CHAPTER 3

ISCHEMIC STROKE

3.1 INTRODUCTION

Less attention is paid towards detection of ischemic stroke with small lesions (lacunar stroke) due to its challenging nature. It is relatively difficult to identify, as it manifests as a small hypodense area of less than 15mm in diameter on CT. Therefore, early detection of ischemic stroke with small lesions is important and this necessitates a more efficient method to improve the detection rate. In this chapter, different techniques used for ischemic stroke detection and classification are discussed and the quantitative analysis is described along with the results of segmentation for diagnosing ischemic stroke.

3.2 PROPOSED APPROACH FOR ISCHEMIC STROKE

Developing an efficient method may help physicians to diagnose ischemic stroke at an appropriate time. The proposed system has been developed using Matlab (The MathWorks, Inc., Natick, MA, USA). Considering CT images as input data, the proposed method has five stages, pre-processing, tracing midline of the brain, extraction of texture features using GLCM and feature selection, classification and segmentation.

In the first stage noise is suppressed using a median filtering and skull bone components of the images are removed by a global thresholding method. In
the second stage, midline shift of the brain is calculated. In the third stage, fourteen texture features are extracted using GLCM. The optimal features are extracted from the left and right side of the brain using GA. The optimal features are used to train the binary classifier, which can automatically infer whether the image is that of a normal brain or an ischemic brain, suffering from a brain lesion. The methodology of the proposed technique for ischemic stroke detection and classification is illustrated in Figure 3.1.

Figure 3.1: Methodology of the proposed technique
In the fourth stage, SVM, ANN and decision tree classifiers have been used to classify normal brain and ischemic brain. Finally, the ischemic stroke region is extracted by using k-means clustering technique.

The research protocol is approved by the Human Ethics Committee of the Rajah Muthiah Medical College, Annamalai University and informed consent is signed by all participants. The dataset of brain CT images is collected from Rajah Muthaiyah Medical College Hospital (RMMCH) in Chidambaram to develop the algorithm. All brain CT scans are performed using a standard protocol according to the guideline, with some variations in slice spacing [Bull2006]. The Field of View (FOV) is 25x25 cm. Each image is 512x512 pixels in size, resulting in an in-plane resolution of 0.488mm per pixel. The original CT number (Hounsfield Unit, HU) is transformed with brain window (center 40HU, width 150HU) into 256 gray levels.

Consecutive Digital Imaging and Communications in Medicine (DICOM) CT slices are imported into the Matlab workspace automatically. The proposed system is experimented and tested using 21 cases. The thickness of brain CT images is 5 mm. All are emergency brain CT scans performed on a single detector CT scanner (Mx 8000 Dual, Philips). All images are axial images obtained parallel to the orbitomeatal line, at 120kV and 80–200mA. A region of ischemic stroke identified by an experienced radiologist is used as gold standard.
3.2.1 PRE-PROCESSING FOR ISCHEMIC STROKE

To detect the abnormality in the brain, the unwanted background information has to be removed. To enhance the accuracy of segmentation and to save computational time, it is useful to eliminate the artifacts that might be present in the image. The following are the examples of the artifacts commonly seen in CT images.

1. Streek artifacts are usually produced by metal and patient motion.
2. A partial volume artifact occurs when a volumetric pixel contains two very different materials like bone and soft tissue.

3.2.1.1 Noise Removal

To reduce noise, median filtering using a 3-by-3 square kernel is applied. Median filter is chosen because it is less sensitive to extreme values and able to remove outliers without reducing sharpness of the image. This produces a more homogeneous background in which abnormalities become more conspicuous. Median filters are quite popular because they provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. The median filter considers each pixel and its neighbors in the image to decide whether or not it is a representation of the surroundings. It replaces the pixel value with the median of the neighboring pixel values. Calculate the median by first sorting all the pixel values from the neighborhood in numerical order and then replace the pixel being considered with the middle pixel value. (If
the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is the median.) It preserves sharp edges of the image.

