International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Issue 10, Volume 6 (October 2019) www.ijirae.com

# A STUDY OF MEDICAL IMAGE PROCESSING AND SEGMENTATION METHODS FOR MEDICAL IMAGE ANALYSIS

**Amrutha K S[1], student**
Department of CSE, ASIET, Kalady, Kerala
amrutha57476@gmail.com

**Asst.Prof.M.Gokilavani [2]**
Department of CSE, ASIET, Kalady, Kerala
gokilavani.cs@adishankara.ac.in

**Manuscript History**
Number: **IJIRAE/RS/Vol.06/Issue10/OCAE10082**
Received: 11, October 2019
Final Correction: 16, October 2019
Final Accepted: 22, October 2019
Published: **October 2019**
**Citation:** Amrutha & Gokilavani (2019). A Study of Medical Image Processing and Segmentation Methods for Medical Image Analysis. IJIRAE::International Journal of Innovative Research in Advanced Engineering, Volume VI, 609-615. **doi://10.26562/IJIRAE.2019.OCAE10082**
**Editor:** Dr.A.Arul L.S, Chief Editor, IJIRAE, AM Publications, India


**Abstract**—MRI is the most important technique, in detecting the tumors in various body parts. In this paper survey of various data mining methods are used for classification of MRI images. A new hybrid technique based on the support vector machine (SVM) and fuzzy c-means for brain tumor classification is studied in this paper. The algorithm is a combination of support vector machine (SVM) and fuzzy c-means, a hybrid technique for prediction of brain tumor. In this algorithm, the image is enhanced using enhancement techniques such as contrast improvement, and mid-range stretch. Fuzzy c-means (FCM) clustering is used for the segmentation of the image to detect the suspicious region in brain MRI image.

**Keywords**— Support vector machine (SVM); MRI;

## I. INTRODUCTION

Data mining may well be a straight forward and robust tool to extract the data from massive dataset [1]. Classification is a branch of data mining field. During this field, many classification techniques are available for medical footage like artificial neural network (ANN), fuzzy c-means (FCM), support vector machine (SVM), decision tree and Bayesian classification. Variety of researchers has been implementing the classification techniques for medical footage classification. Presently many medical imaging techniques like (PET), x-ray, CAT (CT), resonance imaging (MRI), for tumour detection but MRI imaging technique is the smart owing to higher resolution and most researchers have used MRI imaging for designation tumour. During this paper, the MRI images were high during contrast improvement and Mid-Range Stretch techniques. Once the image was improved, segmentation step is usually done simply. Segmentation is a technique to extract suspicious area from footage. In this paper, Segmentation technique was done by Fuzzy C-Mean (FCM) agglomeration [2]. Before applying FCM agglomeration technique, skull masking has been done. Feature extraction means that to induce the information of image. The strategy uses gray Level Run Length Matrix (GLRLM) to extract feature [3]. The reduced GLRLM qualities are outline to support vector machine for coaching and testing. The brain MRI images were differentiating using SVM techniques which widely used for information analysing and pattern recognizing. It creates a hyper plane in between information sets to point that category it belongs to [4]. The foremost objective of this work is to develop a hybrid technique, which could classify the brain MRI images successfully and efficiently via Fuzzy C- implies that and support vector machine (SVM). This work is a cheap classification technique is to observe the tumour in MRI images.


 

Accurate diagnosis for different types of cancer plays an important role in determining and choosing the proper treatment to the doctors to assist them. By using classification techniques, possible errors that might occur due to unskilled doctors can be minimized. Challenge facing medical practitioners makes this study of a much greater significance. Since symptoms appear only in the advanced stages thereby causing the mortality rate of lung cancer to be the highest among all other types of cancer, challenging the detection of cancer in its early stages [1]. The objective of undertaking this project is to facilitate doctors to provide the best possible treatment by providing useful insights with the help of predictive models through analysis and diagnosis of lung cancer treatments. This technique can also examine medical data in a shorter time and more precisely. The critical task is to define and specify a good feature space that means the type of features which will discriminate between malignant and benign.

