non-small cell lung cancer (NSCLC) through radiotherapy. BN based prediction method achieves high predictive rate because it combines both the physical and biological factor for predicating the lung cancer. It also uses SVM to improve classification accuracy.

C. Lung cancer prediction based on ANN

Applying Neural Network based algorithm in medical applications is in practice for a long time. Lakshmanakumar [2], proposed a framework to analyze the CT images using the Optimal Deep Neural Network (ODNN) and Linear Discriminate Analysis (LDA). Here, ODNN extracts the attributes of CT scanned image and LDR is reduces the dimensional of the attribute and classify the tissues as malignant or benign. It gives 94.5% accuracy.

Mandal [8], proposed an ANN based technique to categorize breast cancer data and lung adenocarcinoma. It uses Multilayer Feed Forward (MLFF) neural network to identify cancer cells. The result produced by the ANN method give a high rate of accuracy in classifying the cancer type.

Adetiba [9] presented a comparison between SVM and ANN ensembles and non-ensemble variants of cancer prediction. For extracting the important information from the dataset, two techniques were used. They are Histogram of Oriented Gradient (HOG) and Local Binary Pattern (LBP). Because of combining the ANN and HOG genomic feature, cancer cells are detected in the primitive stage itself. The accuracy rate of this machine learning system is 95.90%. But accuracy reduces when the missing value increases.

III. PROPOSED FRAMEWORK

The proposed approach intends to create a design for prognosticate lung cancer at early stage based on Artificial Neural Network and Support Vector Machine. Both SVM and ANN machine leaning algorithm has different structure to learn from the dataset. ANN uses layers of nodes or neurons and SVM creates an ‘n’ dimensional (n classes) hyper plane from the given data. In general, both algorithms are promising good on predicting the final results.

In this research work, both ANN and SVM are trained differently using the same datasets and the final results are compared to predict lung cancer type. The workflow structure of the proposed combined approach is represented in Fig.1.

To train ANN and SVM, we have used, an open-source lung cancer dataset, which is downloaded from Cancer Imaging Achieve (CIA). The chosen dataset consist of CT scan data of 1019 patients with different cancers. The CT images in this dataset has captured 3D volume of gross tumor.

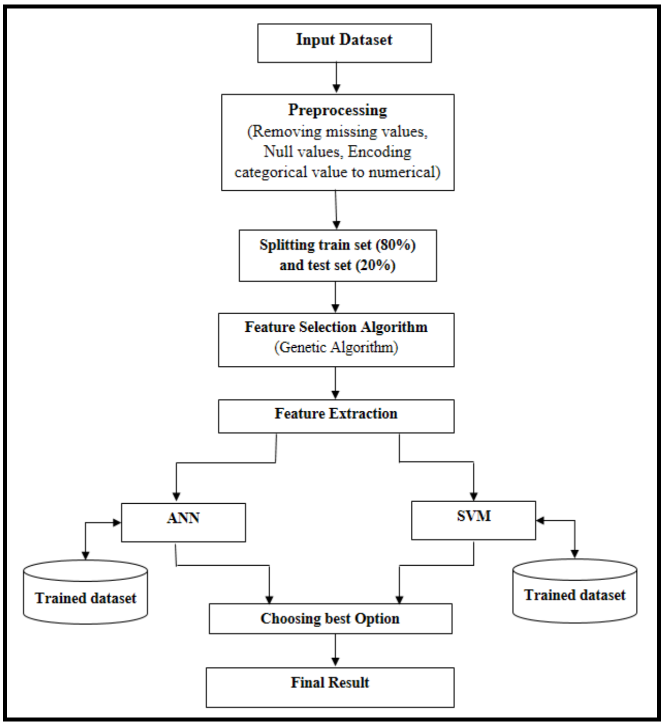


Fig. 1.Workflow of the proposed combined approach

To train machine learning system using Patient Health Information (PHI), raw dataset cannot be given as an input directly. Raw dataset will have missing values, null values, impurities, and irrelevant data in it. To reduce these kind of unwanted data from the dataset, preprocessing is carried out.

A. Data Preprocessing

Pre-processing is the first step in creating a ML algorithm. It transforms raw data into an intelligible format. In general, real world data are incomplete, inconsistent and it will contain several errors. To decrease the errors, preprocessing is carried out on the given dataset. Here, numpy package is used to preprocess. It removes the entire row of the dataset, if it has a void values of a particular feature and column is removed, if it has more than 80% of null values.

To identify the missing or deviated data from a dataset, mean and median for each column. This value is set to be a threshold and missing or deviated values are found from the dataset. After removing the missing values and null values from the dataset, categorical values such as patient-id, histology and stage of caner are encoded into numerical values using pandapy and label_encoder packages.

B. Splitting the dataset

After the data is preprocessed, the filtered dataset is categorized into training dataset and test dataset. Training dataset are given as input to ML algorithm to train and test dataset verifies the results. In the chosen dataset, after preprocessing 398 patient's data (only with NSCLC) are filtered among 1019. In which, 80% of data is taken as training dataset (i.e. 320 patient data). These training dataset are given as an input to ANN and SVM simultaneously. The remaining 78 patient data is chosen as test set.

C. Feature selection algorithm

Feature extraction is performed to decrease the amount unwanted dimensions (i.e. classes) from the pre-processed data. It further reduces the data to a manageable groups of elements. The dimensions are deduced because using the raw dataset directly might complicate the entire learning process. Moreover, it also requires a large computing resources. Feature extraction will create a complete and accurate data groups from the preprocessed data. In this proposed method, Genetic Feature Selection Algorithm (GFSA) is executed to reduce the unwanted dimension from the dataset. Moreover it also reduces the redundant values from the dataset.

GFSA are obtain from the Darwin's idea of natural selection procedure. It only selects the fittest or flawless data from the given dataset. Genetic algorithm has four techniques to scrutinize best-fit values from the dataset. They are selection, inheritance, mutation, and crossover.

Initially, from the given dataset, a couple of candidate-solutions are randomly created. The candidate represent the total population and solution represents individual value. For each solution, the fitness function φ(x) is calculated.

The best fitted individuals are combined together to develop offspring, which later creates another population from the dataset. Later these produced values are given as an input to mutation to find the next best fit values. This procedure is executed continuously to create many generation of values.

In the proposed system, the cross validation of offspring is set to 500 for chosen dataset. So the iteration on finding the fittest values is performed 500times.

D. Artificial Neural Networks

ANN is the main tool used in ML technique for prediction. The dataset inferred from the feature selection module is fed to the input layers of ANN. The inputs are forwarded to hidden layer to process. Hidden layer has several small neuron or units which computes the weight of every neurons and transform the data to the output layer. Hidden layer helps in identifying the complex patterns. The important feature of ANN is backward propagation. It allows the networks to conform the weights of neurons in the hidden layer to improve accuracy of the prediction result. Whereas, ML algorithms like SVM, BN doesn't have the capability of back propagation. Another important feature in ANN is multilayer hidden neurons will extract information from different features. The basic structure of ANN network is represented in Fig.2.

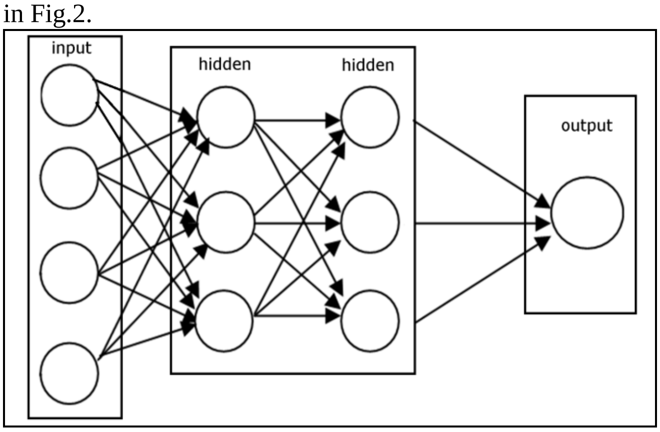


Fig. 2. Stucture of Neural Networks

Every layer in the ANN will have several neurons or nodes in it. Each nodes in the neural network takes numerical values 'X_n' as input. Moreover, each input values are associated with weights 'W_n'.

Net input is calculated using below formula,

$$Net\ input = X_1W_1 + X_2W_2 + \ldots..\ X_nW_n \tag{1}$$

As our scrutinized dataset size for training is 320, we have created 320 input layer neurons to process the data. Each input 'X_n' is fed with the associated weight to the nodes in input layers. The weight for each data input is assigned in a random manner using Pseudo Random Number Generator (PRNG).

While performing forward propagation (linear), the results are equated with the real time values to find the deviations between derived output and exact output. The linear forward propagation value 'Y_n' are obtained from the below equation,

$$Y_1 = X_1W_1 + b \tag{2}$$

Where '*b*' is the bias value (zero for first input).

Now the linear forward value 'Y_1' is given to the first activation function as input.

Activation functions or transferring function are present in hidden layers. It is an important function in deriving the final result. It helps in converting the input signal into output signals. Activation function converts the linear input values into a non-linear inputs, which makes the NN to process complex task in short time.

Here, *tan h* activation functions is applied on the hidden layer neurons. Since, the chosen dataset is of multiclass type, two hidden layers are introduced. The *tan h* function is shown below,

$$tan\ h = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{3}$$

Later, the outputs obtained from the first hidden layer are given to next linear forward propagation Y_2.

After the forward propagation is defined, the backward propagation mechanism has to be created. In neural network, backward propagation computes the gradient of loss function concerning the weights of the network to reduce errors in each neuron.

$$dz_3 = output\ layer - y \tag{4}$$

Loss function concerning weight and bias (for backward propagation – layer 3 to layer 2) is calculated from the below mentioned formulas,

$$\text{Derivative weight } dW_3 = (1/m).\ A_2^T\ .\ dz_3 \tag{5}$$

$$\text{Derivative bias } db_3 = \sum dz_3 \tag{6}$$

The proposed neural network, performs propagations several time to train neural networks. The final prediction is made on the basis of weights and the bias values in output layer.

E. Support Vector Machine

SVM and ANN algorithm are performed in a parallel manner. SVM classifies the given dataset into two or more classes and creates hyper plane between them. In the perspective of dimension space, SVM will draw a line to differentiate the classes based on the features of the objects.

In training phase, training dataset are fed to linear SVM. In the chosen lung cancer dataset, linear SVM creates three classes based on the attributes of the dataset. SVM derived classes are T stage, M stage and N stage. These classes are separated using three hyper lanes as like showed in Fig. 3.

In here, green represent the stage T, cancer blue represents the stage M cancer and red represents the stage N cancer. After the SVM machine learning system is trained, test dataset are fed to the system to check whether the system produces the correct results. Once after the ANN and SVM machine are ready, the final results are produced to the physicians to decide the cancer type.

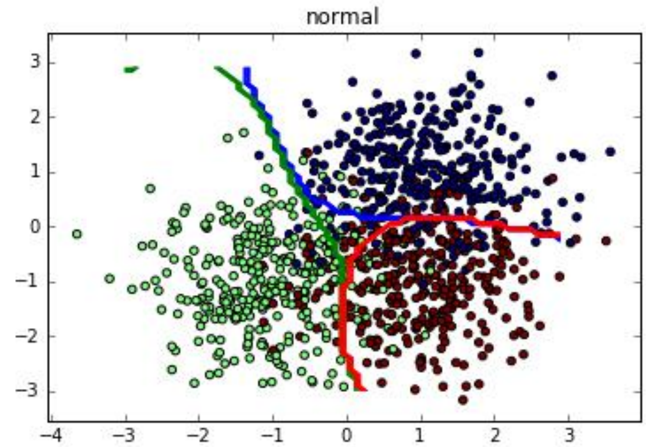


Fig. 3. Three different classes derived from SVM

The final results are decided based on accuracy, sensitiviy and precision. The following section discusses about the results generated by the ANN and SVM.

IV. RESULT AND DISCUSSION

The proposed ANN and SVM machine learning experiments are carried out on the Tensor Flow software, which is a free open-source software developed by Google Inc., The dataset used for the implementation is taken from Cancer Imaging Archives. The chosen dataset consist of CT scan data of 1019 patients with different cancers. Initially, information about the patients, who has affected by the NSCLC cancer is taken out from the given dataset. Around 419 patient records are extracted. Later these, NSCLC cancer data is separated into training dataset and test dataset with the ratio of 70:30. The training dataset are fed as an input to ANN and SVM, simultaneously. They are trained and computed simultaneously for best prediction results. Accuracy is an important performance metric of any classification model. Accuracy is the fraction of predictions that the ML algorithm has predicted accurately. It depends on four possibility of the predicted results, such as, true positive, true negative, false positive, false negative.

The accuracy can be found using the below formula,

$$\text{Accuracy} = \frac{\text{Number of prediction that are correct}}{\text{Total number of prediction made}} \quad (7)$$

Table. 1. Prediction of cancer type using ANN method

Predicted	True/Actual			
		Type 'T'	Type 'M'	Type 'N'
	Cancer Type 'T'	96	8	4
	Cancer Type 'M'	5	89	5
	Cancer Type 'N'	4	5	104

Table. 2. Prediction of cancer type using linear SVM

Predicted	True/Actual			
		Type 'T'	Type 'M'	Type 'N'
	Cancer Type 'T'	101	6	2
	Cancer Type 'M'	7	95	8
	Cancer Type 'N'	8	5	88

Table 1 and Table 2, represent the prediction made through ANN and SVM respectively. Accuracy of ANN model is 90.2%. In 320 total predicted value, ANN has correctly predicted 290 values. However, accuracy of SVM algorithm is 88%, where 284 predictions are made correctly.

The second important performance metric of ML algorithm is precision. It is the fraction of relevant information retrieved (i.e.) in lung cancer type prediction, what fraction of patients belong to a particular cancer type.

In predicting lung cancer type, precision value of type 'x' cancer is found by the fraction of correctly predicted type 'x' cancer from the total prediction. Precision is calculated by the formula specified below,

$$\text{Precision (Type 'x')} = \frac{\text{No. of correctly predicted Type 'x'}}{\text{Total predicted Type 'x' cancer}} \quad (8)$$

So, the precision for type 'T' cancer is 96/108 = 88% for ANN machine learning system and 92% for linear SVM classification. Likewise the precision is calculated for type M and type N cancers and shown in Table 3.

Table. 3. Precision values of ANN and SVM algorithm for given dataset.

Precision values (in percentage)		
	ANN	**SVM**
Cancer Type 'T'	88.8%	92.6%
Cancer Type 'M'	90.8%	86.3%
Cancer Type 'N'	92.6%	87.1%

The third most important metrics of ML algorithm is recall. It is the fraction of relevant information that are retrieved.

$$\text{Recall} = \frac{\text{No. of correctly predicted type 'x' cancer}}{\text{No. of actual type x cancer patients}} \quad (9)$$

The recall value of the ANN and SVM for three classes (types of cancer) are shown in table 4.

Table. 4. Recall value of ANN and SVM for given dataset

Recall values (in percentage)		
	ANN	**SVM**
Cancer Type 'T'	91.4%	87%
Cancer Type 'M'	86.4%	89.2%
Cancer Type 'N'	91.2%	89.7%

After accuracy, precision and recall values are derived from ANN and SVM, physicians can manually go through the results and can find out the exact scenario of the patients. Instead of relying on single ML algorithm, physicians can use both ANN and SVM. It will give an accurate and clear cut information regarding the patient health condition.

V. CONCLUSION

Early exposure of lung cancer and its type is a difficult and complicated process in the medical-machine learning field. Because, the test information gather from patient might not show the exact information of the patients. Some small sized cancer cells are not visible even the images captured through CT scan device. Predicting lung cancer type using a single ML algorithm might provide an erroneous result due to its possibility of false positive errors. To overcome these issues, the proposed design uses both Artificial Neural Networks and Support Vector Machines on the same dataset. For extracting the important information from raw dataset, a genetic (iterative) based method is used. The predictions derived from ANN and SVM will help the physicians to give a better treatment plan to the patients.

APPENDIX

The dataset is downloaded from Cancer Imaging Archives. It depicts a detailed information of 1019 patent records, in which 419 patients are affected by NSCL lung cancer. Information like patient id, age, gender, type of cancer, tissue size, stage, survival days, death status are provided.