For a subset of pixel values of the image how median value calculated is given below. The window can be of any size. Here the window size is 3x3.

\[
\begin{bmatrix}
5 & 6 & 9 \\
2 & 5 & 3 \\
8 & 1 & 2 \\
\end{bmatrix}
\]

Let the input matrix be \( A = \begin{bmatrix} 2 & 5 & 3 \\ 8 & 1 & 2 \end{bmatrix} \)

Now pad the matrix with zeros on all sides.

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 5 & 6 & 9 & 0 \\
0 & 2 & 5 & 3 & 0 \\
0 & 8 & 1 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

Starting from matrix \( A(1,1) \) place the window

\[
\begin{bmatrix}
0 & 0 & 0 \\
\end{bmatrix}
\]

Window = \( A = \begin{bmatrix} 0 & 5 & 6 \\ 0 & 2 & 5 \end{bmatrix} \)

Sort the window matrix.

\[
\begin{bmatrix}
0 & 0 & 0 \\
\end{bmatrix}
\]

\( A = \begin{bmatrix} 0 & 0 & 2 \\ 5 & 5 & 6 \end{bmatrix} \)

After sorting, the matrix is placed with a value of zero at (2, 2) pixel position. The value of the output pixel is found using the median of the neighboring pixels. This procedure is repeated for all the values in the input matrix by sliding the window to the next position i.e. \( A(1, 2) \) and so on.
Output matrix \( A = \begin{bmatrix} 0 & 3 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 0 \end{bmatrix} \)

The input images are shown in Figure 3.2(a-c) and the noise removed images are presented in Figure 3.3(a-c).

![Figure 3.2: Input CT brain images (a-c)](image)

![Figure 3.3: Images after median filtering (a-c)](image)

### 3.2.1.2 Global Thresholding

The bony skull has to be removed with preservation of cranial content of the brain. Skull, by virtue of its exceptionally high attenuation values, is first removed using global thresholding with a threshold value of 110. Intracranial contents, including both the brain and CSF containing spaces, are first segmented. There are two simple anatomical facts to pick out the brain and the
skull in the given CT study: (a) the skull is the largest connected region with bone density in the whole imaging volume; (b) the brain is the largest connected region with brain density within the skull. The CT image is converted to binary image with pixel value (0, 1) using a global threshold. The bony skull of the image is removed with reference to the binary image. Thus in order to detect the abnormality in the brain, the bony skull has to be removed as shown in Figure 3.4(a-c).

![Figure 3.4: Images after skull bone removal (a-c)](image)

level = graythresh(I)  \hspace{1cm} (3.1)

eq. 3.1 computes a global threshold (level) that can be used to convert an intensity image to a binary image. Level is a normalized intensity value that lies in the range [0, 1]. The graythresh function uses the Otsu's method, which chooses the threshold to minimize the intraclass variance of the black and white pixels [Nnbu1979].
3.2.2 MIDLINE OF THE BRAIN

The tracing midline of the brain is a good indicator for measuring the symmetry of the brain. The classification of the images into normal and abnormal depends on the features of left and right side of the image. So in the case of asymmetric brain the left and right features should be extracted accurately to classify it correctly. For automatic detection, the following method is applied for tracing the midline of the brain. The human head is roughly bilaterally symmetric. For many focal brain diseases manifested as intracranial mass clinicians rely on a midline shift to quantify the change of symmetry for diagnosis and outcome prediction. The remarkable feature of a normal human brain is the symmetry. Asymmetry in an axial CT brain image strongly indicates abnormality. Hence the symmetry in non-enhanced CT images are an important feature that needs to be considered in deciding whether the CT image at hand is of a normal or an abnormal brain.

Midline shift is the most important quantitative measure clinicians use to evaluate the severity of brain compression by various pathologies. The deformed midline model used is according to the biomechanical properties of different types of intracranial tissues [Liao2006a, Liao2006b]. There is functional difference between the hemispheres of the brain. Both cerebrum and cerebellum are symmetric with lobes, ventricles and deep nuclei of similar size and shape in both hemispheres. Therefore, they rely on MLS to quantify the change of symmetry for diagnosis and outcome prediction. The most commonly used
landmarks for MLS measurement is septum pellucidum, the structure between both frontal horns of the lateral ventricles. In general, the anatomical midline could only be judged by human eyes. For automatic detection, the following method is used for tracing the midline of the brain.