## II. MEDICAL IMAGE PROCESSING

### A. Algorithm used in medical image processing projects:

Medical Image Processing Projects are developed based on image processing simulation tool named as Matlab. Using the tool processing more medical images of human organs are (Brain, Lung, Kidney, Skin, Retina, Finger, Tissues and Skull). According to the modality results the physician can easily observe the pathologies directly but sometimes it took more time to analysing. Image analysis process can be automated for producing interesting results about human diseases. In medical image processing projects we have to use more algorithms to identify and classify the diseases in the images. Segmentation and classification methods are used to detect the disease and known the status of the human. Some of the commonly used classification algorithms are

- Support Vector Machine
- Fuzzy C-Means Clustering.
- K-NN Classification.
- Naive Bayes Classification.
- Decision Trees.
- Genetic Algorithm.
- Neural network Classification.

Before enter into the process of classification we must do the processes of pre-processing, feature extraction and feature reduction. Pre-processing based on some gray scale conversion methods, noise removal concepts. Feature extraction is the process of extracting features in the images with its pixels. Most commonly extracted features are color, shape, texture, geometric features. In medical image processing greatly define the texture features. Then extracted features are to be reduced using feature reduction or selection method. The particular selected features are used to identify the disease in the human organ. Requirements for medical image processing are image enhancement, changing density range of B/W images, manipulating colors, image line profile display, image restoration, image smoothing, Biomedical image area calculation, detection of contour. Applying the algorithms of filtering and reconstruction automatically identify the 3D image data, it must take more number of object slices for processing.

### B. Image Processing Technique:

Image processing is a method to convert an image into digital form and perform some operations on it in order to get an enhanced image or to extract some useful information from it.Image processing basically includes the following three steps.

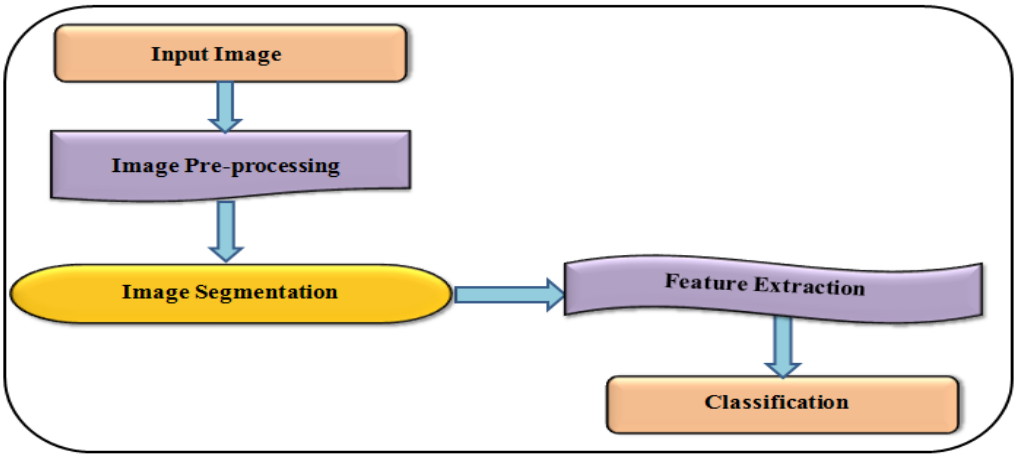


Fig 1: Image processing technique


 

- Importing the image with optical scanner or by digital photography.
- Analyzing and manipulating the image that includes data compression and image enhancement and spotting patterns.
- The last is the output in which result can be altered image.

## III. SVM CLASSIFIER

Next phase in the proposed system is the classification of occurrence and non-occurrence of cancer nodule for the supplied lung image. The classifier used is Support Vector Machine. The aim of classification is to group items that have similar feature values into groups.