REFERENCES

1. K. Kancherla and S. Mukkamala," Feature Selection for Lung Cancer Detection Using SVM Based Recursive Feature Elimination Method", Lecture Notes in Computer Science, Vol. 7256, pp. 168-176,2012.
2. S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N Arunkumar, and G. Ramirez, "Future Generation Computer System, Vol. 92, pp.374-382, 2018
3. A. Trivedi and P. Shukla, "Lung Cancer Diagnosis by Hybrid Support Vector Machine", Communications in Computer and Information Science, Vol. 628, pp. 177-187, 2016.
4. T. Nadira and Z. Rustama, "Classification of Cancer Data Using Support Vector Machines with Features Selection Method Based on Global Artificial Bee Colony", Vol. 2023(1), pp. 1-7, 2018.
5. Arora, P., Boyne, D., Slater, J. J., Gupta, A., Brenner, D. R., and Druzdzel, "Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine, Value in Health",Vol. 22, pp.437-445, 2019.
6. M. B. Sesen, T. Kadir, R. B. Alcantara, J. Fox, and M. Brady," Survival Prediction and Treatment Recommendation with Bayesian Techniques in Lung Cancer", AMIA Annual Symposium, pp. 838-847, 2012

7. K. Jayasurya, G. Fung, S. Yu, C. Dehing-Oberije, D. De Ruysscher, A. Hope, W. De Neve, Y. Lievens, P. Lambin , and A. L. A. J. Dekker, “Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy” Medical Physics, Vol. 37, pp. 1401-1407, 2010.
8. S. Mandal1 and I. Banerjee, “Cancer Classification Using Neural Network”, International Journal of Emerging Engineering Research and Technology Vol. 3(7), pp. 172-178, 2015.
9. E. Adetiba and O. Olugbara, “Lung Cancer Prediction Using Neural Network Ensemble with Histogram of Oriented Gradient Genomic Features” The Scientific World Journal”, Vol. 2015, pp. 1-17, 2015.

AUTHORS PROFILE

Mr. Sheik Mansoor has received his M.C.A from Anna University, Tirunelveli and M.Phil. degree from Manonmanium Sundaranar University, Tirunelveli, Tamilnadu in the year 2010 and 2016 respectively. Currently, he is pursuing his Ph.D. in Sadakathullah Appa College, Tirunelveli which is affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu. His research interest includes machine learning, artificial intelligence, and deep learning.

Dr. M Mohamed Sathik, is an accomplished researcher who has a teaching experience over 35 years. He has published more than 200 research papers in UGC, Scopus and SCI indexed international journals. Currently, he is working as Principal in Sadakathullah Appa College, Tirunelveli, which is affiliated to Manonmanium University, Tirunelveli, Tamilnadu. He has received Ph.D. in arts as well as in engineering studies from Manonmanium Sundaranar University, Tirunelveli in 2006 and 2017 respectively.

Early Exposure of Lung Cancer by Combining ANN and SVM Algorithms

M. Sheik Mansoor, M. Mohamed Sathik

Abstract*: Lung cancer is a lethal type of cancers as its rate of spreading is very high compared to the other cancers. Patient who have been affected from Small Cell Lung Cancer (SCLC) has fast outspread rate. Even at initial stage, around 67-75% of cancer victims with SCLC will have fast outspreads and serious damages to the nearby physical parts. Moreover, World Health Organization (WHO) has predicted the count of lung cancer deaths will reach 9.6 million in 2020. Identifying such a lethal type of cancer early can be lifesaving one. Because, cancer cells in lungs are capable of traveling to other body parts even before the doctor detects them in lungs. In this research work, we have designed a combined approach to prognosticate lung cancer and its type using Artificial Neural Networks (ANN) and Support Vector Machine (SVM). To train both the ML algorithm, an open access patient health dataset published by cancer imaging archives is used. The dataset has the information like pretreatment CT scans, 3D image details of tumor and clinical outcomes. The results produced by ANN and SVM algorithm are compared to predict the type of the lung cancer accurately. The result holds good for a real time implementation.

Keywords: Neural Networks, Lung Cancer Prediction, Cancer diagnosis, Support Vector Machine.

I. INTRODUCTION

Lung cancer is a dangerous kind of cancer which originates from the lungs and outspreads to other nearest physical parts in a short time. Lung cancer is the second most cancer which affects men and fifth most cancer which affects women. As stated by Global Cancer Observatory (GCO), every 5.4 person has lung cancer among one million peoples in India. The alarming issue in the raise of lung cancer is, it has very low survival rate compare to any other cancer diseases. In India, 25% of cancer victims loses their life every year. Due to late stage diagnosis and fast outspread, deaths rate of lung cancer is too high compared to other prostate, colorectal, skin, kidney and breast cancers.

In general, lung cancers can be majorly categorized into Non-Small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC). Here, NSCL cancer is further categorized into three major types, such as adenocarcinoma, squamous and large cell carcinoma. The second type that is SCLC is the dangerous cancer, in which cancer cells spreads to different body parts in short time period through lymphatic nodes. The recent report from National Institute of Cancer Prevention and Research (NICPR) says, around 85-90% are of NSCL cancer and 10-15% of cancer are of SCLC.

Identifying lung cancer cells and its forceful exertion in primitive stage is the only possible way to meliorate the patient's chance of survival. But, in the primitive stage, performing traditional way of histopathology using microscopes will not produce the clear cut results.

Sometimes, lesions in small size can't be effectively detected by CT scan device. To overcome these issues and to identify the cancer type in early stage, automatic ML techniques are used on the patient data. It helps the physicians to acquire a clear cut knowledge about condition of the patients. Moreover, it also helps physicians in identifying the type and vigorous of the cancer cells.

In this research work, we have designed a combined approach to prognosticate the lung cancer and its type using Artificial Neural Networks (ANN) and Support Vector Machine (SVM).

ANN maintains an interconnected nodes, called as neurons to gather information by identifying relationships and new pattern between the data. It has three layer such as, input neuron layer, hidden neuron layer and output neuron layer. Neurons in each layers will receive the input data, performs operations and forwards the data to the nearby connected neurons. Each neurons and the edge which connects the neurons has a particular weight. The weight will change on the neurons based on the learnings. ANN allows both forward and backward propagation for learning.

The final result of ANN are produced based on the maximum probability of neurons present in output layer. Even there exist several algorithms to predict the early state lung cancer, using ANN will produce an accurate result as it allows both forward and backward propagation of learning.

SVM is a discriminative classifier, which draws a hyper plane to differentiate the classes that are derived as outputs. The hyper planes are the decision boundaries. The maximum accuracy can be attained only if the SVM draws the hyper plane separating all the objects to its classes correctly.

This paper is organized in a manner such that, Section 2, describes the related works that are performed in predict lung cancer. Section 3, explains the proposed combined ANN and SVM lung cancer prediction approach. Section 4, evaluates the performance of the proposed approach and Section 5 concludes and discusses about the future work of the proposed combined approach.

Revised Manuscript Received on November 22, 2019

M Sheik Mansoor, Research Scholar (Reg. No. 17221192161007), Sadakathullah Appa College, Tirunelveli, affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India. Email: sheikmansoormsm@gmail.com

Dr. M Mohamed Sathik, Principal and Research Supervisor, Sadakathullah Appa College, Tirunelveli, affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India Email: mmdsadiq@gmail.com

II. RELATED WORKS

Machine learning algorithms are the subgroup of Artificial Intelligence (AI), which learns from data samples and generate insights to identify the pattern in given dataset. Every machine learning has two major learning process such as, identifying the dependencies between data and use the identified dependencies to prefigure the output result.

ML algorithms are been used in number of biomedical applications. The important application of ML in medical field is cancer prediction. Right after the evolvement of Personal Health Information (PHI) and increased computation power, ML algorithms are implemented to identify the cancer type. This section shows an elaborate study of related works that are carried out on the ML to predict cancers, especially on lungs.

The popular ML algorithms used in predicting cancers are Support Vector Machine (SVM) Bayesian Networks (BN) and Neural networks (NN).

A. Cancer prediction based on SVM

SVM is a supervised learning technique which performs classification and regression to identify the associations between the data in the given dataset. Kancherla [1], proposed a cancer prediction method using Support Vector Machine (SVM) based Recursive Feature Elimination (RFE) method. Here, image processing algorithms are executed on the dataset to draw out the basic features that differentiates the cancer cells from the normal cells. After extracting the basic attributes from the dataset, Tetrakis Carboxy Phenyl Porphine (TCPP) is added to the image file to identify the defected cells. For differentiation, SVM is executed. This SVM-RFE method attains accuracy up to 87.5%.

Trivedi [3], proposed a hybrid ML algorithm to identify the cancer defected cells in the lungs using SVM. This particular research combines three major categories of SVMs, such as, linear SVM, Radical Basic Function (RBF) SVM and polynomial SVM to categorize the cancerous and non-cancerous cells. The final optimized results are derived from all the three SVMs. Combination of three SVM are mentioned as hybrid classification. Hybrid SVM attains accuracy upto 77.27%.

Nadira [4], proposed a Global Artificial Bee Colony (GABC) SVM based technique for classifying the cancer type. GABC - SVM predicts important feature from dataset, improves run-time of the ML algorithm. The accuracy of the GABC based SVC classifiers gives accuracy rate as 96.42% (calculated from confusion matrix). The GABC SVM system uses k-folded cross validation method to ensure the maximum rate of testing and training.

B. Lung cancer prediction based on BN

Bayesian networks is probabilistic model which is formed by creating a Direct-Acyclic Graph (DAG). In here, nodes represent the variable and edge represent the dependency between the variables. Depending upon the relations between the nodes, results are derived. BN models can effectively handle uncertainty and missing values compared to the SVM.

Arora [5], presented a detailed survey on knowledge representation and ML tool for risk estimation using Bayesian Network (BN) models. Later, Sesen [6], proposed an idea by implementing Bayesian Networks for predicting the survival days and selecting the perfect treatment plan to cancer patients. The treatments are selected based on different treatment plan option on this estimated survival. It also uses a CAMML hybrid casual discovery algorithm to improve Receiver Operation Curve (ROC) rate while classification. Though the system has several advantages like data encoding, handling large patient data. The proposed method does not provide a significant improvement in prediction models with respect to SVM.

Jayasurya [7] proposed a framework to foresee the survival of the patients who are affected from the non-small cell lung cancer (NSCLC) through radiotherapy. BN based prediction method achieves high predictive rate because it combines both the physical and biological factor for predicating the lung cancer. It also uses SVM to improve classification accuracy.

C. Lung cancer prediction based on ANN

Applying Neural Network based algorithm in medical applications is in practice for a long time. Lakshmanakumar [2], proposed a framework to analyze the CT images using the Optimal Deep Neural Network (ODNN) and Linear Discriminate Analysis (LDA). Here, ODNN extracts the attributes of CT scanned image and LDR is reduces the dimensional of the attribute and classify the tissues as malignant or benign. It gives 94.5% accuracy.

Mandal [8], proposed an ANN based technique to categorize breast cancer data and lung adenocarcinoma. It uses Multilayer Feed Forward (MLFF) neural network to identify cancer cells. The result produced by the ANN method give a high rate of accuracy in classifying the cancer type.

Adetiba [9] presented a comparison between SVM and ANN ensembles and non-ensemble variants of cancer prediction. For extracting the important information from the dataset, two techniques were used. They are Histogram of Oriented Gradient (HOG) and Local Binary Pattern (LBP). Because of combining the ANN and HOG genomic feature, cancer cells are detected in the primitive stage itself. The accuracy rate of this machine learning system is 95.90%. But accuracy reduces when the missing value increases.

III. PROPOSED FRAMEWORK

The proposed approach intends to create a design for prognosticate lung cancer at early stage based on Artificial Neural Network and Support Vector Machine. Both SVM and ANN machine leaning algorithm has different structure to learn from the dataset. ANN uses layers of nodes or neurons and SVM creates an ‘n’ dimensional (n classes) hyper plane from the given data. In general, both algorithms are promising good on predicting the final results.

In this research work, both ANN and SVM are trained differently using the same datasets and the final results are compared to predict lung cancer type. The workflow structure of the proposed combined approach is represented in Fig.1.

To train ANN and SVM, we have used, an open-source lung cancer dataset, which is downloaded from Cancer Imaging Achieve (CIA). The chosen dataset consist of CT scan data of 1019 patients with different cancers. The CT images in this dataset has captured 3D volume of gross tumor.

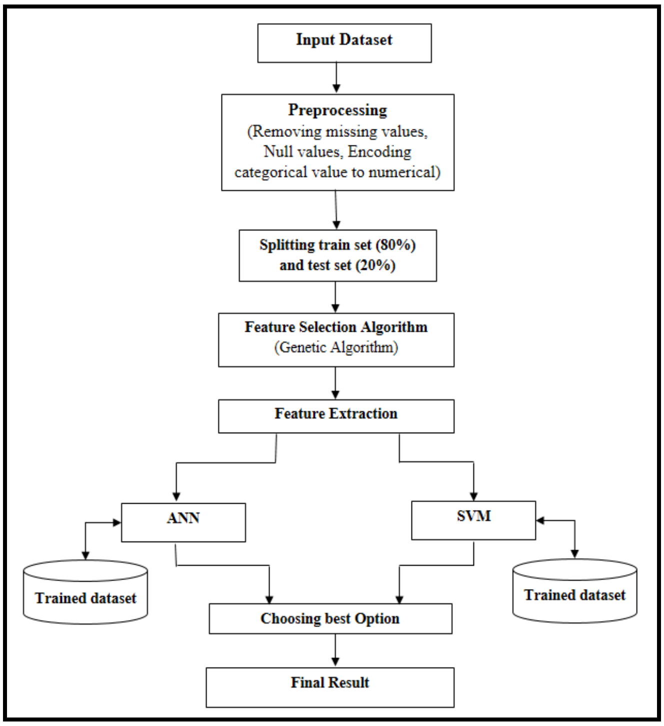


Fig. 1.Workflow of the proposed combined approach

To train machine learning system using Patient Health Information (PHI), raw dataset cannot be given as an input directly. Raw dataset will have missing values, null values, impurities, and irrelevant data in it. To reduce these kind of unwanted data from the dataset, preprocessing is carried out.

A. Data Preprocessing

Pre-processing is the first step in creating a ML algorithm. It transforms raw data into an intelligible format. In general, real world data are incomplete, inconsistent and it will contain several errors. To decrease the errors, preprocessing is carried out on the given dataset. Here, numpy package is used to preprocess. It removes the entire row of the dataset, if it has a void values of a particular feature and column is removed, if it has more than 80% of null values.

To identify the missing or deviated data from a dataset, mean and median for each column. This value is set to be a threshold and missing or deviated values are found from the dataset. After removing the missing values and null values from the dataset, categorical values such as patient-id, histology and stage of caner are encoded into numerical values using pandapy and label_encoder packages.

B. Splitting the dataset

After the data is preprocessed, the filtered dataset is categorized into training dataset and test dataset. Training dataset are given as input to ML algorithm to train and test dataset verifies the results. In the chosen dataset, after preprocessing 398 patient's data (only with NSCLC) are filtered among 1019. In which, 80% of data is taken as training dataset (i.e. 320 patient data). These training dataset are given as an input to ANN and SVM simultaneously. The remaining 78 patient data is chosen as test set.

C. Feature selection algorithm

Feature extraction is performed to decrease the amount unwanted dimensions (i.e. classes) from the pre-processed data. It further reduces the data to a manageable groups of elements. The dimensions are deduced because using the raw dataset directly might complicate the entire learning process. Moreover, it also requires a large computing resources. Feature extraction will create a complete and accurate data groups from the preprocessed data. In this proposed method, Genetic Feature Selection Algorithm (GFSA) is executed to reduce the unwanted dimension from the dataset. Moreover it also reduces the redundant values from the dataset.

GFSA are obtain from the Darwin's idea of natural selection procedure. It only selects the fittest or flawless data from the given dataset. Genetic algorithm has four techniques to scrutinize best-fit values from the dataset. They are selection, inheritance, mutation, and crossover.

Initially, from the given dataset, a couple of candidate-solutions are randomly created. The candidate represent the total population and solution represents individual value. For each solution, the fitness function φ(x) is calculated.

The best fitted individuals are combined together to develop offspring, which later creates another population from the dataset. Later these produced values are given as an input to mutation to find the next best fit values. This procedure is executed continuously to create many generation of values.

In the proposed system, the cross validation of offspring is set to 500 for chosen dataset. So the iteration on finding the fittest values is performed 500times.

D. Artificial Neural Networks

ANN is the main tool used in ML technique for prediction. The dataset inferred from the feature selection module is fed to the input layers of ANN. The inputs are forwarded to hidden layer to process. Hidden layer has several small neuron or units which computes the weight of every neurons and transform the data to the output layer. Hidden layer helps in identifying the complex patterns. The important feature of ANN is backward propagation. It allows the networks to conform the weights of neurons in the hidden layer to improve accuracy of the prediction result. Whereas, ML algorithms like SVM, BN doesn't have the capability of back propagation. Another important feature in ANN is multilayer hidden neurons will extract information from different features. The basic structure of ANN network is represented in Fig.2.