3.2.2.1 Symmetry Map

The abnormality mostly occurs only on one side of the brain. The difference along the horizontal direction is considered because of bilateral symmetry of the brain and the head. The symmetry metric in an image of a point is quantized by measuring the intensities of points on both sides of it. Typically, the diameter of the brain is about 300 pixels. Weighted sum of squared difference of the 48 pixels on each side is calculated for each brain pixel value. This corresponds to about 2.4 cm laterally and usually covers anatomical structures around the midline, such as frontal horns of the lateral ventricles and the basal ganglia. Let \( n \) denote the number of pixels on each side to be considered near a point \( (i, j) \), then the symmetric metric at that point is calculated using eq. (3.2) [Liao2006a]. In this calculation \( n \) is assigned to be 48.

\[
S_{i,j} = \sum_{k=1}^{n-1} ((P_{i+k,j} - P_{i-k,j})^2 \ast (n - k) / n); \ k = 1 \ to \ 47
\]

where \( P_{i,j} \) is the intensity of a point \( (i, j) \) of the image and \( S \) forms a symmetry map that can be used to trace the deformed midline. Based on the regional property, upper and lower points of the midline of the brain image are obtained. A midpoint is calculated from upper and lower points. Using these three control
points (N=3) [Liao2006a] deformed midline is drawn using Bezier curves. The
symmetry map of the image is shown in Figure 3.5(a-c).

![Figure 3.5: Generation of the symmetry map (a-c)](image)

3.2.2.2 Bezier curve

The upper and lower control points represent the parts of tough meninges
separating the two hemispheres; it can be treated as straight lines and the curved
segment in between are fitted by a quadratic Bezier curve representing the
intervening soft brain tissue. This quadratic Bezier curve is drawn using a De
Casteljau method [Sede2011] for N=3 using eq. (3.3). The label of the control
point is represented by P. τ is the slope of the line. The deformed midline
indicated by dashed line is shown in Figure 3.6(a-c).

\[
P_{i+1}^j = (1 - \tau)P_{i}^{j-1} + \tau P_{i+1}^{j-1}; \quad j = 1, ..., n; \quad i = 0, ..., n - 1
\]  

(3.3)

![Figure 3.6: Generation of Bezier curve (a-c)](image)
3.2.3 TEXTURE FEATURE EXTRACTION AND SELECTION

Feature extraction is the process of extracting certain characteristic attributes and generating a set of meaningful descriptors from an image. The purpose of the feature extraction component in a computer-aided diagnosis system of ischemic stroke is to extract various features from a given brain image which best characterizes a given lesion as normal or abnormal.

The types of features are colour, texture and shape feature. With respect to medical imaging, colour features are mostly inapplicable. However, in the medical field, the majority of images are gray-level images. Lesions in medical images do not take any particular shape and hence shape feature is not suitable for classification. Medical images are often highly textured and texture analysis becomes crucial in medical image retrieval. Many images may present an irregularity in a local region, but actually display some kinds of regularity in whole, which is usually called texture.

Methods of texture analysis are usually divided into two major categories. The first is the structural approach, where texture is considered as a repetition of some primitives, with a specific rule of placement. The traditional fourier spectrum analysis and wavelet based analysis are often used to determine the primitives and the placement rule. The second major approach in texture analysis is statistical method. Its aim is to characterize the stochastic properties of the spatial distribution of gray levels in an image. The statistical approach to texture
analysis is more useful than the structural one. Texture based analysis has been very successful in the diagnosis of medical images.

Texture is a repeating pattern of local variations in image intensity. The co-occurrence matrix is a statistical method used for texture analysis. As the name suggests, the co-occurrence matrix is constructed from the image by estimating the pairwise statistics of pixel intensity. The use of the co-occurrence matrix is based on the hypotheses that the same grey-level configuration is repeated in a texture. This pattern will vary more by fine textures than by coarse textures. The co-occurrence matrix $P(i, j|d, \theta)$ counts the co-occurrence of pixels with grey values $i$ and $j$ at a given distance $d$ and in a given direction $\theta$. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics.