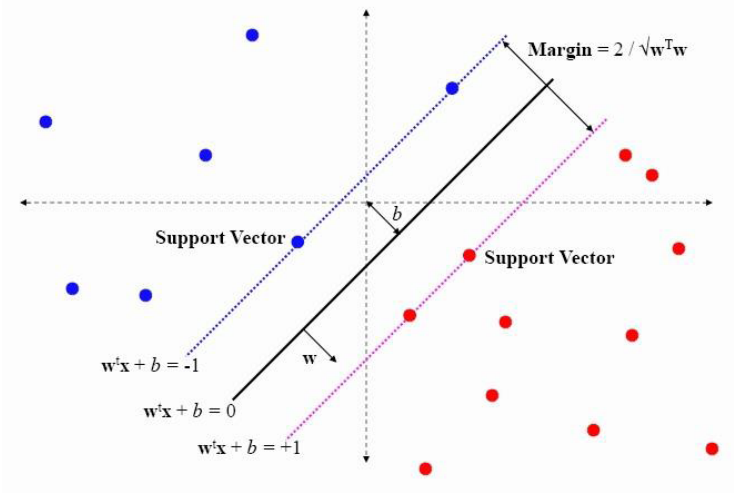


Fig 2: SVM Classifier

Classifier achieves this by making a classification decision based on the value of the linear combination of the features. SVM is a binary classification method that takes as input labeled data from two classes and outputs a model file for classifying new unlabeled/labeled data into one of two classes.

Training an SVM involves feeding known data to the SVM along with previously known decision values, thus forming a finite training set. It is from the training set that an SVM gets its intelligence to classify unknown data. In SVM, for two class classification problem, input data is mapped into higher dimensional space using RBF kernel. Then a hyper plane linear classifier is applied in this transformed space utilizing those patterns vectors that are closest to the decision boundary, [16], shown in fig-3.

Consider the pattern classifier, which uses a hyper plane to separate two classes of patterns based on given examples {x (i), y (i)}, i=....n .Where (i) is a vector in the input space I=Rk and y (i) denotes the class index taking value 1 or 0. A support vector machine is a machine learning method that classifies binary classes by finding and using a class boundary the hyper plane maximizing the margin in the given training data. The training data samples along the hyper planes near the class boundary are called support vectors, and the margin is the distance between the support vectors and the class boundary hyper planes. The SVM are based on the concept of decision planes that define decision boundaries.

### A. Training the classifier

In the training phase, known data is given and the classifier is trained. Given training data (xi, yi) for i = 1 ...N, with xi ∈Rd and yi∈ {−1, 1} .The training points satisfy the following conditions.

$$F(x) = W^T x_i + b \geq +1 \text{ for } y_i = +1$$

$$F(x) = W^T x_i + b \leq -1 \text{ for } y_i = -1$$

### B. Testing the data

In testing phase, unknown data are given and the classification is performed using trained classifier. Classification is done by using following decision function.

$$F(x, \{w, b\}) = \text{sign}(w.x+b)$$

Every input x is initially mapped into a higher dimension feature space F, by z=φ(x) through a nonlinear mapping φ: Rn→F. W is the normal to the line, x is the feature vector and b the bias. W is known as the weight vector and b is bias.


 

## IV. RELATED WORK

### A. A Survey on Image Classification Approaches and Techniques

A support vector machine [5] builds a hyper plane or set of hyper planes in a high- or infinite dimensional space. Classification is achieved by selecting the hyper plane that has the largest distance to the nearest training data point of any class. SVM uses Nonparametric with binary classifier approach. Non parametric approach has no assumptions about the data and no statistical parameters are used for classification. Binary classifier Classifies to 2 classes. Performance and accuracy depends upon the hyper plane selection and kernel which is a mathematical function used for performing SVM.

### B. The Melanoma Skin Cancer Detection and Classification using Support Vector Machine

In order to classify using SVM [6], we need to find the best plane which separates the data in an efficient possible manner. Most cases, the plane used will be linear if the data can be linearly separated- represents a line which separates points of one class with points of another class. If the data cannot be linearly separated, the kernel (i.e., mathematical function) used is Radial basis function. Radial basis function is a function whose value depends on the distance from the origin or from some point. Gaussian Kernel is of the following format,

$$K(X_1, X_2) = exponent(-\gamma \|X_1 - X_2\|^2)$$

||X1 — X2 || = Euclidean distance between X1 & X2.