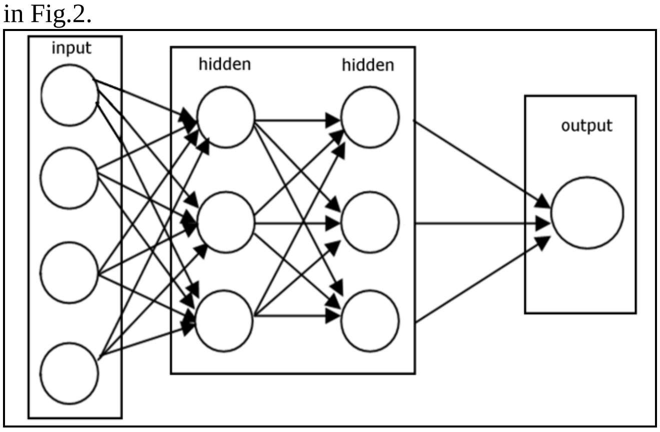


Fig. 2. Stucture of Neural Networks

Every layer in the ANN will have several neurons or nodes in it. Each nodes in the neural network takes numerical values 'X_n' as input. Moreover, each input values are associated with weights 'W_n'.

Net input is calculated using below formula,

$$Net\ input = X_1W_1 + X_2W_2 + \ldots.. X_nW_n \quad (1)$$

As our scrutinized dataset size for training is 320, we have created 320 input layer neurons to process the data. Each input 'X_n' is fed with the associated weight to the nodes in input layers. The weight for each data input is assigned in a random manner using Pseudo Random Number Generator (PRNG).

While performing forward propagation (linear), the results are equated with the real time values to find the deviations between derived output and exact output. The linear forward propagation value 'Y_n' are obtained from the below equation,

$$Y_1 = X_1W_1 + b \quad (2)$$

Where '*b*' is the bias value (zero for first input).

Now the linear forward value 'Y_1' is given to the first activation function as input.

Activation functions or transferring function are present in hidden layers. It is an important function in deriving the final result. It helps in converting the input signal into output signals. Activation function converts the linear input values into a non-linear inputs, which makes the NN to process complex task in short time.

Here, *tan h* activation functions is applied on the hidden layer neurons. Since, the chosen dataset is of multiclass type, two hidden layers are introduced. The *tan h* function is shown below,

$$tan\ h = \frac{e^{2x} -1}{e^{2x} +1} \quad (3)$$

Later, the outputs obtained from the first hidden layer are given to next linear forward propagation Y_2.

After the forward propagation is defined, the backward propagation mechanism has to be created. In neural network, backward propagation computes the gradient of loss function concerning the weights of the network to reduce errors in each neuron.

$$dz_3 = output\ layer - y \quad (4)$$

Loss function concerning weight and bias (for backward propagation – layer 3 to layer 2) is calculated from the below mentioned formulas,

$$\text{Derivative weight } dW_3 = (1/m).\ A_2^T .\ dz_3 \quad (5)$$

$$\text{Derivative bias } db_3 = \sum dz_3 \quad (6)$$

The proposed neural network, performs propagations several time to train neural networks. The final prediction is made on the basis of weights and the bias values in output layer.

E. Support Vector Machine

SVM and ANN algorithm are performed in a parallel manner. SVM classifies the given dataset into two or more classes and creates hyper plane between them. In the perspective of dimension space, SVM will draw a line to differentiate the classes based on the features of the objects.

In training phase, training dataset are fed to linear SVM. In the chosen lung cancer dataset, linear SVM creates three classes based on the attributes of the dataset. SVM derived classes are T stage, M stage and N stage. These classes are separated using three hyper lanes as like showed in Fig. 3.

In here, green represent the stage T, cancer blue represents the stage M cancer and red represents the stage N cancer. After the SVM machine learning system is trained, test dataset are fed to the system to check whether the system produces the correct results. Once after the ANN and SVM machine are ready, the final results are produced to the physicians to decide the cancer type.

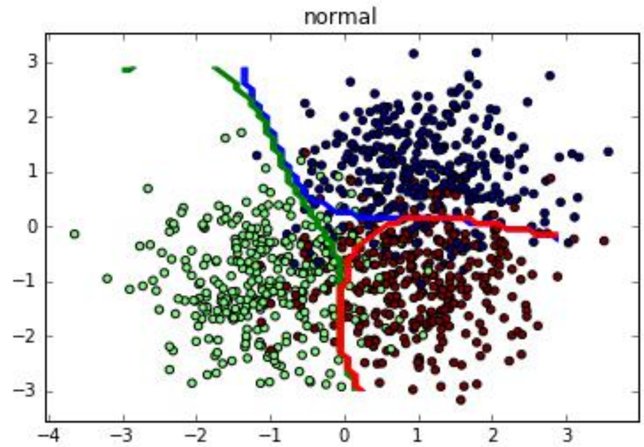


Fig. 3. Three different classes derived from SVM

The final results are decided based on accuracy, sensitiviy and precision. The following section discusses about the results generated by the ANN and SVM.

IV. RESULT AND DISCUSSION

The proposed ANN and SVM machine learning experiments are carried out on the Tensor Flow software, which is a free open-source software developed by Google Inc., The dataset used for the implementation is taken from Cancer Imaging Archives. The chosen dataset consist of CT scan data of 1019 patients with different cancers. Initially, information about the patients, who has affected by the NSCLC cancer is taken out from the given dataset. Around 419 patient records are extracted. Later these, NSCLC cancer data is separated into training dataset and test dataset with the ratio of 70:30. The training dataset are fed as an input to ANN and SVM, simultaneously. They are trained and computed simultaneously for best prediction results. Accuracy is an important performance metric of any classification model. Accuracy is the fraction of predictions that the ML algorithm has predicted accurately. It depends on four possibility of the predicted results, such as, true positive, true negative, false positive, false negative.

The accuracy can be found using the below formula,

$$\text{Accuracy} = \frac{\text{Number of prediction that are correct}}{\text{Total number of prediction made}} \quad (7)$$

Table. 1. Prediction of cancer type using ANN method

Predicted	True/Actual			
		Type ‘T’	Type ‘M’	Type ‘N’
	Cancer Type ‘T’	96	8	4
	Cancer Type ‘M’	5	89	5
	Cancer Type ‘N’	4	5	104

Table. 2. Prediction of cancer type using linear SVM

Predicted	True/Actual			
		Type ‘T’	Type ‘M’	Type ‘N’
	Cancer Type ‘T’	101	6	2
	Cancer Type ‘M’	7	95	8
	Cancer Type ‘N’	8	5	88

Table 1 and Table 2, represent the prediction made through ANN and SVM respectively. Accuracy of ANN model is 90.2%. In 320 total predicted value, ANN has correctly predicted 290 values. However, accuracy of SVM algorithm is 88%, where 284 predictions are made correctly.

The second important performance metric of ML algorithm is precision. It is the fraction of relevant information retrieved (i.e.) in lung cancer type prediction, what fraction of patients belong to a particular cancer type.

In predicting lung cancer type, precision value of type ‘x’ cancer is found by the fraction of correctly predicted type ‘x’ cancer from the total prediction. Precision is calculated by the formula specified below,

$$\text{Precision (Type ‘x’)} = \frac{\text{No. of correctly predicted Type ‘x’}}{\text{Total predicted Type ‘x’ cancer}} \quad (8)$$

So, the precision for type ‘T’ cancer is 96/108 = 88% for ANN machine learning system and 92% for linear SVM classification. Likewise the precision is calculated for type M and type N cancers and shown in Table 3.

Table. 3. Precision values of ANN and SVM algorithm for given dataset.

Precision values (in percentage)		
	ANN	**SVM**
Cancer Type ‘T’	88.8%	92.6%
Cancer Type ‘M’	90.8%	86.3%
Cancer Type ‘N’	92.6%	87.1%

The third most important metrics of ML algorithm is recall. It is the fraction of relevant information that are retrieved.

$$\text{Recall} = \frac{\text{No. of correctly predicted type ‘x’ cancer}}{\text{No. of actual type x cancer patients}} \quad (9)$$

The recall value of the ANN and SVM for three classes (types of cancer) are shown in table 4.

Table. 4. Recall value of ANN and SVM for given dataset

Recall values (in percentage)		
	ANN	**SVM**
Cancer Type ‘T’	91.4%	87%
Cancer Type ‘M’	86.4%	89.2%
Cancer Type ‘N’	91.2%	89.7%

After accuracy, precision and recall values are derived from ANN and SVM, physicians can manually go through the results and can find out the exact scenario of the patients. Instead of relying on single ML algorithm, physicians can use both ANN and SVM. It will give an accurate and clear cut information regarding the patient health condition.

V. CONCLUSION

Early exposure of lung cancer and its type is a difficult and complicated process in the medical-machine learning field. Because, the test information gather from patient might not show the exact information of the patients. Some small sized cancer cells are not visible even the images captured through CT scan device. Predicting lung cancer type using a single ML algorithm might provide an erroneous result due to its possibility of false positive errors. To overcome these issues, the proposed design uses both Artificial Neural Networks and Support Vector Machines on the same dataset. For extracting the important information from raw dataset, a genetic (iterative) based method is used. The predictions derived from ANN and SVM will help the physicians to give a better treatment plan to the patients.

APPENDIX

The dataset is downloaded from Cancer Imaging Archives. It depicts a detailed information of 1019 patent records, in which 419 patients are affected by NSCL lung cancer. Information like patient id, age, gender, type of cancer, tissue size, stage, survival days, death status are provided.

REFERENCES

1. K. Kancherla and S. Mukkamala,” Feature Selection for Lung Cancer Detection Using SVM Based Recursive Feature Elimination Method”, Lecture Notes in Computer Science, Vol. 7256, pp. 168-176,2012.
2. S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N Arunkumar, and G. Ramirez, ”Future Generation Computer System, Vol. 92, pp.374-382, 2018
3. A. Trivedi and P. Shukla, “Lung Cancer Diagnosis by Hybrid Support Vector Machine”, Communications in Computer and Information Science, Vol. 628, pp. 177-187, 2016.
4. T. Nadira and Z. Rustama, “Classification of Cancer Data Using Support Vector Machines with Features Selection Method Based on Global Artificial Bee Colony”, Vol. 2023(1), pp. 1-7, 2018.
5. Arora, P., Boyne, D., Slater, J. J., Gupta, A., Brenner, D. R., and Druzdzel, “Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine, Value in Health”,Vol. 22, pp.437-445, 2019.
6. M. B. Sesen, T. Kadir, R. B. Alcantara, J. Fox, and M. Brady,” Survival Prediction and Treatment Recommendation with Bayesian Techniques in Lung Cancer”, AMIA Annual Symposium, pp. 838-847, 2012

7. K. Jayasurya, G. Fung, S. Yu, C. Dehing-Oberije, D. De Ruysscher, A. Hope, W. De Neve, Y. Lievens, P. Lambin , and A. L. A. J. Dekker, “Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy” Medical Physics, Vol. 37, pp. 1401-1407, 2010.
8. S. Mandal1 and I. Banerjee, “Cancer Classification Using Neural Network”, International Journal of Emerging Engineering Research and Technology Vol. 3(7), pp. 172-178, 2015.
9. E. Adetiba and O. Olugbara, “Lung Cancer Prediction Using Neural Network Ensemble with Histogram of Oriented Gradient Genomic Features” The Scientific World Journal”, Vol. 2015, pp. 1-17, 2015.

AUTHORS PROFILE

Mr. Sheik Mansoor has received his M.C.A from Anna University, Tirunelveli and M.Phil. degree from Manonmanium Sundaranar University, Tirunelveli, Tamilnadu in the year 2010 and 2016 respectively. Currently, he is pursuing his Ph.D. in Sadakathullah Appa College, Tirunelveli which is affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu. His research interest includes machine learning, artificial intelligence, and deep learning.

Dr. M Mohamed Sathik, is an accomplished researcher who has a teaching experience over 35 years. He has published more than 200 research papers in UGC, Scopus and SCI indexed international journals. Currently, he is working as Principal in Sadakathullah Appa College, Tirunelveli, which is affiliated to Manonmanium University, Tirunelveli, Tamilnadu. He has received Ph.D. in arts as well as in engineering studies from Manonmanium Sundaranar University, Tirunelveli in 2006 and 2017 respectively.

Journal of Engineering Sciences
Vol 11, Issue 5,May/2020
ISSN NO:0377-9254

A COMPARATIVE STUDY OF SUPERVISED AND UNSUPERVISED MACHINE LEARNING TECHNIQUES ON LUNG CANCER PREDICTION

M. Sheik Mansoor[1] and M. Mohamed Sathik[2]

[1]*Research Scholar r (Reg. No. 17221192161007), Sadakathullah Appa College, Affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India*

[2]*Principal and Research Supervisor, Sadakathullah Appa College, Tirunelveli, Affiliated to Manonmanium Sundaranar University, Tirunelveli, Tamilnadu, India*

***Abstract*— Lung cancer is one of the most dangerous type of cancers which has the high spread rate. Lung cancer metastases spreads through fluid lymph nodes and bloodstreams to other organs like bone, glands and brains. Due to the air and industrial pollution the rate of people who has affected by the lung cancer is increasing enormously. According to the prediction reports of World Health Organization (WHO) the number of lung cancer deaths will reach 9.6 million in 2020, which is an alarming issue. Diagnosis the lung cancer at its earlier stage could help the physicians to treat the patients. Though the manual analysis of CT scan exists in the medical field, it is too hard for the medical advisors to predict the exact stage of the disease using the CT scan images. Hence, the medical informatics research community has created several machine learning model to predict the lung cancer and its type in the earlier stage. In this comparative research study, we have downloaded the lung cancer dataset from the Cancer Image Archive and given as the input to the two most accepted machine learning models such as, Artificial Neural Networks (ANN), Support Vector Machine (SVM) from supervised learning method and another unsupervised dataset as input for Apriori and K-means model from unsupervised learning to observe the changes. The final results and the performance metrics of the machine learning algorithms such as accuracy, precision and recall are compared with each other and tabulated.**

***Keywords*— Machine Learning; Lung Cancer Prediction; Supervised Learning; Cancer Diagnosis.**

1. INTRODUCTION

Lung cancer is a type cancer which starts in the cells of the lungs and spreads to the other parts of the human body [3]. Likewise, cancer cells such as breast, mouth and kidney can also spread to the lungs via lymph nodes or bloodstreams[2, 3]. The lung are is made up of sponge like structure in the chest of the human. The main objective of the lungs is to take oxygen into the body and release the carbon dioxide [1]. While breathing air passes through pipe like structure called trachea and propagates through bronchi nodes to enter lungs and come outs in the same path. The small sized holes in the bronchi nodes called alveoli passes the oxygen to the blood and takes out the carbon dioxide out from the blood [4,5].

At initial stage of lung cancer, DNA of the patient will change or damage and mutate the genes. Mutated genes will not work properly because they will not get any instruction from DNA properly or in a correct manner. This will cause the cells in the lung to divide and grow out of control in and around the lungs and causes the lung cancer [6].

As stated by Global Cancer Observatory (GCO), every 5.4 person has lung cancer among one million peoples in India. The alarming issue in the raise of lung cancer is, it has very low survival rate compare to any other cancer diseases. In India, 25% of cancer victims loses their life every year. Due to late stage diagnosis and fast outspread, deaths rate of lung cancer is too high compared to other prostate, colorectal, skin, kidney and breast cancers [7]. Accurately identify the lung cancer cells in its initial stage through manual analysis of CT scan is not possible. It makes difficult for medical advisors to predict the exact stage of the cancer using the CT scan images.

To overcome these issues and to identify the cancer type in early stage, Machine Learning techniques are used on the patient data. It helps the physicians to acquire a clear cut knowledge about condition of the patients. Moreover, it helps physicians in identifying the type and vigorous of the cancer cells [8, 9].

Machine learning techniques can be classified into two major types based on its application and working nature. While considering the lung cancer prediction several research contributions and prediction methods were been introduced. In this research work, we have taken two supervised learning methods such as Artificial Neural Networks (ANN), Support Vector Machine (SVM) and two unsupervised learning methods Apriori and K-means for this comparative study. The datasets were downloaded from the open- source Cancer Imaging archives and given as the training set to these machine learning algorithm. The preprocessing, feature extraction and selection are kept same for all these four methods.

This comparative study paper is organized in such a manner such that, Section 2, describes the difference between the supervised learning and unsupervised learning. Section 3, explains the preprocessing, feature extraction and selection. Section 4, evaluates the performance of the ANN, SVM, Apriori and K-means and Section 5 concludes and discusses about the future work of the comparative study.

2. SUPERVISED LEARNING AND UNSUPERVISED LEARNING

In supervised learning, the machine learning system is trained with a well labeled information, which means that some data is already tagged with the correct answer. So it can be directly compared with learning process. A supervised learning algorithm learns from labeled training data, helps you to predict outcomes for unforeseen data. Where as in unsupervised machine learning technique, the dataset will not have a clear label. Instead it should be programmed in a manner, in which it should discover the information on its own. Unsupervised machine learning techniques can perform more complex processing tasks compared to the supervised learning algorithm but the results of the unsupervised machine learning are unpredictable compared to other deep learning and natural learning process.