The GLCM method is a way of extracting second order statistical texture features [Hara1973]. However, the performance of a given GLCM based feature, as well as the ranking of the texture features, depends on the number of gray levels used. The following notations are used: $\mu$ is the mean value of $P$, $\mu_x$, $\mu_y$, $\sigma_x$ and $\sigma_y$ are the means and standard deviations of $P_x$ and $P_y$. $G$ is the size of the co-occurrence matrix. Here the number of rows and columns of the co-occurrence matrix is equal. The following GLCM features are extracted in this research work: ASM, contrast, entropy, correlation, sum of squares, difference entropy,
IDM, inertia, cluster prominence, cluster shade, energy, homogeneity, dissimilarity and difference in variance. They are defined in eqs. (3.4)-(3.17).

1. Angular Second Moment

\[
ASM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{P(i,j)\}^2
\]  

ASM is a measure of homogeneity of the image. A homogeneous image will contain only a few gray levels, GLCM gives only a few but relatively high values of \(P(i,j)\). Thus, the sum of squares also will be high.

2. Contrast

\[
Contrast = \sum_{n=0}^{G-1} n^2 \left( \sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \right) \text{, } |i - j| = n
\]  

Contrast is a measure of the local variations present in an image. This measure of contrast will favour contributions from \(P(i,j)\) away from the diagonal, i.e. \(i = j\). If there is a large amount of variations in an image, the \(P[i,j]\)'s will be concentrated away from the main diagonal and the contrast will be a high value.

3. Inverse Difference Moment

\[
IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i-j)^2} P(i,j)
\]  

IDM is also influenced by the homogeneity of the image. Because of the weighting factor \((1+(i-j)^2)^{-1}\) IDM will get small contributions from inhomogeneous areas \((i,j)\). The result is a low IDM value for inhomogeneous images and higher value for homogeneous images.
4. Entropy

\[
Entropy = - \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \times \log(P(i,j)) \tag{3.7}
\]

Entropy statistic measures the disorder or complexity of an image. Complex textures tend to have high entropy. Entropy is strong, but inversely correlated to energy.

5. Correlation

\[
Correlation = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i \times j) \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \tag{3.8}
\]

Correlation is a measure of gray level linear dependence between the pixels at the specified positions relative to each other. The correlation will be higher, if an image contains a considerable amount of linear structure.

6. Sum of Squares, Variance

\[
Variance = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - \mu)^2 P(i,j) \tag{3.9}
\]

The variance is a measure of the dispersion of the gray level differences at a certain distance, \(d\). This feature puts relatively high weights on the elements that differ from the average value of \(P(i,j)\).

7. Difference Entropy

\[
DEnt = - \sum_{i=0}^{G-1} P_{x+y}(i) \log\left( P_{x+y}(i) \right) \tag{3.10}
\]

Difference entropy is a measure of histogram content and logical value between two images. If two images are identical the difference entropy will be
high otherwise low. Where $P_x(i)$ is the $i^{th}$ entry in the marginal-probability matrix obtained by summing the rows of $P(i,j)$ and $P_y(i)$ is obtained by summing the columns of $P(i,j)$.

8. Inertia

$$\text{Inertia} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - j)^2 \times P(i,j)$$

The inertia indicates the distribution of gray scales in the image.

9. Cluster Shade

$$\text{Shade} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i + j - \mu_x - \mu_y)^3 \times P(i,j)$$

The image is not symmetric when shade is high.

10. Cluster Prominence

$$\text{Prom} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i + j - \mu_x - \mu_y)^4 \times P(i,j)$$

The image is not symmetric when prominences are high.

11. Energy

$$\text{Energy} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i,j)^2$$

The energy of a texture describes the uniformity of the texture. Energy is 1 for a constant image.
12. Homogeneity

\[
Homogeneity = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{P(i,j)}{1 + |i - j|} \tag{3.15}
\]

Homogeneity returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for a diagonal GLCM. A homogeneous image will result in a co-occurrence matrix with a combination of high and low \( P[i,j] \)'s. A heterogeneous image will result in an even spread of \( P[i,j] \)'s.