In the training stage, the classifier was built using cross validation procedure to find the optimized parameters of the hyper plane to avoid biasing with over fitting. The full set of features (11 features) and the selected 5 features using PCA (Principal Component Analysis- used for feature extraction) are fed into the SVM model which is used to classify the image into binary classes benign and malignant. SVM produced results with 92.1% accuracy, with full set of features and also with the PCA selected features.

### C. A Fast SVM classification learning algorithm used to large training set

For Fast SVM [7], the sample points which are not in the same class are eliminated first and then the relative boundary vectors (RBVs) are computed. Not only the RBV sample itself, but a near RBV sample whose distance to the RBV is smaller than a certain value will also be selected for SVM training in order to prevent the loss of some critical sample points for the optimal hyper plane. First it will find the distance between two training samples as

$$K_{CPD}(x_i, x_j) = -\|x_i - x_j\|^q + 1 \quad , \qquad 0 < q \le 2$$

Now, the distance in the mapping space can be simplified

;

$$d_\varphi(x_i, x_j) = \sqrt{\|x_i - x_j\|^q}$$

If the value of $q$ is taken as 2, then

$$d_\varphi(x_i, x_j) = \|x_i - x_j\|$$

It also find the relative boundary vector and then selects a pruning strategy for the training samples.

### D. Recognizing Human Actions: Local SVM Approach

SVM method used for classifying local features through spatial recognition. Local features are being taken as human actions. Here the hyper plane is $w \cdot x + b = 0$ in some space H and that have no prior knowledge about the data distribution, and then the optimal hyper plane is the one which maximizes the margin. The optimal values for w and b can be found by solving a constrained minimization problem, using Lagrange multipliers αi (i = 1,...m). It follows that LF with local SVM [8] gives the best performance for all training sets while the performance of all methods increases with the number of scenarios used for training.

### E. Classification of Structural images via High-Dimensional Image Warping, Robust Feature Extraction and SVM

The results from a group of 61 brain images of female normal controls and schizophrenia patients demonstrate not only high classification accuracy (91.8%) and steep ROC curves, but also exceptional stability with respect to the number of selected features and the SVM kernel size. SVM constructs a maximal margin linear classifier in a high (often infinite) dimensional feature space, by mapping the original features via a kernel function. The Gaussian radial basis function kernel is used in our method. SVM [9] is not only empirically demonstrated to be one of the most powerful pattern classification algorithms, but also has provided many theoretic bounds on the leave one-out error to estimate its capacity, for example, the radius/margin bound, which could be utilized in feature selection. Another reason for us to select SVM as a classifier is its inherent sample selection mechanism, i.e., only support vectors affect the decision function, which may help us find subtle differences between groups. The classification result on the testing subject using the trained SVM classifier was compared with the ground-truth class label, to evaluate the classification performance.


 

Absolutely all feature selection and training steps were cross-validated, i.e. the testing image had no influence on the construction of the classifier. By repeatedly leaving each subject out as testing subject, we obtained the average classification rate from 61 leave-one-out experiments. Finally, these experiments were repeated for different numbers of features, in order to test the stability of the results. The best average correct classification rate was 91.8% by using 37 features. This adaptive regional feature extraction method aims at overcoming the limitations of the traditional ROI methods that need prior knowledge of what specific regions might be affected by disease, and the limitations of the voxel based morphometric (VBM) methods that use an identical isotropic filter to collect regional morphological information in all brain locations.

### F. Large-scale Training Image Classification: Fast Feature Extraction and SVM Training

For SVM training in large scale image classification, developed a parallel averaging stochastic gradient descent (ASGD) algorithm for training one-against-all 1000-class SVM classifiers [10] . The ASGD algorithm is capable of dealing with terabytes of training data and converges very fast – typically 5 epochs are sufficient. As a result, we achieve state-of-the-art performance on the Image Net 1000-class classification, i.e., 52.9% in classification accuracy and 71.8% in top 5 hit rate.