A. Supervised learning algorithm: Artificial Neural Networks (ANN)

ANN maintains an interconnected nodes, called as neurons to gather information by identifying relationships and new pattern between the data. It has three layer such as, input neuron layer, hidden neuron layer and output neuron layer. Neurons in each layers will receive the input data, performs operations and forwards the data to the nearby connected neurons. Each neurons and the edge which connects the neurons has a particular weight. The weight will change on the neurons based on the learnings. ANN allows both forward and backward propagation for learning.

The final result of ANN are produced based on the maximum probability of neurons present in output layer. Even there exist several algorithms to predict the early state lung cancer, using ANN will produce an accurate result.

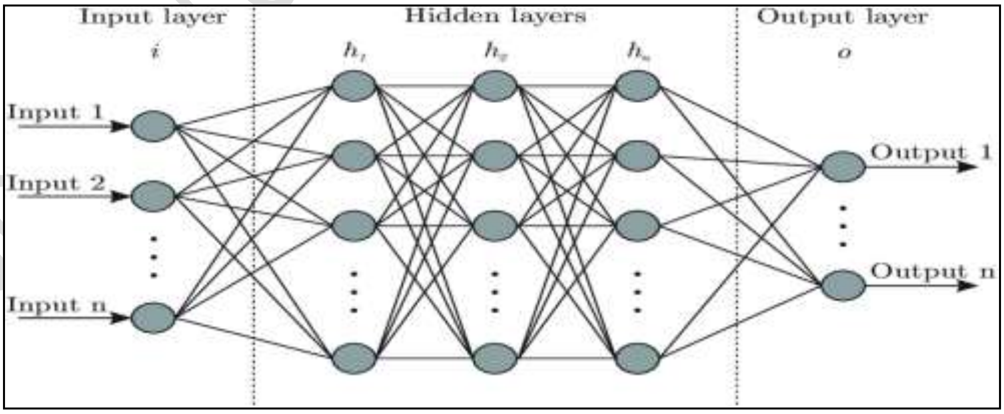


Fig. 1. Stucture of Neural Networks

B. Supervised learning algorithm : Support Vector Machine (SVM)

SVM is a supervised learning technique which performs classification and regression to identify the associations between the data in the given dataset. SVM is a discriminative classifier, which draws a hyper plane to differentiate the classes that are derived as outputs. The hyper planes are the decision boundaries. The maximum accuracy can

Journal of Engineering Sciences
Vol 11, Issue 5,May/2020
ISSN NO:0377-9254

be attained only if the SVM draws the hyper plane separating all the objects to its classes correctly.

In here, support vectors are data that are very closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, the margin of the classifier can be maximized to get the clear idea. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build accurate SVM model.

In here, support vectors are data that are very closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, the margin of the classifier can be maximized to get the clear idea. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build accurate SVM model.

C. Unsupervised learning algorithm – Apriori Algorithm

The Apriori algorithm is a classical frequent item sets generation algorithm and a milestone in the development of data mining. It is used for finding frequent item in a dataset for Boolean association rule. Apriori algorithm uses prior knowledge of frequent item properties. An iterative approach or level-wise search where k-frequent item are used to find k+1 item.

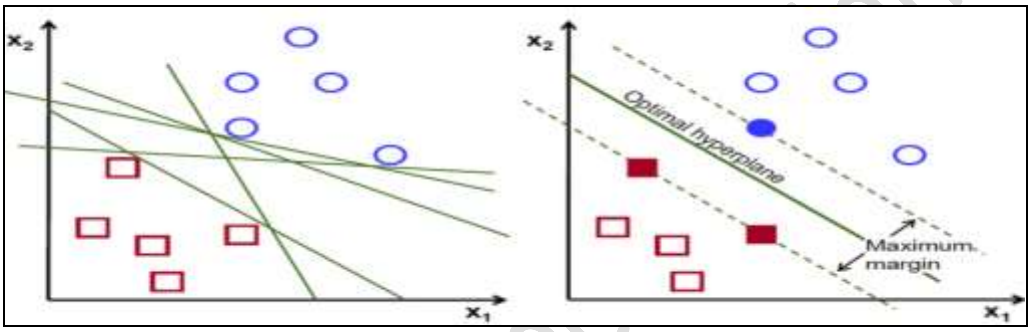


Fig. 2. Hyperplane of SVM

To improve the efficiency of level-wise generation of frequent item, an important property is used called Apriori property which helps by reducing the search space. Apriori property states that, all non-empty subset of frequent item set must be frequent.

D. Unsupervised learning algorithm – K Means

K-means algorithm is an iterative algorithm that tries to partition the dataset into K pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group. It tries to make the inter-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

3. HANDLING DATA

A. Pre-processing

The pre-processing is an important task that is used for transforming the raw data into a useful and efficient data. The pre- processing include several steps such as data cleaning, transformation and reduction. Data cleaning is a process in which the missing values are replaced or removed and involves in removing the noise in the data through several methods such as regression, clustering or binning method. The transformation of data includes normalization of data, selection of attributes, transferring of continuous dataset to discrete through discretization and generation of hierarchy. The data reduction includes several actions such as aggregation of data as per the need, subset selection section in the particular attribute, replacement of original data into a data representation through parametric or non-parametric numerosity reduction and reduction of dimensionality.

B. Feature Selection

Feature selection is a dimension reduction method which is used to select the relevant feature for constructing the model. It includes four important approaches such as wrapper, filter, embedded and hybrid approaches for selecting the features. Wrapper approach is an approach which is highly complex computation. It selects the feature through classification and uses a learning algorithm for calculating the accuracy of the classification. The filter approaches select the subset of the feature without using any leaners. The database with higher dimension can use this type of feature selection approach. The embedded approach selects the feature during the training of the data and it uses applied learning algorithms for deriving the specificity of the approach. The hybrid approach is another approach where the filter and wrapper approaches are used in combination for selecting the feature. The feature is selected through the filter approach and are tested with wrapper approach. Thus it uses both the advantages for feature selection.

C. Feature Extraction

Feature extraction is another dimensionality reduction method through which the raw data will be transformed into a group of manageable data for further processing. It plays an important role in image processing as multiple parameters are needed to process the images. It includes low level extraction, edge level extraction, curvature extraction, shape detection, motion detection and so on. Here the low level processing of images includes several detection such as detection of edges, detection of corners, blob detection for detecting the regions in the images, ridge detection for extracting the thin line which is brighter than the nearby regions and feature transform through difference in scales of images. The curvature extraction intends to extract the direction of edges. It also identifies the change in intensity of images and the autocorrelation. The shape detection involves in finding the threshold of the images, region extraction and template matching. It also includes hough transformation which involves in extract the imperfect features of the objects by comparing it within the class through voting procedure. The motion detection model involves in extracting the motion of images and the optical flow by admiring the area of the images.

4. Performance evaluation of ANN, SVM, Apriori and K- means

A. Performance comparison of ANN and SVM

The ANN and SVM machine learning experiments are carried out on the Tensor Flow software, which is a free open-source software developed by Google Inc., The dataset used for the implementation is taken from Cancer Imaging Archives. The chosen dataset consist of CT scan data of 1019 patients with different cancers. Initially, information about the patients, who has affected by the NSCLC cancer is taken out from the given dataset. Around 419 patient records are extracted. Later these, NSCLC cancer data is separated into training dataset and test dataset with the ratio of 70:30. The training dataset are fed as an input to ANN and SVM, simultaneously. They are trained and computed simultaneously for best prediction results.

		True/Actual		
		Type 'T'	Type 'M'	Type 'N'
Predicted	Cancer Type 'T'	96	8	4
	Cancer Type 'M'	5	89	5
	Cancer Type 'N'	4	5	104

Table. 1. Prediction of cancer type using ANN method

		True/Actual		
		Type 'T'	Type 'M'	Type 'N'
Predicted	Cancer Type 'T'	101	6	2
	Cancer Type 'M'	7	95	8
	Cancer Type 'N'	8	5	88

Table. 2. Prediction of cancer type using linear SVM

Table 1 and Table 2, represent the prediction made through ANN and SVM respectively. Accuracy of ANN model is 90.2%. In 320 total predicted value, ANN has correctly predicted 290 values. However, accuracy of SVM algorithm is 88%, where 284 predictions are made correctly.

The second important performance metric of ML algorithm is precision. It is the fraction of relevant information retrieved (i.e.) in lung cancer type prediction, what fraction of patients belong to a particular cancer type.

In predicting lung cancer type, precision value of type 'x' cancer is found by the fraction of correctly predicted type 'x' cancer from the total prediction. Precision is calculated by the formula specified below,

Precision (Type 'x') = (No. of correctly predicted Type 'x') / (Total predicted Type 'x' cancer)

Precision values (in percentage)		
	ANN	SVM
Cancer Type 'T'	88.8%	92.6%
Cancer Type 'M'	90.8%	86.3%
Cancer Type 'N'	92.6%	87.1%

Table. 3. Precision values of ANN and SVM algorithm for given dataset.

The recall can be derived through the below mentioned formula,

Recall = (No. of correctly predicted type 'x' cancer) / (No. of actual type x cancer patients)

Table. 4. Precision values of ANN and SVM algorithm for given dataset.

On comparing both the algorithms we can observe that the ANN is more effective than SVM in many of the cases. On observing the precision value ANN works better than SVM but in few cases such as in detecting cancer type T the precision value is good for SVM. On comparing the recall value we can observe that ANN results good than SVM but in few cases such as predicting cancer type M, the SVM computes better than ANN as the correctness of result is good.

B. Comparison of Apriori and K-Means

This research work used a dataset which is needed to extract to achieve useful information about the effect of k-means algorithm to apriori algorithm from computation time and rule achieved. The dataset used consists of 8243 disease diagnose data. Medical data variables consist of disease diagnosis, age group, gender, the status of care. The partial data used can be seen in Table 5.

In the first approach, directly apply the apriori algorithm in the dataset to 4 input variables, namely disease diagnosis, age group, gender, the status of care in order to obtain confidence values, rules and computational time on apriori algorithms. The test results obtained from the Apriori algorithm can be seen in Table 6.

This rule information obtained in the Large Itemset 4 results in two rules, namely the diagnosis of another allergic rhinitis with the age group of female children and outpatient status. Then, the diagnosis of postoperative disease with the adult age group gender male and outpatient status with each confidence value of 69%. From these results, it can be seen that the information obtained from the Apriori algorithm is still lacking.

Recall values (in percentage)		
	ANN	SVM
Cancer Type 'T'	91.4%	87%
Cancer Type 'M'	86.4%	89.2%
Cancer Type 'N'	91.2%	89.7%

As shown in Table 7 above, the combination of the K-Means algorithm and the Apriori algorithm produces more complete and detailed information compared to the results obtained by the application of a priori algorithm only.

S. No.	Disease Diagnosis	Age Cluster	Gender	Status of Care
1	Observation of Febris	Baby	Male	Outpatient
2	Observation of Febris	Baby	Female	Outpatient
3	Observation of Febris	Baby	Male	Outpatient
4	Observation of Febris	Baby	Female	Outpatient
5	Paronychia	Adult	Female	Outpatient
6	Hnp Lumbalis	Adult	Male	Outpatient
:	:	:	:	:
8243	Disputes with the counselor	Toddlers	Female	Outpatient

Table 5. Sample patient diagnosis data in 2016

Using Apriori							
Full Data							
Disease Diagnose	Cataracts not Specified	Cataracts not Specified	Cataracts not Specified	Cataracts not Specified	Another Allergic Rhinitis	Another Allergic Rhinitis	Post Operation
Age Cluster	--	Elder	Elder	Elder	--	Child	Adult
Gender	Male	--	Female	--	Female	Female	Male
Status of Care	Out	Out	Out	Out	Out	Out	Out
Confidence (%)	69	76	60	66	69	69	69

Table 6. Data processing using Apriori algorithm

K-Means + Apriori				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Disease Diagnose	Cataracts not Specified	Cataracts not Specified	Another allergic rhinitis	Post Operation
Age Cluster	Elder	Elder	Child	Adult
Gender	Male	Female	Female	Male
Status of Care	Outpatient	Outpatient	Outpatient	Outpatient
Confidence (%)	66	66	92	93

Table 7. Data processing using K- Means and Aprior

Meanwhile, the computation time of K-Means and Apriori algorithms combinations are faster than the Apriori algorithm, where the total time from K-Means algorithm and Apriori algorithms combinations are 17.41 minutes while the total time of the Apriori algorithm is 21.93 minutes.

5. CONCLUSION

In this comparative research study, we have downloaded the lung cancer dataset from the Cancer Imaging archives and given as the input to the two most accepted machine learning models such as, Artificial Neural Networks (ANN), Support Vector Machine (SVM) and another patient dataset is given as input for unsupervised learning method such as Apriori and K-means model for observing the performance difference. The final results and the performance metrics of the machine learning algorithms such as accuracy, precision and recall are compared with each other and tabulated. Thus the comparison of unsupervised and supervised algorithms are compared.

Reference

1. K. Kancherla and S. Mukkamala," Feature Selection for Lung Cancer Detection Using SVM Based Recursive Feature Elimination Method", Lecture Notes in Computer Science, Vol. 7256, pp. 168-176,2012.
2. S.K. Lakshmanaprabu, S.N. Mohanty, K. Shankar, N Arunkumar, and G. Ramirez, "Future Generation Computer System, Vol. 92, pp.374-382, 2018
3. A. Trivedi and P. Shukla, "Lung Cancer Diagnosis by Hybrid Support Vector Machine", Communications in Computer and Information Science, Vol. 628, pp. 177-187, 2016.
4. T. Nadira and Z. Rustama, "Classification of Cancer Data Using Support Vector Machines with Features Selection Method Based on Global Artificial Bee Colony", Vol. 2023(1), pp. 1-7, 2018.
5. Arora, P., Boyne, D., Slater, J. J., Gupta, A., Brenner, D. R., and Druzdzel, "Bayesian Networks for Risk Prediction Using Real-World Data: A Tool for Precision Medicine, Value in Health",Vol. 22, pp.437-445, 2019.

Vol 11, Issue 5,May/2020
ISSN NO:0377-9254

6. M. B. Sesen, T. Kadir, R. B. Alcantara, J. Fox, and M. Brady," Survival Prediction and Treatment Recommendation with Bayesian Techniques in Lung Cancer", AMIA Annual Symposium, pp. 838-847.
7. K. Jayasurya, G. Fung, S. Yu, C. Dehing-Oberije, D. De Ruysscher, A. Hope, W. De Neve, Y. Lievens, P. Lambin , and A. L. A. J. Dekker, "Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy" Medical Physics, Vol. 37, pp. 1401-1407, 2010.
8. Dharshinni N P, Mawengkang H and Nasution M K M 2018. Mapping of medicine data with k-means and apriori combinations based on patient diagnosis. International Conference on Computing and Applied Informatics.Vol 2 (978).
9. E. Adetiba and O. Olugbara, "Lung Cancer Prediction Using Neural Network Ensemble with Histogram of Oriented Gradient Genomic Features" The Scientific World Journal", Vol. 2015, pp. 1-17, 2015.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

A New Innovative Approach for Natural Image Denoising Using Genetic Algorithm and Thresholding

S.Aaisha Nazleem, L.S.Subbulakshmi

Programmer, Dept. of I.T., Sadakathulla Appa College, Tirunelveli, Tamilnad, India

Assistant Professor, Dept. of I.T., Sadakathulla Appa College, Tirunelveli, Tamilnad, India

ABSTRACT: The image de-noising naturally corrupted by noise is a classical problem in the field of signal or image processing. Additive random noise can easily be removed using simple threshold methods. De-noising of natural images corrupted by Speckle noise, salt & pepper noise and Poisson using wavelet techniques are very effective because of its ability to capture the energy of a signal in few energy transform values. Natural Image De-noising plays a cardinal role in the field of image pre-processing. Natural Image is recurrently debauched by noise in its acquisition and transmission. Natural Image De-noising is the process of undesirable noise in order to reinstate the original image. This thesis presents, image de-noising scheme based on Wavelet Transform. First the input natural image is taken and then the noise is applied in the image. Among different types of noises, this thesis focuses only the speckle noise, poisson noise and salt & pepper noise. And then apply Wavelet Transform on to the noisy natural image to produce the decomposed image representation. This thesis uses four different types of Wavelet Families such as COIF4, COIF5, RBio6.8 and Sym8. Finally threshold shrinkage methods are applied to de-noise the noisy coefficients and then apply the inverse transform to get the de-noised image. Among several shrinkage methods this thesis takes only four shrinkage methods such as Visu Shrink, Neigh Shrink, Sure Shrink and Modineighshrink. After the denoising process are completed the performance are analysed by using the performance metric such as Peak Signal to Noise Ratio(PSNR), Root Mean Square Error(RMSE), Mean Structural Similarity Index Measure(MSSIM), Mean Absolute Error(MAE),Normalized Cross Correlation(NCC) , Normalized Absolute Error(NAE).