13. Dissimilarity

\[
Dissimilarity = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} |i - j|P(i,j) \tag{3.16}
\]

Dissimilarity is a measure of evenness between two groups.

14. Difference in variance

\[
Variance = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i - \mu)^2P(i,j) \tag{3.17}
\]

Difference in variance is the sum of the difference between the intensity of the central pixel and its neighborhood.

Normalized GLCM \( N[i,j] \), is defined in eq. (3.18).

\[
N[i,j] = \frac{P[i,j]}{\sum_i \sum_j P[i,j]} \tag{3.18}
\]

which normalizes the co-occurrence values to lie between 0 and 1.

Feature selection algorithms are categorized into exponential, randomized and sequential algorithms. Exponential algorithms (e.g. branch and bound,
exhaustive) have exponential complexity in the number of features and are expensive to use (i.e. they have complexity $O(2^d)$ where ‘d’ is the number of features). Randomized algorithms include genetic and simulated annealing search methods. They give the optimum solution in very large search space. Sequential search algorithms have polynomial complexity (i.e. $O(d^2)$). There are two kinds of feature selection processes, one is based on feature selection criterion of classifier and another one is based on filtering approach.

In a classification problem, the number of features can be quite large; many of them can be irrelevant or redundant. All the features extracted from the input images does not account for high accuracy. So the optimal solution is not occurring. To solve this problem, feature reduction is introduced to improve the classification by searching for the best feature subset, from the fixed set of the original features, according to a given processing goal and a feature evaluation criterion: classification accuracy. In this research work GA is proposed to select the optimal features.

GA is an adaptive method of global-optimization searching and simulates the behaviour of the evolutionary process in nature [Sied1989]. The total features extracted are 14. The features selected using GA is 7. Therefore, optimized feature selection reduces data dimensionalities and computational time and increase the classification accuracy. The GA uses three basic operators to evolve the population: selection, crossover and mutation. From the experiments
conducted for feature selection, it is found that the optimal feature set which gives good classification performance are ASM, contrast, entropy, correlation, IDM, variance and dissimilarity details of the image. Hence for this application domain, the above 7 features are enough for classifying images into normal and abnormal.

A typical feature selection process consists of four basic steps. They are named subset generation, subset evaluation, stopping criterion and result validation. It is shown in Figure 3.7. $T$ denotes the subset of selected features and $V$ denotes the set of remaining features. So, $S = T \cup V$ at any time. $J(T)$ denotes a function evaluating the performance of $T$. $J$ depends on the particular application. Here $J(T)$ denotes the classification performance using the set of features in $T$.

![Feature selection process diagram]

**Figure 3.7:** Feature selection process

In the feature selection formulation of the GA, individuals are composed of chromosomes: a 1 in the bit position indicates that feature should be selected; 0 indicates this feature should not be selected. As an example the chromosome
00101000 means the 3rd and 5th features are selected. The 1, 2, 4, 6, 7, 8 features are not selected. That is the chromosome represents $T=\{3,5\}$ and $V=\{1,2,4,6,7,8\}$. Fitness function and penalty which is used to set constraint for the given chromosome $T$ is defined in eqs. (3.19)-(3.20).

\[
\text{Fitness} (T) = J(T) - \text{penalty}(T)
\] (3.19)

\[
\text{Penalty} (T) = w \times (|T| - d)
\] (3.20)

Here ‘$w$’ is the penalty coefficient. The size value ‘$d$’ is taken as a constraint and a penalty is imposed on the chromosomes breaking this constraint. The chromosome selection for the next generation is done on the basis of fitness. The fitness value decides whether the chromosome is good or bad in a population.