### G. Image classification using Support Vector Machine and Artificial Neural Network

Firstly, we separate the image into many sub-images based on the features of images. Each sub-image is classified into the responsive class by an ANN. Finally, SVM [11] has been compiled all the classified result of ANN. Our proposal classification model has brought together many ANN and one SVM denoted as ANN_SVM. ANN_SVM has been applied for Roman numerals recognition application and the precision rate is 86%. The experimental results show the feasibility of our proposal model.

### H. Image Classification Optimization Algorithm based on SVM

For the image classification optimization algorithm includes neural network, Bayesian and Fuzzy sets, etc. But these algorithms have high training complexity, low convergence speed, etc. In view of this, this paper proposed an image classification optimization algorithm based on support vector machine (SVM) [11] . This algorithm is finished mainly by the following steps:

- Select the proper kernel function;
- Confirm parameters of kernel function with the method of grid search;
- Make a feature extraction based on colour and vein to images, which can be taken as input to achieve the classification of images;
- The experiment shows the effectiveness of optimization algorithm of image classification based on SVM proposed in this study.

Meanwhile, support vector machine is built on the basis of structural risk minimization principle of statistical learning theory and VC-dimensional theory. SVM can map the input vectors of nonlinear samples into a high-dimensional feature space through introducing a kernel function when solving nonlinear classification problem, in order to achieve the transformation from linearity to nonlinearity, and then construct one optimal classification hyper plane in the above feature space. If it is linear separable, support vector machine (SVM) can use the hyper plane directly to classify. When dealing with most problems, however, it is nonlinear. In this case, we need to use the core functions to form transformation, and convert raw data from low dimensional space to high dimension space, so that it can achieve the goal of linearly separable. Support vector machine (SVM) is a binary classifier. When solving multiple-class problem, we need to combine with multiple binary classifiers to complete multiple classification purposes. In general there are two ways of dealing with the multiple classification problems:

- One-against-all. When one-against-all is in dealing with k kinds of problems, it will produce k support vector machines. Among them, the way of producing the support vector machine is to marked the data to "+ 1", the others are marked as "-1".
- One-against-one. When the one-against-one method processes the classification problem, every binary data will create a SVM.

Therefore, for the classification problems which have k categories, it will have k k (1) / 2 support vector machines. Namely, if there are five types of data, it will be divided into 10 support vector machines. Selection of kernel function has a decisive influence on classification ability and type of SVM classifier. RBF kernel function has excellent learning ability under the circumstance of large sample and low dimension, so it has been widely used

### I. Data Classification using support Vector Machine

For the image classification optimization algorithm includes neural network, Bayesian and Fuzzy sets, etc. But these algorithms have high training complexity, low convergence speed, etc. In view of this, this paper proposed an image classification optimization algorithm based on support vector machine (SVM) [11] . This algorithm is finished mainly by the following steps:


 

- The RBF kernel nonlinearly maps samples into a higher dimensional space unlike to linear kernel
- The RBF kernel has less hyper parameters than the polynomial kernel.
- The RBF kernel has less numerical difficulties.

Rough set is a new mathematical tool to deal with un-integrality and uncertain knowledge. It can effectively analyze and deal with all kinds of fuzzy, conflicting and incomplete information, and finds out the connotative knowledge from it, and reveals its underlying rules. Rough Set Theory is successfully used in feature selection and is based on finding a reduce from the original set of attributes. Use LIBSVM with different kernel linear, polinomial, sigmoid and RBF and two parameters, the RBF kernel parameter γ and the cost parameter C, to be set. The combinations of (C, γ) are the appropriate for the data classification problem with respect to prediction accuracy. The choice of kernel function and best value of parameters for particular kernel is critical for a given amount of data and the best kernel is RBF for infinite data and multi class.