KEYWORDS: Image Denoising,Natural Image, DFT,DWT

I. INTRODUCTION

Image noise means unwanted signal. It is random variation of color information and brightness in images, and is usually an aspect of electronic noise. It is an undesirable by-product of image capture that adds spurious and extraneous information. Speckle is a granular 'noise' that inherently exists in and degrades the quality of the active radar, synthetic aperture radar (SAR), natural ultrasound and optical coherence tomography images. Speckle noise in conventional radar results from random fluctuations in the return signal from an object that is no bigger than a single image-processing element. It increases the mean grey level of a local area. Speckle noise in SAR is generally serious, causing difficulties for image interpretation. Shot noise or Poisson noise is a type of electronic noise which can be modelled by a Poisson process. In electronics shot noise originates from the discrete nature of electric charge. Shot noise also occurs in photon counting in optical devices, where shot noise is associated with the particle nature of light .Salt-and-pepper noise–Fat-tail distributed or "impulsive" noise is sometimes called salt-and pepper noise. Any image having salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions. In salt-and-pepper noise corresponding value for black pixels is 0 and for white pixels the corresponding value is 1.

II. RELATED WORK

Curvelet and Wavelet Image Denoising [1]this paper describes the image denoising of Curvelet and Wavelet Image Denoising by using 4 different additive noises like Gaussian noise, Speckle noise, Poisson noise and Salt & Pepper noise and also by using 4 different threshold estimators like heursure, rigrsure, mini-maxi and squawolog for wavelet

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

and curvelet transform both. It offer exact reconstruction, stability against perturbation, ease of implementation and low computational complexity. The curvelet reconstruction offering visual sharp image and in particular, higher quality recovery of edges and of faint linear and curvilinear features . Image Denoising Using Wavelet Thresholding [2] This paper proposes and explore different wavelets methods in digital image denoising. Using several wavelets threshold technique such as SURE Shrink, Visu Shrink, and Bayes Shrink in search for efficient image denoising method. This paper extend the existing technique and provide a comprehensive evaluation of the proposed method. Wiener filtering technique is the proposed method which was compared and analysed, while the performance of all the techniques were compared to ascertain the most efficient method. Image Denoising Techniques[3] This paper is to provide a review of some of those techniques that can be used in image processing (denoising). This paper outlines the brief description of noise, types of noise, image denoising and then the review of different techniques and their approaches to remove that noise . The aim of this paper is to provide some brief and useful knowledge of denoising techniques for applications using images to provide an ease of selecting the optimal technique according to their needs. Speckle Noise is a natural characteristic of medical ultrasound images. Speckle Noise reduces the ability of an observer to distinguish fine details in diagnostic testing. It also limits the effective implementation of image processing such as edge detection, segmentation and volume rendering in 3 D. Therefore; treatment methods of speckle noise were sought to improve the image quality and to increase the capacity of diagnostic medical ultrasound images. Such as median filters, Wiener and linear filters (Persona & Malik, SRAD ... ..).The method used in this work is 2-D translation invariant forward wavelet transform, it is used in image processing, including noise reduction applications in medical imaging[4]. Mohammad Ali says [5] A novel method for image denoising which relies on the DBNs' ability in feature representation. This work is based upon learning of the noise behavior. Generally, features which are extracted using DBNs are presented as the values of the last layer nodes. The nodes in the last layer of trained DBN are divided into two distinct groups of nodes. After detecting the nodes which are presenting the noise. A reduction of 65.9% in average mean square error (MSE) was achieved when the proposed method was used for the reconstruction of the noisy images. A novel self-learning based image decomposition framework. Based on the recent success of sparse representation, the proposed framework first learns an over-complete dictionary from the high spatial frequency parts of the input image for reconstruction purposes. This method perform unsupervised clustering on the observed dictionary atoms (and their corresponding reconstructed image versions) via affinity propagation, which allows us to identify image-dependent components with similar context information. The proposed and are able to automatically determine the undesirable patterns (e.g., rain streaks or Gaussian noise) from the derived image components directly from the input image, so that the task of single-image denoising can be addressed[6]. In Adaptive Multi-Column Deep Neural Networks with Application to Robust Image Denoising[7] Stacked sparse denoising autoencoders (SSDAs) have recently been shown to be successful at removing noise from corrupted images. However, like most denoising techniques, the SSDA is not robust to variation in noise types beyond what it has seen during training. This paper eliminate the need to determine the type of noise, let alone its statistics, at test time and even show that the system can be robust to noise not seen in the training set. It show that state-of-the-art denoising performance can be achieved with a single system on a variety of different noise types. Additionally, we demonstrate the efficacy of AMC-SSDA as a preprocessing (denoising) algorithm by achieving strong classification performance on corrupted MNIST digits. Contourlet Based Image Denoising[8] This paper proposed contour let based image denoising algorithm which can restore the original image corrupted by salt and pepper noise, Gaussian noise, Speckle noise and the poisson noise. The noisy image is decomposed into sub bands by applying contour let transform, and then a new thresholding function is used to identify and filter the noisy co efficient and take inverse transform to reconstruct the original image. This contourlet technique is computationally faster and gives better results compared to the existing wavelet technique. But this proposed method is not well suited for the removal of salt and pepper noise from the original image. Salt and Pepper Noise Removal[9] Images may be corrupted by salt and pepper impulse noise due to noisy sensors or channel transmission errors. A denoising method by detecting noise candidates and enforcing image sparsity with a patch-based sparse representation is proposed. Compared with traditional impulse denoising methods, including adaptive median filtering, total variation and Wavelet, the new method shows obvious advantages on preserving edges and achieving higher structural similarity to the noise-free images. Parallel Edge Preserving Algorithm for Salt and Pepper Image Denoising [10] this paper a two-phase filter for removing "salt and pepper" noise is proposed. In the first phase, an adaptive median filter is used to identify the set of the noisy pixels; in the second phase, these pixels are restored according to a regularization method, which contains a data-fidelity term reflecting the impulse noise characteristics.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

III.METHODOLOGY

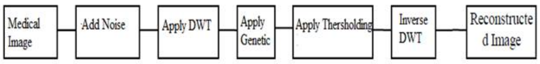


Fig.1 Image denoising block diagram for DWT & Contourlet transform

The overall block diagram of the proposed method is shown in Fig.1.1. In this paper decompose the image using discrete wavelet and then applied genetic algorithm for feature selection and threshold for noise removal. The image de-noising naturally corrupted by noise is a classical problem in the field of signal or image processing. Additive random noise can easily be removed using simple threshold methods. In this paper Visu Shrink, Neigh Shrink, Sure Shrink and Modineighshrink.Min-Max Shrink are used. De-noising of natural images corrupted by Speckle noise and Poisson using wavelet techniques are very effective because of its ability to capture the energy of a signal in few energy transform values. The further details of these modules are discussed below:

Step 1: Input Image Choosing

This is the first step of the proposed method. In this step the input image is get from the user via open dialog box control.

Step 2: Apply Noise

This is the second step of the proposed method. In this step the input is corrupted by noise. Image noise is random (not present in the object imaged) variation of brightness or color information in images, and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an undesirable by-product of image capture that adds spurious and extraneous information.In this paper two types of noises are used. They are Speckle Noise and Poisson Noise.

Speckle is a granular 'noise' that inherently exists in and degrades the quality of the active radar, synthetic aperture radar (SAR), natural ultrasound and optical coherence tomography images. The vast majority of surfaces, synthetic or natural, are extremely rough on the scale of the wavelength. Poisson noise is a type of electronic noise which can be modelled by a Poisson process. In electronics shot noise originates from the discrete nature of electric charge. Shot noise also occurs in photon counting in optical devices, where shot noise is associated with the particle nature of light.In this paper, we have used three types of noises. There are speckle noise, Poisson noise, salt and pepper noise.

Step 3:Apply Discrete Wavelet Transform

This is the third step of the proposed method. In this step the noisy image is decomposed using discrete wavelet transform. The Discrete Wavelet Transform (DWT) of image signals produces a non- redundant image representation, which provides better spatial and spectral localization of image formation, compared with other multi scale representations such as Gaussian and Laplacian pyramid. Recently, Discrete Wavelet Transform has attracted more and more interest in image de-noising. The DWT can be interpreted as signal decomposition in a set of independent, spatially oriented frequency channels. The signal S is passed through two complementary filters and emerges as two signals, approximation and Details. This is called decomposition or analysis. The components can be assembled back into the original signal without loss of information. This process is called reconstruction or synthesis. The mathematical manipulation, which implies analysis and synthesis, is called discrete wavelet transform and inverse discrete wavelet transform. An image can be decomposed into a sequence of different spatial resolution images using DWT. These are also known by other names, the sub-bands may be respectively called a1 or the first average image, h1 called horizontal fluctuation, v1 called vertical fluctuation and d1 called the first diagonal fluctuation.

The wavelet transform has gained widespread acceptance in signal processing and image compression. Recently the JPEG committee has released its new image coding standard, JPEG-2000, which has been based upon DWT. Wavelet transform decomposes a signal into a set of basis functions. These basis functions are called wavelets. Wavelets are obtained from a single prototype wavelet called mother wavelet by dilations and shifting. The DWT has been introduced as a highly efficient and flexible method for sub band decomposition of signals. The 2D-DWT is nowadays established as a key operation in image processing .It is multi-resolution analysis and it decomposes images into wavelet coefficients and scaling function. In Discrete Wavelet Transform, signal energy concentrates to specific wavelet coefficients. This characteristic is useful for compressing images.

The sub-image a1 is formed by computing the trends along rows of the image followed by computing trends along its columns. In the same manner, fluctuations are also created by computing trends along rows followed by

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

trends along columns. The next level of wavelet transform is applied to the low frequency sub band image LL only. The Speckle noise will nearly be averaged out in low frequency wavelet coefficients. Therefore, only the wavelet coefficients in the high frequency levels need to be thresholded. Several families are available in DWT. Among those this paper consider four families such as coif4, coif5, rbio6.8 and sym8.

Step 4: Feature Selection using Genetic Algorithm

This is the fourth step of the proposed method. In this step the noisy wavelet coefficient feature are selected using genetic algorithm. In a genetic algorithm, a population of strings (called chromosomes or the genotype of the genome), which encode candidate solutions (called phenotypes) to an optimization problem, evolves toward better solutions. Solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. The algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

A typical genetic algorithm requires: A genetic representation of the solution domain

Step 5: Apply Thersholding

This is the fifth step of the proposed method. In this step the image denoised by using thresholding approach. Here, the threshold plays an important role in the denoising process. Finding an optimum threshold is a tedious process. A small threshold value will retain the noisy coefficients whereas a large threshold value leads to the loss of coefficients that carry image signal details. Normally, hard thresholding and soft thresholding techniques are used for such de-noising process. Hard thresholding is a keep or kill rule whereas soft thresholding shrinks the coefficients above the threshold in absolute value. It is a shrink or kill rule. The following are the methods of threshold selection for image de-noising based on wavelet transform.

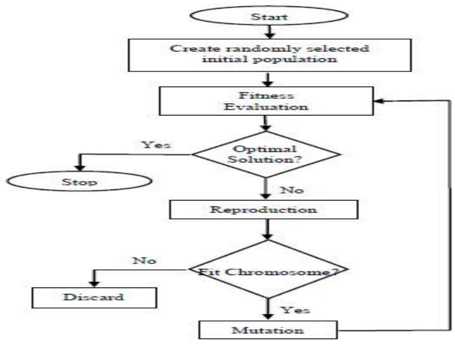


Fig.1.2. Flow Chart of the Genetic Algorithm

Method1: SureShrink

SureShrink is a thresholding technique in which adaptive threshold is applied to sub band, but a separate threshold is computed for each detail sub band based upon SURE (Stein.s Unbiased Estimator for Risk), a method for estimating the loss in an unbiased fashion. The optimal λ and L of every sub band should be data-driven and should minimize the Mean Squared Error (MSE) or risk of the corresponding sub band. Fortunately, Stein has stated that the MSE can be estimated unbiased from the observed data. Neighshrink can be improved by determining an optimal threshold and neighbouring window size for every wavelet sub band using the Stein"s Unbiased Risk Estimate (SURE). For ease of notation, the Ns noisy wavelet coefficients from sub band s can be arranged into the 1-D vector. Similarly, the unknown noiseless coefficients from subband „s" is combined with the corresponding 1-D vector. Stein shows that, for almost any fixed estimator based on the data , the expected loss (i.e risk) $E\{\|\underline{\theta}^s - \theta^s\|_s^2\}$ can be estimated unbiasedly. Usually, the noise standard deviation σ is set at 1, and then

$$E\{\|\overline{O}_s - O_s\|_2^2\} - N_s + E\{\|g(w_s)\|_2^2 + 2\nabla . g(w_s)\}$$

$$g(w_s) = \{g_n\}_{n=1}^{N_s} = \overline{\theta}_s - w_s, \nabla . g = \sum_n \partial g_n / \partial w_n$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Method 2: Visushrink

Heshol can be calculated using formula T= σ√2logn2. This method performs well under a number of applications because wavelet transform has the compaction property of having only a small number of large coefficients. All the rest wavelet coefficients are very small. This algorithm offers the advantages of smoothness and adaption.

Method 3: Neighshrink

Let d(i,j) denote the wavelet coefficients of interest and B(i,j) is a neighborhood window around d(i,j). Also let S2=Σd2(i,j) over the window B(i,j). Then the wavelet coefficient to be thresholded is shrinked according to the formulae, d(i,j)= d(i,j)* B(i,j) where the shrinkage factor can be defined as B(i,j) = (1- T2/ S2(i,j))+, and the sign + at the end of the formulae means to keep the positive value while set it to zero when it is negative.

Method 4: Mod neighshrink

During experimentation, it was seen that when the noise content was high, the reconstructed image using Neighshrink contained mat like aberrations. These aberrations could be removed by wiener filtering the reconstructed image at the last stage of IDWT. The cost of additional filtering was slight reduction in sharpness of the reconstructed image. However, there was a slight improvement in the PSNR of the reconstructed image using wiener filtering. The de-noised image using Neighshrink sometimes unacceptably blurred and lost some details. This problem will be avoided by reducing the value of threshold itself. So, the shrinkage factor is given by B(i,j) = (1- (3/4)*T2/ S2(i,j))

Step 6: Apply Inverse Wavelet Transform

This is the final step of the proposed method. In this step the inverse wavelet transform is applied and get the denoised image.

IV.EXPERIMENTAL RESULTS

Genetic Algorithm and Thresholding to verify its effectiveness. One is use objective data such as RMS,PSNR,MSSIM to objective analyzed its performance. Experimental results were conducted to denoise a normal image such as cameraman shown in Fig.1 and Fig 2. Speckle, Poisson noise and Salt & Pepper noise were considered. Genetic Algorithm and Thresholding used and their various denoised images is shown in Fig.3 and Fig.4.

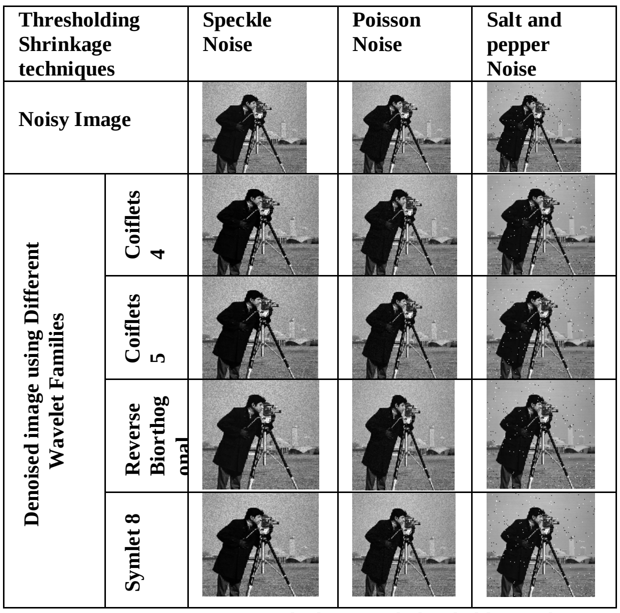


Fig 1 De-noising using wavelet bases with thresholding techniques

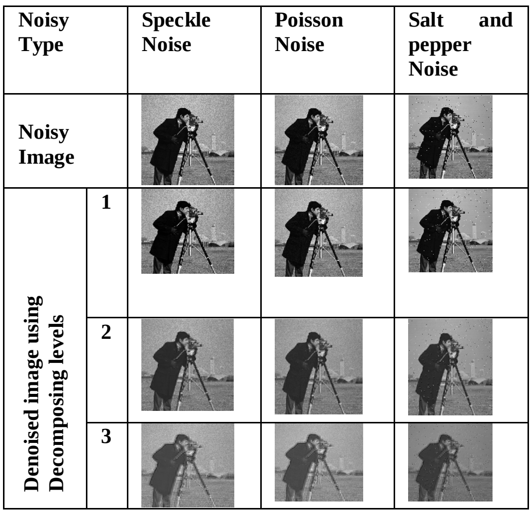


Fig 2Denoising image using Decomposing levels with threshold techniques

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

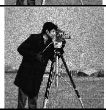
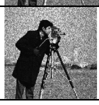
Noise Variance	Noisy Image	
	Speckle Noise	Salt and pepper noise
0.01		
0.02		
0.04		
0.06		

Fig 3 Noisy image with noise variance for speckle and salt and pepper noise

Noise Variance	Noisy Image	
	Speckle Noise	Salt and pepper noise
0.01		
0.02		
0.04		
0.06		

Fig 4 Denoised image with noise variance for speckle and salt and pepper noise

V. PERFORMANCE EVALUATION

PERFORMANCE METRICS

i. Peak Signal-to-Noise-Ratio

The peak signal-to-noise ratio (PSNR) is used to evaluate the quality between the denoised image and the original image. The PSNR formula is defined as follows:

$$\mathbf{PSNR} = 10 \times \log 10 \frac{255 \times 255}{\frac{1}{H \times W} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} [f(x,y) - g(x,y)]^2} \mathbf{dB}$$

where H and W are the height and width of the image, respectively; and f(x,y) and g(x,y) are the grey levels located at coordinate (x,y) of the original image and denoised image, respectively.