The selection mechanism should ensure that fitter chromosomes have a higher probability survival. So, the design adopts the rank-based roulette-wheel selection scheme. If the mutated chromosome is superior to both parents, it replaces the similar parent. If it is in between the two parents, it replaces the inferior parent; otherwise, the most inferior chromosome in the population is replaced. The normal feature values are given in Table 3.1 for the proposed method and in Table 3.2 for the existing method.
### Table 3.1: Image feature values of left by right ratio and right by left ratio for the proposed method

<table>
<thead>
<tr>
<th>Image features</th>
<th>Proposed Method</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L/R</td>
<td>R/L</td>
<td></td>
</tr>
<tr>
<td>ASM</td>
<td>0.50228</td>
<td>0.51911</td>
<td>0.96758</td>
<td>1.0335</td>
<td></td>
</tr>
<tr>
<td>Contrast</td>
<td>0.36185</td>
<td>0.38849</td>
<td>0.93143</td>
<td>1.0736</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>0.94689</td>
<td>0.9431</td>
<td>1.004</td>
<td>0.996</td>
<td></td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>0.18908</td>
<td>0.1881</td>
<td>1.0052</td>
<td>0.99484</td>
<td></td>
</tr>
<tr>
<td>Entropy</td>
<td>2.0349</td>
<td>1.9796</td>
<td>1.0279</td>
<td>0.97285</td>
<td></td>
</tr>
<tr>
<td>IDM</td>
<td>0.92005</td>
<td>0.92259</td>
<td>0.99725</td>
<td>1.0028</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>0.007797</td>
<td>0.0080625</td>
<td>0.96707</td>
<td>1.0341</td>
<td></td>
</tr>
</tbody>
</table>

L=Left side; R= Right side

### Table 3.2: Image feature values of left by right ratio and right by left ratio for the existing method

<table>
<thead>
<tr>
<th>Image features</th>
<th>Existing method</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L/R</td>
<td>R/L</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>0.95446</td>
<td>0.95627</td>
<td>0.9981</td>
<td>1.0019</td>
<td></td>
</tr>
<tr>
<td>Entropy</td>
<td>3.1121</td>
<td>3.1808</td>
<td>0.9784</td>
<td>1.0221</td>
<td></td>
</tr>
<tr>
<td>IDM</td>
<td>0.87917</td>
<td>0.8753</td>
<td>1.0044</td>
<td>0.9956</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>0.0030802</td>
<td>0.0029877</td>
<td>1.0309</td>
<td>0.96998</td>
<td></td>
</tr>
</tbody>
</table>

L=Left side; R= Right side
In the proposed method all the features of left by right ratio and right by left ratio are close to 1 compared to the circular adaptive region of interest method. Therefore in the proposed method the midline is the best indicator for measuring the degree of symmetry.

### 3.2.4 SEGMENTATION: K-MEANS CLUSTERING TECHNIQUE

Clustering methods are one of the most commonly used techniques for image segmentation. Clustering is an unsupervised way of data grouping with a given measure of similarity. Clustering techniques can be divided into hierarchical and partitional algorithms, where the main difference between them is that hierarchical methods produce a nested series of partitions while partitional methods produce only a single partition. As hierarchical methods are more accurate, partitional methods can be used in applications involving large datasets, like the ones related to images, as the use of nested partitions is computationally prohibitive. The traditional partitional clustering algorithm is the k-means algorithm, which is characterised by simple implementation and low complexity. The clusters produced by k-means clustering do not overlap.

In particular, many existing image segmentation algorithms fail to provide satisfactory results when the boundaries of the desired objects are not clearly defined by the image intensity information. K-means clustering algorithm was developed by J. MacQueen in the year 1967. K-means clustering is often suitable for biomedical image segmentation since the number of clusters is usually known
for images of particular regions of the human anatomy. In this research work the number of clusters is 6. This method has a number of advantages, such as its ability to handle large amounts of data points and its ability to work with compact clusters. However, it has its own set of limitations as well, such as the variables must be commensurable, the number of clusters should be known beforehand and it is sensitive to outliers and noise.

Segmentation plays a crucial role in many medical imaging applications. Medical image segmentation is the process of labeling each pixel in a medical image dataset to indicate its tissue type or anatomical structure. The labels that result from this process have a wide variety of applications in medical research and visualization. The desired output, or segmentation, contains the labels that classify the input greyscale pixels.