### J. Method for Melanoma Skin Cancer Detection using Dermoscopy Images

SVM with linear function and radial basis function is used to classify the data into malignant or benign. 10 features has been extracted and given to SVM [13] classifier. The results are obtained using colour, texture and shape features. Sensitivity, accuracy and specificity is calculated using SVM classifier. This is applied with varied % of training data and testing data, so that to find out which kernel function would best classify the data set. A binary classifier has 4 outcomes as

- True positive (TP): correct positive prediction
- False positive (FP): incorrect positive prediction
- True negative (TN): correct negative prediction
- False negative (FN): incorrect negative prediction

Error rate (ERR) is calculated as the number of all incorrect predictions divided by the total number of the dataset. Accuracy (ACC) is calculated as the number of correct predictions divided by the total number of the dataset. The best accuracy is 1.0, whereas the worst is 0.0. It can also be calculated by 1 – ERR. Sensitivity (SN) is calculated as the number of correct positive predictions divided by the total number of positives. It is also called recall (REC) or true positive rate (TPR). Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. It is also called true negative rate (TNR).

## V. CONCLUSIONS

This paper attempts to study and provides a brief knowledge about the different image classification approaches and different classification methods. Most common approaches for image classification can be categories as supervised and unsupervised, or parametric and nonparametric or object-oriented, subpixel, per-pixel and perfield or spectral classifiers, contextual classifiers and spectral-contextual classifiers or hard and soft classification. This survey gives theoretical knowledge about different classification methods and provides the advantages and disadvantages of various classification methods.

## ACKNOWLEDGMENT

The authors are thankful to Asst.Prof.M.Gokilavani CSE doing research in the area of Medical image processing and working as assistant professor.

## REFERENCES

1. Pathak, Akhanda Nand, And Ramesh Kumar Sunkaria. "Multiclass Brain Tumor Classification Using Svm." International Journal Of Computer Applications 97.23 (2014).
2. Kailash Sinha, G.R.Sinha, “Efficient Segmentation Methods for Tumor Detection in MRI Images”, 2014 IEEE Student’s Conference on Electrical, Electronics and Computer Science, 978-1-4799-2526- 1/14/$31.00 ©2014 IEEE.
3. R. S. Raj Kumar and G. Niranjana, “Image Segmentation and Classification of MRI Brain Tumor Based on Cellular Automata and Neural Networks”, IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013 .ISSN: 2320 – 8791.
4. H. B. Nandpuru, Dr. S. S. Salankar, Prof. V. R. Bora, MRI brain cancer classification using support vector machine. 2014 IEEE Students 'Conference on Electrical, Electronics and Computer Science, 978-1- 4799-2526-1/14/$31.00 ©2014 IEEE.
5. Pooja Kamavisdar, Sonam Saluja, Sonu Agrawal,” A Survey on Image Classification Approaches and Techniques”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2013.
6. Diwakar Gautam, Mushtaq Ahmed.” Melanoma Detection and Classification Using SVM Based Decision Support System”, IEEE INDICON 2015.



 

7. Qihua Xu, Shuai Geng ,"A Fast SVM Classification Learning Algorithm Used to Large Training Set ",January 2012 ISDEA '12: Proceedings of the 2012 Second International Conference on Intelligent System Design and Engineering Application ,IEEE Computer Society .
8. Christian Sch¨uldt, Ivan Laptev, Barbara Caputo," Recognizing Human Actions: A Local SVM Approach", Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04),IEEE.
9. Yong Fan, Dinggang Shen, and Christos Davatzikos," Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM", © Springer-Verlag Berlin Heidelberg 2005.
10. Liangliang Cao and Thomas Huang," Large-scale Image Classification: Fast Feature Extraction and SVM Training",2010.
11. Le Hoang Thai, Tran Son Hai, Nguyen Thanh Thuy," Image Classification using Support Vector Machine and Artificial Neural Network", I.J. Information Technology and Computer Science, 2012, 5, 32-38.
12. C. H. Qian et al., "An Image Classification Algorithm Based on SVM", Applied Mechanics and Materials, Vols. 738-739, pp. 542-545, 2015.
13. Durgesh K. Srivastava, 2lekha Bhambhu," Data Classification Using Support Vector Machine", Jati, 2009.
14. Rosy Kumari," SVM Classification an Approach on Detecting Abnormality in Brain MRI Images", IJERA ,Vol. 3, Issue 4, Jul-Aug 2013, pp.1686-1690.
15. Shinde, Swati," A Method for Melanoma Skin Cancer Detection Using Dermoscopy Images", 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).


 