To analysis the performance of the three methods by using the performance metrics which are mentioned above. This is shown in the below tables and graphs.

It gives the ratio between possible power of a signal and the power of corrupting noise present in the image.

$$PSNR = 20 \log 10(255/RMSE)$$

Higher the PSNR gives lower the noise in the image i.e.,higher the image quality[9,25].

ii. Root Mean Square Error(RMSE)

Mean square error (MSE) is given by

$$MSE = \sum_{i=j=1}^{N} [(i,) - F(i,j)]^2/N^2$$

Where, f is the original image F is the image denoised with some filter and N is the size of image.

$$RMSE = \sqrt{MSE}$$

iii. Mean Structural Similarity Index Measure(MSSIM)

The Structural Similarity Index between two images is computed as :

$$SSIM(x,y) = (2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)/(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)$$

Where $\mu_x = \sum_{i=1}^{N} w_i x_i$

$\sigma_x = \left(\sum_{i=1}^{N} w_i (x_i - \mu_x)^2\right)^{1/2}$, $\sigma_{xy} = \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)$, $C_1 = (K_1 L)^2, and C_2 = (K_2 L)^2$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Where L is the range of pixel values(255 for 8-bit grayscale images). And K1<<1 is a small constant and also K2<<1

$$MSSIM = \sqrt{SSIM}$$

To analysing the performance by using performance metrics which are shown in the above. The result is shown in below tables.

Table.1.Decomposition levels with poisson Noise

Metrics	Wavelet Decomposition Level	Poisson Noise			
		Visu Shrink	**Neigh Shrink**	**Sure Shrink**	**Modneigh Shrink**
PSN	1	27.4216	27.4337	27.4173	27.3855
	2	22.0409	22.0656	22.0747	22.0615
	3	18.6575	18.6569	18.6484	18.6621
MSS	1	0.64586	0.64767	0.6464	0.64336
	2	0.52498	0.52782	0.52894	0.52467
	3	0.42249	0.42463	0.42552	0.42316
NAE	1	0.068464	0.06852	0.068696	0.06893
	2	0.11241	0.11193	0.11158	0.11237
	3	0.15545	0.15525	0.15534	0.15522
NCC	1	1.0002	1.0002	0.99985	0.99994
	2	0.98849	0.98726	0.98812	0.9878
	3	0.96832	0.96889	0.96853	0.9685

Table.2.Noise variance with speckle noise

Metircs	**Speckle noise**			
Noice Variance	**0.01**	**0.02**	**0.04**	**0.06**
PSNR	25.6569	22.6331	19.6118	17.8568
MSSIM	0.61287	0.53271	0.45561	0.41565
NAE	0.08546	0.12135	0.17191	0.21066
NCC	0.99865	0.99945	0.99898	0.99805

Table.3. Noise variance with salt & pepper noise

Metircs	**Salt & pepper noise**			
Noice Variance	0.01	0.02	0.04	0.06
PSNR	25.4053	22.1805	19.1857	17.1537
MSSIM	0.79437	0.63342	0.46082	0.33628
NAE	0.01005	0.02181	0.04226	0.06653
NCC	0.99806	0.99673	0.99433	0.99045

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

VI. CONCLUSION AND FUTURE WORK

In this thesis, Wavelet based different thresholding techniques are used to enhance the quality of the input natural Image. Mainly in the case of existence of Speckle noise, Salt & Pepper and Poisson Noise, the shrinkage approaches are very much needed with the intention of improving the natural image diagnostic examination. The Wavelet based Thershlding techniques for noise removal gives superior quality of denoising effect with enhanced effect of denoised images. The thersholding technique is applied on every subband of the Wavelet coefficient images for enhancing the denoising performance. Among several shrinkage methods this thesis considers only Visu Shrink, Neigh Shrink, Sure Shrink and Modineighshrink. Experiments were performed to analyse the best suitable shrinkage methods for Wavelet against different noises(Speckle noise, Poisson noise and Salt & Pepper noise). And also this thesis use four different types of Wavelet Families such as COIF4, COIF5, RBio6.8 and Sym8The shrinkage approaches take account of the use of nearest coefficients. So it supply the better worth of denoising images. Performance Metrics such as Peak Signal to Noise Ratio(PSNR), Mean Structural Similarity Index Measure(MSSIM), Normalized Cross Correlation(NCC) , Normalized Absolute Error(NAE) are used to evaluate the denoising effect of output images. It is observed from Wavelet decomposed subband with the help of thersholding approached. From the conducted experimental research one thing is clearly proved that the wavelet shrinkage approach performs well and enhance the denoising performance. In speckle noice wavelet coif4 and threshold neigh shrink gives the best result. In salt and pepper noice wavelet sym8 and threshold visu shrink gives the best result. In poisson noice wavelet sym8 and threshold neigh shrink gives the best result.

REFERENCES

[1]Monika Shukla, Dr. Soni Changlani, Mr. Ayoush Johari " Study of Curvelet and Wavelet Image Denoising by Using Different Threshold Estimators" International Journal of Electrical, Electronics ISSN No. (Online): 2277-2626 and Computer Engineering 3(1): 48-54(2014).

[2] Okuwobi Idowu Paul Yonghua Lu "Image Denoising Using Wavelet Thresholding Techniques" International Journal of Education and Research Vol. 2 No. 2 February 2014.

[3] Monika Raghav, Sahil Raheja "Image Denoising Techniques" International Journal Of Engineering And Computer Science ISSN:2319-724Volume 3 Issue 5 May, 2014 Page No. 5637-5641.

[4] R. Bouchouareb, N. Hedjazi, and D.Benatia "Denoising Medical Ultrasound Images and Error Estimate by Translation-Invariant Wavelets" International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.7, No.1 (2014), pp.193-202 http://dx.doi.org/10.14257/ijsip.2014.7.1.18 .

[5] Mohammad Ali ,Keyvanrad,Mohammad Mehdi Homayounpour"Deep Belief Networks for Image Denoising" arXiv:1312.6158v2 [cs.LG] 2 Jan 2014.

[6] De-An Huang, Li-Wei Kang, Yu-Chiang Frank Wang, Chia-Wen Lin, "Self-Learning Based Image Decomposition With Applications to Single Image Denoising" IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 16, NO. 1, JANUARY 2014 83.

[7] Forest Agostinelli Michael R. Anderson Honglak Lee "Adaptive Multi-Column Deep Neural Networks with Application to Robust Image Denoising"

[8] Sakthivel K."Contourlet Based Image Denoising Using New-Threshold Function" International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014.

[9] Di Guo, Xiaobo Qu, Xiaofeng Du, Keshou Wu, and Xuhui Chen "Salt and Pepper Noise Removal with Noise Detection and a Patch-Based Sparse Representation" Hindawi Publishing Corporation Advances in Multimedia Volume 2014, Article ID 682747, 14 pages.

[10] M. Aldinucci, C. Spampinato, M. Drocco, M. Torquati, S. Palazzo "A Parallel Edge Preserving Algorithm for Salt and Pepper Image Denoising"

[11] Hari Om, Mantosh Biswas "An Improved Image Denoising Method Based on Wavelet Thresholding" Journal of Signal and Information Processing, 2012, 3, 109-116 .

[12] Kamlesh Dixit, Mohit Khandelwal "Digital Image Denoising Using Histogram and Dynamic Filters" InternationalJournal of Computer & Organization Trends – Volume 5 – February 2014

ISSN 2347-7644

Sadakath

A Research Bulletin

UGC APPROVED JOURNAL (NO. 64130) . INDEXED IN GOOGLE SCHOLAR

Volume VI | Number 2 | January 2019

Published by

Sadakathullah Appa College

(Autonomous)

(Reaccredited by NAAC at an 'A' Grade with a CGPA of 3.40 in the III Cycle)

(An ISO 9001:2015 Certified Institution)

Rahmath Nagar, Tirunelveli - 627 011, Tamil Nadu, India.

Phone Number : 0462 - 2540763

Website : www.sadakath.ac.in

Cardiac MRI Segmentation Techniques – An Overview

Gomathi, G.,[1] *and Subha, V.*[2]

Abstract

Segmentation of LV, RV in Cardiac MRI plays a paramount role in medical imaging. Cardiovascular diseases are the main cause of death in present century. The study of cardiac motion is one of the important subjects in biomedicine field. The right ventricle (RV) and left ventricle (LV) are the two lower chambers of the heart that receive blood from the two upper chambers of the heart and pump it into the arteries by contraction or tightening of the chamber walls .Cardiac MRI provide the flexible and accurate information on morphology, tissue viability and blood flow. Manual segmentation is hectic and time-consuming job for radiologist and cardiologist. The paper describes an automated framework for LV, RV segmentation task. Assorted methods are available for segmentation of heart chambers.Cardiac MRI segmentation techniques are classified into three main categories. That are Threshold based, Region Based and Edge Based. Segmentation accuracy encompasses the wealth of computerized analysis. In the splitting of an image into meaningful structures, image segmentation, is often an essential step in image analysis, object representation, visualization, and many other image processing tasks

Keywords: *Cardiac MRI, Left and right ventricle segmentation.*

Introduction:

Magnetic resonance imaging (MRI), computed tomography (CT), digital mammography, and other imaging modalities provide an effective means for noninvasively mapping the anatomy of a subject. These technologies have greatly increased knowledge of normal and diseased anatomy for medical research and are a critical component in diagnosis and treatment planning. Magnetic Resonance Imaging (MRI) is a test that has been useful for decades in diagnosing problems of the brain, spine, joints, and other stationary organs. Accurate segmentation is the important step in the workflow of cardiac function evaluation process. Ejection fraction (EF), left ventricle volume, and muscle wall thickness are some of the parameters need to be measured precisely to evaluate the cardiac function. Calculation of

1 Research Scholar, Department of CSE, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India.
2 Assistant Professor, Department of CSE, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India.

these parameters requires an accurate delineation of muscle and ventricle contours. Manual heart contouring by expertise is one solution to this problem but it is an extremely tedious task, time consuming, and highly affected by the observe variability.

Therefore fully automatic segmentation is an important Cardiac MRI (CMR) that offers greater contrast and image clarity than CT and does not require use of a contrast agent and allows radiation-free perfusion imaging.

Cardiac Ventricle Segmentation:

The status of cardiac physiology in health and disease is quantified with multiple important indicators of cardiac function, including global performance indexes (e.g., wall mass, end-systolic and end-diastolic volumes, and ejection fraction), and local function indexes (e.g., wall thickness and thickening) [1]. These characteristics help clinicians to detect myocardial dysfunction, specifically major adult heart conditions leading to heart attack or failure, and to optimize therapies for individual patients. Basic ventriculometrics require accurate segmentation of myocardial borders on cardiac images. Segmentation of the left ventricle (LV) in cardiac magnetic resonance (CMR) images for 2-D, 3-D, or 4-D analysis has been long explored. However, it still remains a challenge due to the difficulties of delineating the inner border from the cavity, protrusions of papillary muscle structures into the cavity, partial influence of adjacent structures, such as the diaphragm, and artifacts from blood moving within the ventricular cavity [2]. Thus, the improvement of popular existing segmentation methods (e.g., deformable models [3]) for efficient, robust, and accurate assessment of cardiac images remains an area of considerable interest.

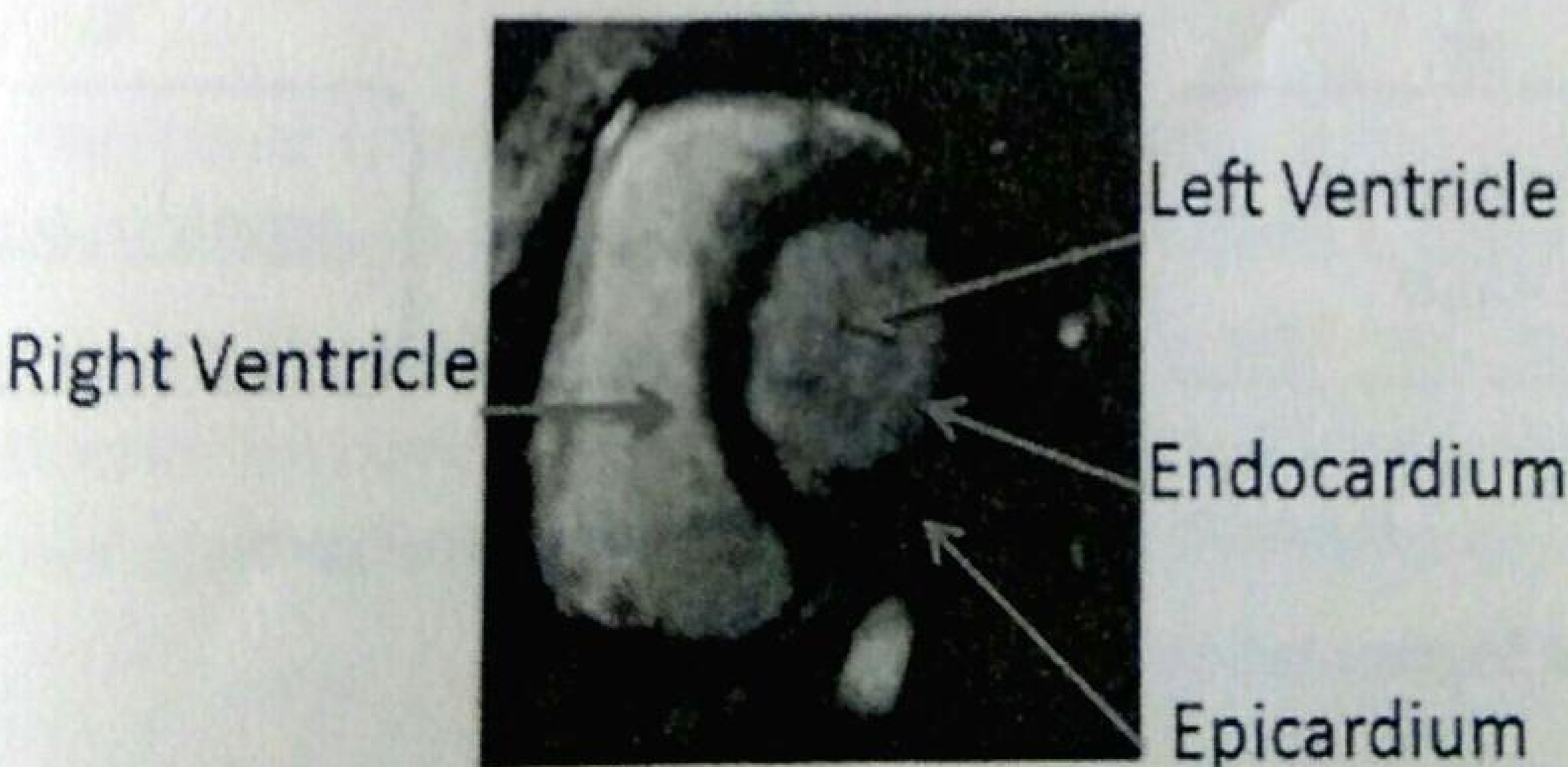


Figure 1: Cardiac Left and Right Ventricle

The main goal of image segmentation is domain independent partitioning of an image into a set of disjoint regions that are visually different and has meaning with respect to some characteristics or computed property, such as grey level, texture or color to enable easy image analysis. Image segmentation has played an important role in Computer vision it is used for object tracking and to identify image boundaries. The different algorithms used in Image segmentation are Region-based, Edge-based and Threshold-based.