Manual segmentation of medical image by the radiologist is not only a tedious and time consuming process, but also not very accurate especially with the increasing medical imaging modalities and unmanageable quantity of medical images that need to be examined. It becomes therefore necessary to review current methodologies of image segmentation using automated algorithms that are accurate and require as little user interaction as possible especially for medical images. However, this consumes time and usually subjective, depending on the experience of the radiologist. Consequently the use of computer-aided systems becomes very necessary to overcome these limitations. Therefore,
segmentation of medical images is very important for clinical research, diagnosis and applications, leading to requirement of robust, reliable and adaptive segmentation techniques. Thus a robust algorithm for the segmentation of ischemic stroke region based on a k-means clustering is proposed.

In the abnormal brain imaging, the ischemic stroke region is extracted using k-means clustering technique [Hart1979, Hema2011]. The k-means clustering algorithm includes the following steps:

1. Select the number of clusters k with initial cluster centroid \( v_i, \ i = 1,2,\ldots,k \).

2. Partition the input data points into k clusters by assigning each data point \( x_j \) to the closest cluster centroid \( v_i \) using the selected distance measure, for example, Euclidean distance, defined as

\[
d_{ij} = \| x_j - v_i \| \tag{3.21}
\]

where \( X = \{x_1, x_2, \ldots x_n\} \) is the input data set.

3. Compute a cluster assignment matrix \( U \) representing the partition of the data points with the binary membership value of the \( j^{th} \) data point to the \( i^{th} \) cluster such that \( U = [u_{ij}] \), where \( u_{ij} \in \{0,1\} \) for all \( i,j \)

\[
\Sigma_{i=1}^{k} u_{ij} = 1 \text{ for all } j \text{ and } 0 < \Sigma_{j=1}^{n} u_{ij} < n \text{ for all } i. \tag{3.22}
\]

4. Re-compute the centroid using the membership values by

\[
v_i = \frac{\Sigma_{j=1}^{n} u_{ij} x_j}{\Sigma_{j=1}^{n} u_{ij}} \text{ for all } i. \tag{3.23}
\]

5. If a cluster centroid or the assignment matrix does not change from the previous iteration, stop; otherwise go to step 2.

‘The k-means clustering method’ optimizes the sum-of-squared-error-
based objective function $J_w (U, v)$ then

$$J_w(U, v) = \sum_{i=1}^{k} \sum_{j=1}^{n} \|x_j - v_i\|^2 \quad (3.24)$$

It can be spotted from the above algorithm that the k-means clustering method is quite sensitive to the initial cluster assignment and the choice of the distance measure. The additional criterion like within-cluster and between-cluster variances can be included in the objective function as constraints to force the algorithm to adapt the number of clusters $k$, as necessary for optimization of the objective function. The extracted stroke region using k-means clustering technique is shown in Figure 3.8(a-c).

![Ischemic stroke region using k-means clustering technique](a-c)

**Figure 3.8:** Ischemic stroke region using k-means clustering technique (a-c)

### 3.2.5 QUANTITATIVE ANALYSIS

The automatic segmentation results are compared to the manual results. The results are quantitatively evaluated by a human expert. A region of ischemic stroke identified by an experienced radiologist is the ground truth. Each pixel belongs to one of the four classes accordingly: True Positive (TP) is correctly
classified as positive pixels; False Positive (FP) is incorrectly classified as negative pixels; True Negative (TN) is correctly classified as negative pixels and False Negative (FN) is incorrectly classified as positive pixels. Three quantitative measures are employed to evaluate the segmentation result: precision, recall and overlap metric. They are defined in the following eqs. (3.25)-(3.27).

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (3.25)
\]

\[
\text{Recall or Sensitivity} = \frac{TP}{TP + FN} \quad (3.26)
\]

\[
\text{Overlap metric} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (3.27)
\]

The average overlap metric, average precision and the average recall between the results obtained using the proposed approach and the ground truth are given in Table 3.3.

| Table 3.3: Quantitative measures for ischemic stroke |
|-----------------------------|------------------|------------------|
| Overlap metric | Precision | Recall |
| 0.98 | 0.99 | 0.98 |

3.3 RESULTS AND DISCUSSION FOR SEGMENTATION OF ISCHEMIC STROKE DETECTION

The images are collected from 15 cases of acute ischemic stroke and 6 normal cases from the Department of Radiology, RMMCH. All the input data set used for ischemic stroke detection consisted of non-enhanced 512 × 512 CT brain images. Then it is resized to 256 × 256 pixel size to reduce the time complexity of further processing. The CT brain images collected from the
patients are acquired on Mx 8000 Dual, Philips CT scanner. The proposed method has been evaluated on a dataset of 21 patients (400 image slices). The proposed method consists of five stages. In the first stage noise is suppressed using median filtering and skull bone components of the images are removed by a global thresholding method. In the second stage, midline shift of the brain is calculated. Symmetry map is calculated at this stage and Bezier curve is drawn for the control points.

In the third stage, fourteen texture features are extracted using GLCM. The following GLCM features are extracted in this research work: ASM, entropy, correlation, sum of squares, difference entropy, inertia, cluster prominence, IDM, cluster shade, energy, homogeneity, dissimilarity and difference in variance. The 7 optimal features are extracted from the left and right side of the brain using GA. The optimal features are ASM, contrast, entropy, correlation, IDM, variance and dissimilarity. The optimal features are used to train the binary classifier, which can automatically infer whether the image is that of a normal brain or an ischemic brain suffering from a brain lesion. In the next stage, SVM, ANN and decision tree classifiers have been used to classify normal brain and ischemic brain. Finally, the stroke region is extracted using k-means clustering technique. The original CT brain image without stroke is shown in Figure 3.9(a). CT brain image which is stained with ischemic stroke lesion is shown in Figure 3.9(b-f).
Figure 3.9: Original CT brain images: (a) Normal brain image and (b-f) Abnormal brain images (Ischemic stroke region marked in green colour)

The results are quantitatively evaluated by a human expert. The results are compared with recent results [Fuk2011] for the same CT image datasets. In the circular adaptive region of interest method eight features are extracted and the accuracy of the ANN is 89%. In the proposed method the accuracy of SVM, ANN and decision tree are 97%, 94% and 91%. Thus SVM is the best classifier for this dataset. The normal image feature values of left by right ratio and right by left ratio is shown in Table 3.1 for the proposed method. The normal image feature values of left by right ratio and right by left ratio is shown in Table 3.2 for the circular adaptive region of interest.
In the proposed method all the features of left by right ratio and right by left ratio are close to 1 compared to the circular adaptive region of interest method. Therefore in the proposed method the midline is the best indicator for measuring the degree of symmetry.

The average overlap metric, average precision and average recall between the results obtained using the proposed approach and the ground truth are 0.98, 0.99 and 0.98. The performance of the algorithm is very good. The application of the proposed method for early detection of ischemic stroke is demonstrated to improve efficiency and accuracy of clinical practice. The proposed method helps to increase the accuracy in detection of ischemic stroke and hence it decreases the risk of misdiagnosis and mismanagement. The output images of size 256 × 256 of the proposed method are given in Figure 3.10(a-h).
Figure 3.10: An ischemic stroke detection process in brain CT image:
a) Original brain CT image; b) Image after skull bone removal; c) Generation of
the symmetry map; d) Generation of Bezier curve; e), f) Left and right image
extraction; g) Ischemic stroke region using proposed method and h) Ischemic
stroke region using manual method
3.4 SUMMARY

The following gives a brief summary of the entire procedure used to detect ischemic stroke in CT images. In this chapter, a new approach to automated detection of ischemic stroke, using k-means clustering technique which separates the lesion region from healthy tissues in CT images are discussed. The experimental results obtained from different stages of ischemic stroke detection are discussed. This thesis presents a new approach to automated classification of ischemic stroke using texture features. The results are quantitatively evaluated by a human expert. The average overlap metric, average precision and average recalls between the results obtained using the proposed approach and the ground truth are 0.98, 0.99 and 0.98.