Cardiac MRI Segmentation Techniques:

Cardiac MRI imaging sequences is considered an important tool that is used for evaluating cardiac function. Evaluation of cardiac function requires calculation of different cardiac parameters (i.e. ejection fraction (EF), left ventricle mass (LVM), left ventricle volume, wall thickness, or wall thickening). All of these parameters depend on segmenting the endocardial, and epicardial contours of the left ventricle from the image sequences that are acquired from cardiac imaging technique. The manual segmentation of these contours from all dataset (i.e. all time frames per all slices) takes a lot of time and effort from the cardiologist. Therefore, the automatic segmentation is very important, and can be considered as a challenging task.

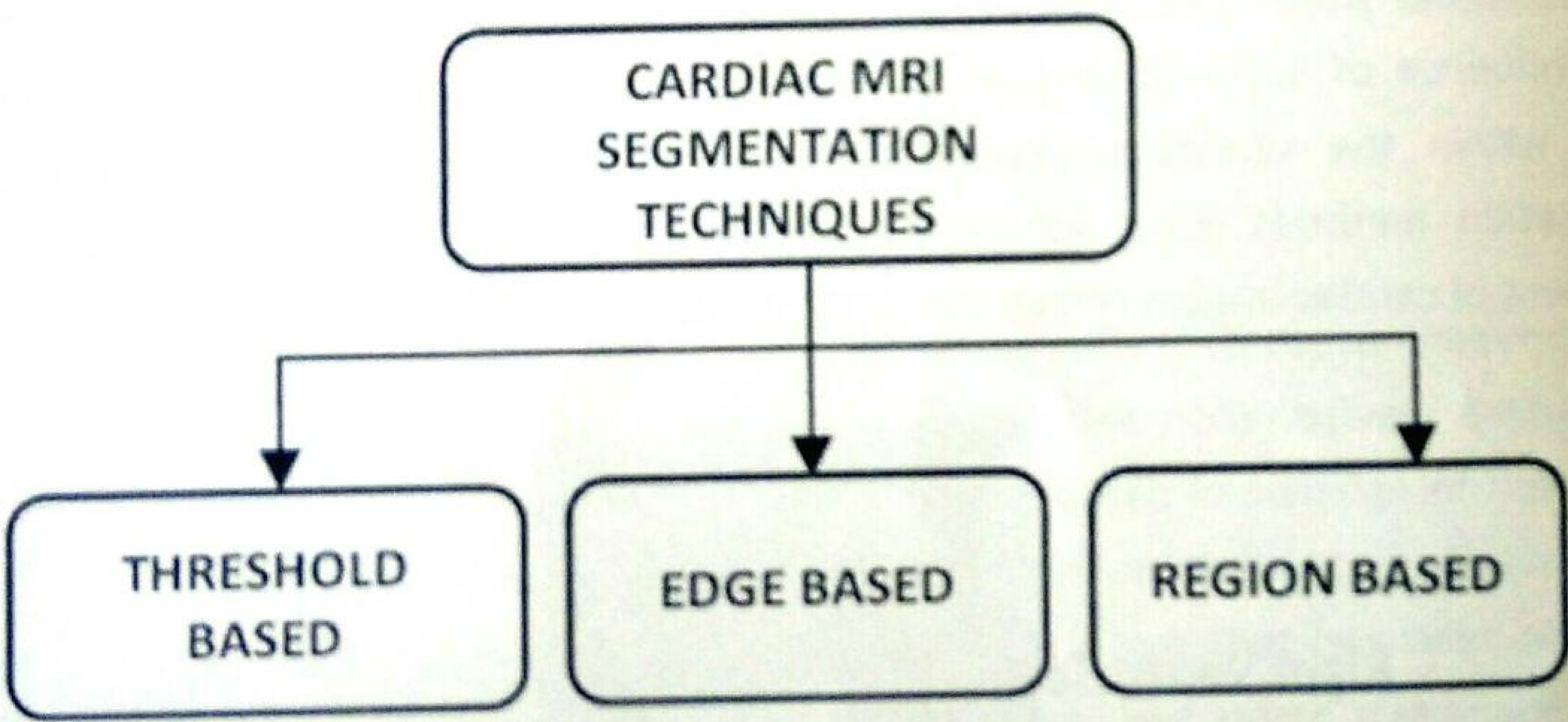


Figure 2: Cardiac MRI Segmentation Types

Thresholding Based Segmentation:

Image thresholding is a simple, yet effective, way of partitioning an image into a foreground and background. This image analysis is a type of mage segmentation that isolates objects by converting greyscale images into binary images. Thresholding techniques are used in the case where pixel within segments has similar intensities. It is one of the popular techniques as it is very simple to implement. Threshold-based techniques are generally used for gray-scaled images. Here the intensity of each pixel is compared with a threshold value. For gray scale image f(x,y), consider the image is divided into two parts: background and foreground. The foreground is defined as the region of interest and the background as the rest. Threshold value T is first calculated by analyzing all image pixels intensity.

Any pixel (x,y) for which $f(x,y) > T$ is called object point, otherwise, that point is called background point. Thus, intensity level is compared to the background image and a threshold value decides if the pixel differs enough to belong to the foreground or not.

There are three types of thresholding algorithms.

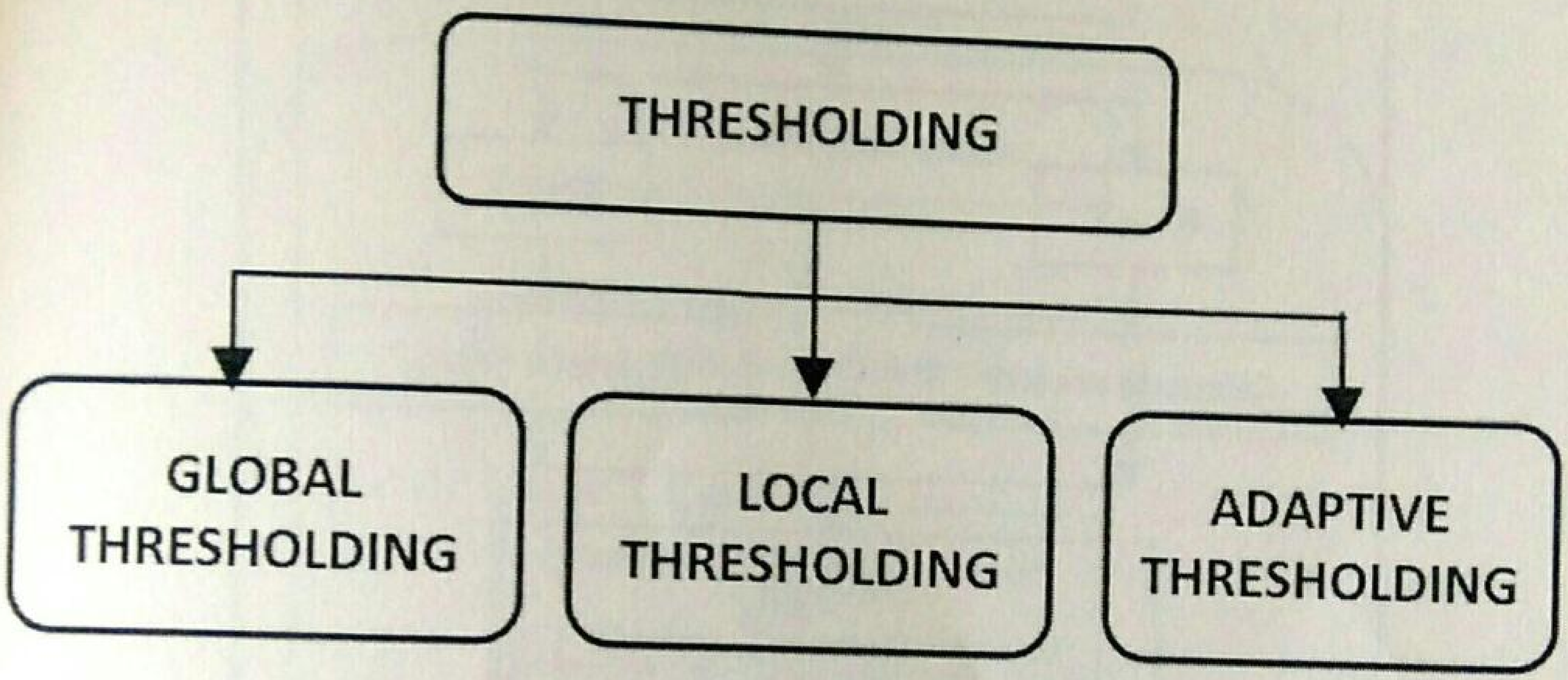


Figure 3: Types of Thresholding

In adaptive thresholding, different threshold values for different local areas are used.

Global Thresholding:

The global threshold is applicable when the intensity distribution of objects and background pixels are sufficiently distinct. In the global threshold, a single threshold value is used in the whole image. The global threshold has been a popular technique for many years. When the pixel values of the components and that of background are fairly consistent in their respective values over the entire image, global thresholding could be used.

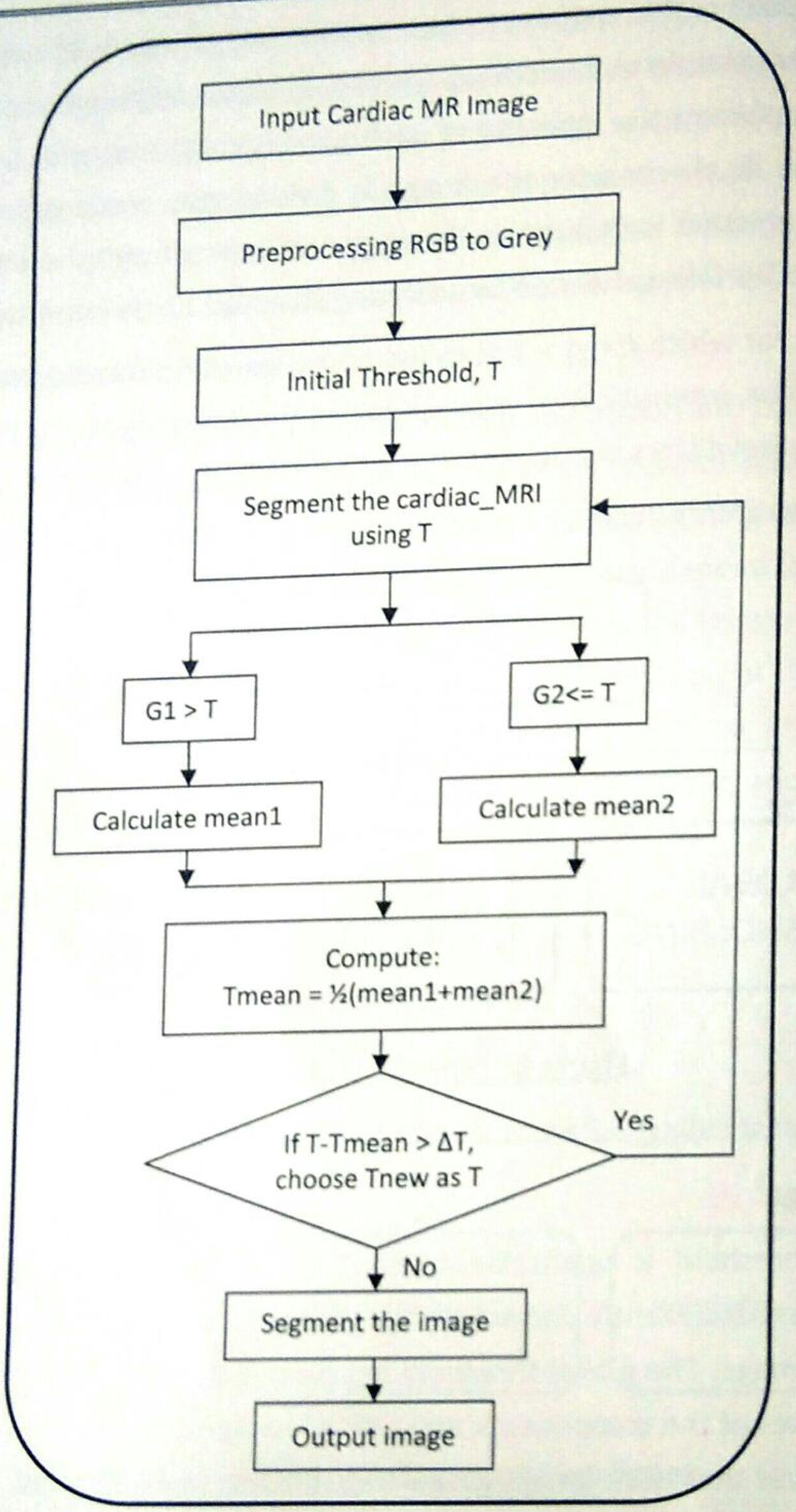


Figure 4: Global Thresholding Flow Chart.

Algorithm:

Step 1: Select an initial estimate for T.

Step 2: Segment the image using T. This will produce two groups of pixels. G1 consisting of all pixels with grey level values >T and G2 consisting of pixels with values <=T.

Step 3: Compute the average grey level values mean1 and mean2 for the pixels in regions G1 and G2.

Step 4: Compute a new threshold value T=(1/2)(mean1 +mean2)

Step 5: Repeat steps 2 through 4 until difference in T in successive iterations is smaller than a predefined parameter T0, i.e., if |T-Tnew| > ΔT, back to step 2, otherwise stop

Local Thresholding:

Local thresholding is an image thresholding technique that segments an image on the basis of local threshold, i.e. it compares threshold for every pixel in an image. The main idea is that each pixel is compared to an average of the surrounding pixels. Specifically, an approximate average of the last s × s window of pixels centered around each pixel is calculated while throughout the image which means, it considers neighbouring pixels on all sides in the region. If the value of the current pixel is t percent lower than the average then it is set to black, otherwise it is set to white

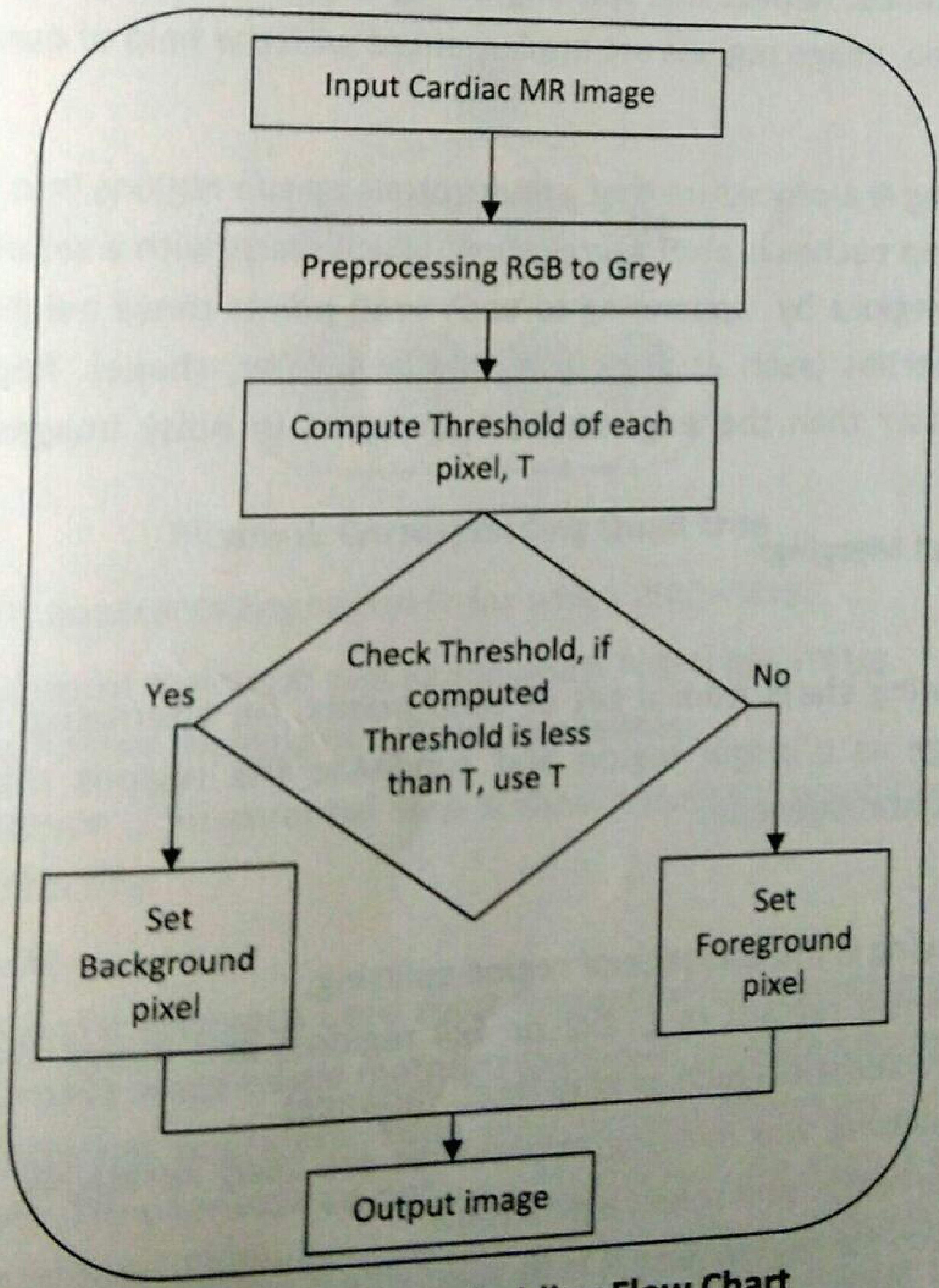


Figure 5: Local Thresholding Flow Chart

Region Based Technique:

Region-based method can be classified into two: Region growing and Region splitting-merging procedures. In region growing procedure it groups pixels or sub-regions into large regions based on certain predefined criteria. Initially set of seed points are created and from this point remaining regions grow up if neighbouring pixels have similar properties of that of `seed point. Selection of seed points is critical procedure for coloured images if prior information is not available. Hence set of descriptors based on intensity levels and spatial properties are required to determine the pixel intensities. In region splitting-merging, an image is subdivided into arbitrary, disjoint regions and then either merge or split operation is performed to satisfy prerequisite constraints. This algorithm is iterative. First, split the given image into four disjoint quadrants, then merge any adjacent regions which satisfy the prerequisite constrained. Repeat this splitting of regions and merging till no further merging or splitting is possible. Image regions are implemented with the help of quad tree.

Region Growing:

Region growing is a procedure that groups pixels or sub regions into larger regions. The simplest of these approaches is pixel aggregation, which starts with a set of "seed" points and from these grows regions by appending to each seed points those neighbouring pixels that have similar properties (such as gray level, texture, color, shape). Region-growing-based techniques are better than the edge-based techniques in noisy images where edges are difficult to detect.

Region Splitting and Merging:

Region Splitting

- Region growing starts from a set of seed points. An alternative is to start with the whole image as a single region and subdivide the regions that do not satisfy a condition of homogeneity.

Region Merging

- Region merging is the opposite of region splitting.
- Start with small regions (e.g. 2x2 or 4x4 regions) and merge the regions that have similar characteristics (such as gray level, variance).
- Typically, splitting and merging approaches are used iteratively. Let R represent the entire image region and select a predicate.
- One approach for segmenting R is to subdivide it successively into smaller and smaller quadrant regions so that, for Ri, P (Ri) = TRUE.

- If P(R) FALSE divide the image into quadrants.
- If P is FALSE for any quadrant, subdivide that, quadrants and so on.
- This particular splitting technique has a convenient representation in the form called quad tree.

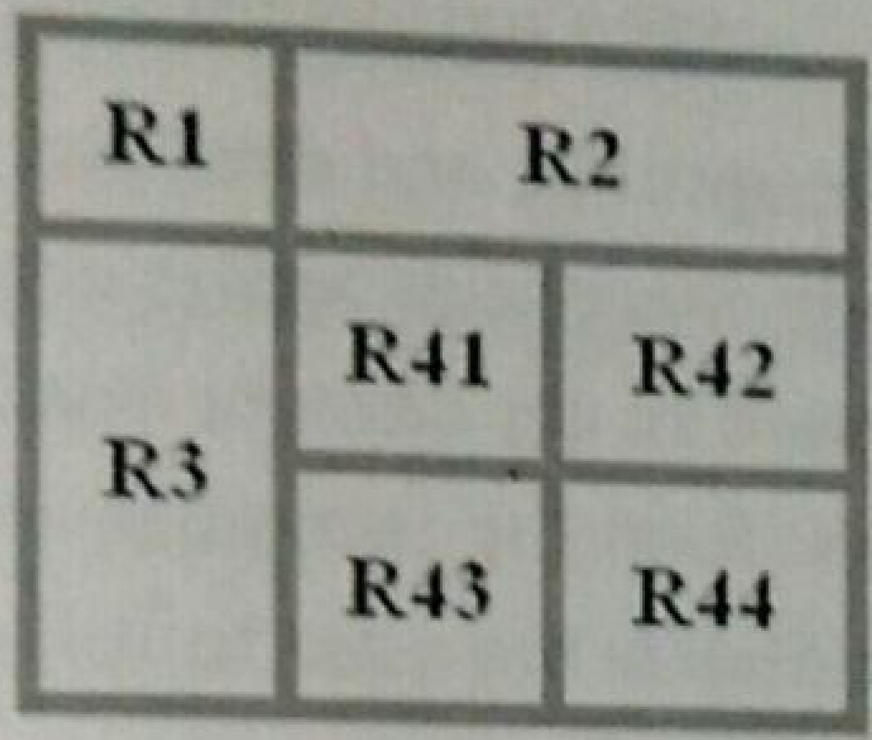


Figure 5: Partitioned Image

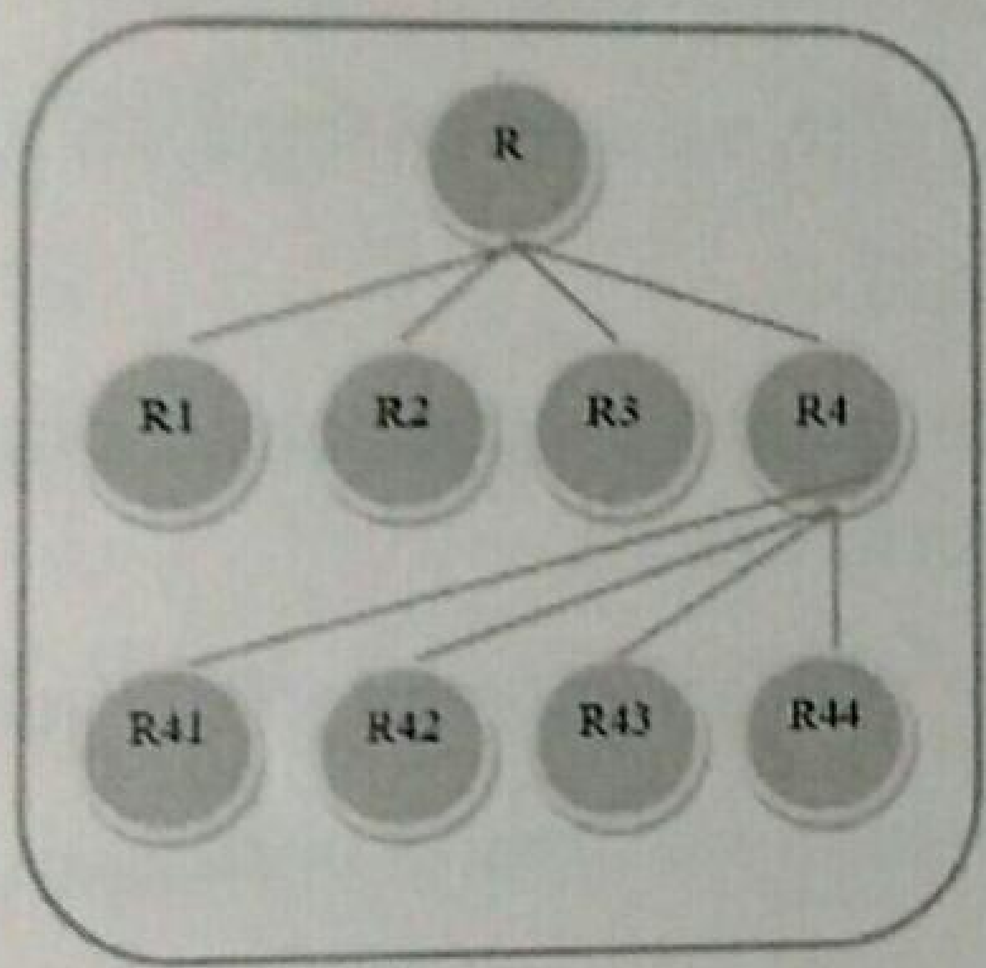


Figure 6: Corresponding Quad Tree

Split into four disjoint quadrants any region Ri for which P(Ri)=FALSE.

- Merge any adjacent regions Rj and Rk for which P(Rj U Rk) =TRUE.
- Stop when no further merging or splitting is possible.

Image segmentation is an essential step in most automatic graphic pattern recognition and scene analysis problems.

Edge-Based Technique:

Edge detection is the approach used most frequently for segmenting images based on abrupt changes in intensity. Edge based methods are based on edge information found in the image using edge detection operators. Edge plays a very important role in many image processing applications. They provide an outline of the object. An edge is a set of connected pixels that lies on the boundary between two regions that differ in grey value. These pixels on the edge are called edge points. There are four types of edge model that are classified according to their intensity profile.

Canny Edge Detection:

Canny edge operator is a second order derivative operator for edge detection and is considered as superior edge detection operator among the available operators based on the experimental results as it determines strong and weak edges in the image. Image is first smoothed by using circular two-dimension Gaussian function, computing the gradient of the result and then using the gradient magnitude and direction to approximate edge strength and direction at every point. The gradient magnitude arrays consists of undesirable ridges around local maxima and are to be suppressed to get discrete orientations of the edge normal by the process of non maxima suppression. Then the technique of double thresholding is employed to reduce false fragments. Two thresholds are used to solve the purpose T1 and T2 where T2 ≈ 2T1.

Algorithm:

Step1: Convolve image f(r, c) with a Gaussian function to get smooth image

f ^ (r, c). f ^ (r, c)=f(r,c)*G(r,c,6)

Step2: Apply first difference gradient operator to compute edge strength then edge magnitude and direction are obtain as before.

Step3: Apply non-maximal or critical suppression to the gradient magnitude.

Step4: Apply threshold to the non-maximal suppression

Sobel Edge Detector:

This is widely used first derivative operator to find edges and is modification of Prewitt"s operator, as will be discussed next. It changes the center coefficient by '2'.

The Sobel operatiors are given as

-1	-2	-1
0	0	0
+1	+2	+1

Gx

-1	0	-1
-2	0	+2
-1	0	+1

Gy

Algorithm:

Step1: Apply mask Gx, Gy to the input image.

Step2: Apply Sobel edge detection algorithm and the gradient.

Step3: Masks manipulation of Gx, Gy separately on the input image. Step4: Results combined to find the absolute magnitude of the gradient.

$$|G| = \sqrt{G_x^2 + G_y^2}$$

Step5: the absolute magnitude is the output edges

Prewitt Edge Detector:

This operator uses 3 X 3 mask to find the edges. The mask used along x and y direction corresponding they are:

-1	-1	-1
0	0	0
+1	+1	+1

Gx

-1	0	+1
-1	0	+1
-1	0	+1

Gy

Prewitt edge detector adequately works much better in the same way as Sobel operator

Conclusion:

Segmentation of cardiac image is imperative in surgical planning and treatment planning in the field of medicine. . In this study, the efficient techniques to segment the inner border, outer border of the LV, RV are analyzed. According to those methods the edges are detected..Fully automated segmentation technique is fast, accurate, robust and beneficial for quantification of Cardiac MRI in clinical practice. As the result, image segmentation is affected by lots of factors, such as homogeneity of images, spatial characteristics of the image continuity, texture and image content. As the result, image segmentation is affected by lots of factors, such as homogeneity of images, spatial characteristics of the image continuity, texture and image content.

References:

[1]. Amandeep Kaur, Rakesh Kumar, Kamaljeet Kainth, *Review Paper on Image Segmentation Techniques*, IJARCSSE, Volume 6, Issue 7, July 2016.

[2]. Apurv Vashisht, Shiv Kumar, *Analysis of Image Segmentation Techniques: A Survey*, IJETR- ISSN: 2321-0869 (O) 2454-4698 (P), Volume-4, Issue-1, January 2016

[3]. Basavaraj D., Dr. Govardhan A. and Dr. Premchand P., *"Active contours and Image segmentation: The current state of the art"*, Global Journal of computer science and technology graphics and vision, Vol. 12, Iss. 11, Ver 1.0, 2012.

[4]. Bhargavi K. and Jyothi S., *"A survey on threshold based segmentation technique in image processing"*, IJIRD: Vol. 03, Iss. 12, Nov 2014.

[5]. Fahmi Khalifa, Garth M. Beache, Georgy Gmel'farb, Guru Prasad A. Giridharan and Ayman El-Baz, *"Accurate automatc analysis of cardiac cine images"*, IEEE Transactions on biomedical engg, Vol. 59, Iss. 02, Feb 2012.

[6]. Faliu Y and Inkyu Moon, *"Image segmentation: A survey of graph-cut methods"*, ICSA 2012.

[7]. Grosgeorge D., Petitjean C., Dacher J.N. and Ruan S., *"Graph cut segmentation with a statistical shape model in cardiac MRI"*, Computer vision and image understanding 117, 2013. (1027-1035).

[8]. Heldah Paul R. and Jeeva J.B., *"Left Ventricle segmentation from cardac cine MRI to detect cardiac abnormalities"*, International conference on communication and signal processing, April 3-5 2014, India.

[9]. Jothilakshmi G. R., Sharmila P. and Arun Raaza, *Mammogram Segmentation using Region based Method with Split and Merge Technique*, Indian Journal of Science and Technology, Vol 9(40), DOI: 10.17485/ijst/2016/v9i40/99589, October 2016

[10]. Michael R. Avendi, Arash Kheradvar and Hamid Jafarkhani, *"Automatic segmentation of the right ventricle from cardiac MR using a learning based approach"*, Magnetic Resonance n Medicine 00:00-00, 2017.

[11]. Mohamad Mujtaba Lone and Saddam Hussain, *"A survey on image segmentation techniques in digital image processing"*, IRJET: Vol. 05, Iss. 09, Sep 2018.

[12]. Princy Mishra, Shikha Agarwal and Usha Kiran, *Survey Paper Based on Medical Image Segmentation and Clustering Techniques*, IJARCET, Volume 2, Issue 12, December 2013

[13]. Roy Jackson Monteiro, Dhanush J.K and Nausheeda B.S, *"Comparison of various segmentation algorithms in image processing"*, JLTET: SACAIM 2016, pg. 241-247.

[14]. Shameena N, Rahna Jabbar, *A Study of Preprocessing and Segmentation Techniques on Cardiac Medical Images*, IJERT, Vol. 3 Issue 4, April – 2014.

[15]. Shilpa Kamdi and Krishna R.K., *Image Segmentation and Region Growing Algorithm*, IJCTEE - Volume 2, Issue 1, ISSN 2249-6343.

[16]. Sneha Kumari, *Detail and Comparative Study on Various Segmentation Techniques*, IJCSMC, Vol. 6, Issue. 5, May 2017, pg.308 – 315.

DESIGN ENGINEERING

Semantic Segmentation of Ventricular and Myocardium Regions in Cardiac MRI

G. Gomathi, Dr. V.Subha

Keywords: Deep Learning, Unet, Segnet, FCN (Fully ConvolutionalNetwork), Cardiac MRI(Magnetic Resonance Imaging).

ABSTRACT

Deep Learning has been most widely used in Cardiac MRI segmentation as it comprises of stack of layers. Manual demarcation is a time consuming and tiresome operation whereas automated process makes it easier and significant to identify the clinical parameters of Cardiac MRI. Over the last few decades, extensive study has been conducted on automation of the necessities. This study aims to compare the three deep learning subsets Unet, Segnet, and Fully ConvolutionalNetwork (FCN) that are being implemented to segment the ventricle regions Left Ventricle (LV), Right Ventricle (RV) and Myocardium (MYO) in systolic and diastolic phases of cardiac cycle. The performance analysis is done based on the values obtained using Dice Coefficient and Haudorff Distance. While comparing the three methods, Unetarchitecture provides the best outcome. The results are affirmative to the statement-'Unet is the multiclass segmentor'.

PDF

HOW TO CITE

Dr. V.Subha, G. G. (2021). Semantic Segmentation of Ventricular and Myocardium Regions in Cardiac MRI. *Design Engineering*, 8510-8522. Retrieved from http://www.thedesignengineering.com/index.php/DE/article/view/7957

More Citation Formats

ISSUE

Vol 2021: Issue 09

SECTION

Articles

MAKE A SUBMISSION

CONTACT US

Editorial Office of Design Engineering.
Address: 4143 Danforth Avenue Toronto, ON M4K 1A6.
Email: editor@thedesignengineering.com

Downloads

Paper Template Download

Information and Guidelines

- Author Guidelines
- Competing Interest Statement
- Copyright Notice
- Publication and Peer Review Processes
- Published Statement of Human and Animal Rights guidelines

- Published Statement of Informed Consent

Subscribe

Journal print copy or article reprints are available for order, please contact: editor@thedesignengineering.com

Online Access
This is a fully open access journal, the full texts (in HTML and PDF) of all articles can be viewed online for free immediately after publishing.

Permission
For permission, please contact the editorial office directly:
Email: editor@thedesignengineering.com

SUBSCRIPTION

Login to access subscriber-only resources.

INFORMATION

For Readers

For Authors

For Librarians