Support Vector Machine for MRI Stroke Classfication

A.S.Shanthi
Assistant Professor, Department of Computer Science and Engineering
Tamilnadu College of Engineering, Karumathampatti
Coimbatore, India
babushanthi@gmail.com

M.Karthikeyan
Professor and Head of Department of Electronics and Communication
Tamilnadu College of Engineering, Karumathampatti
Coimbatore, India
mkarthikn@rediffmail.com

***Abstract*—Magnetic resonance imaging (MRI) is a low-risk, non-invasive imaging technique without ionizing radiation hazard, providing high quality/high contrast images and functional images of anatomical structures and organs. In this study, a method to classify the MRI images of the brain related to stroke is presented. Stroke is a sudden development of neurological damage. In this paper, the MRI images for stroke classification use Gabor filters and Histograms to extract features from the images. In the proposed feature extraction method, the features from Gabor filter and histogram features are fused. The extracted features are classified using Support Vector Machine (SVM) with various kernels. Experimental results are shown that the presented method achieves satisfactory classification accuracy**

Keywords- Image Classification, MRI, Stroke, Gabor Filters, Histograms, Support Vector Machine (SVM)
***Introduction* (Heading 1)**

I.INTRODUCTION

Magnetic resonance imaging (MRI) is a low-risk, without ionizing radiation hazard, non-invasive imaging technique providing high quality/high contrast images and functional images of anatomical structures and organs. With MRI soft tissue structures like, heart, lungs, liver and brain are seen in clear detail when compared to other medical imaging modalities. MRI's non-invasive nature and provision of rich information ensures it wide use for both diagnosis and treatment [1].

Interpretation of medical image includes 3 key tasks: (i) image findings perception, (ii) its interpretation to ensure diagnosis/differential diagnosis, and (iii) recommendations for clinical management including biopsy and follow up or more imaging if a clear diagnosis is not evident [2].CBIR (content base image retrieval) is of two types including global image retrieval and object/region based retrieval. In the former, query images are matched with database images through consideration of the whole image as a single feature set, whereas in the latter, image objects are segmented and these segmented subsets are matched to images in a dataset. Though object segmentation from images at the semantic level with many real world images is impractical, they can still be separated into constituent regions with homogeneous properties. In reality, regional properties are related to low level features like color, shape and texture, the latter in fact, plays a major role in discriminating regions [3].

Of the two components of a generic CBIR system, the first represents visual information in image pixels as image features/descriptors aiming to bridge the visual content and numerical representation gap. Such representations encode an image's color and texture properties, an object's spatial layout and geometric shape characteristics of perceptually coherent structures in an image. The second component assesses image feature similarities based on mathematical analyses comparing descriptors across images [4, 5].

Due to its task's nature CBIR technology is about 2 intrinsic problems, (1) how to describe an image mathematically and (2) how to assess an image pair similarity based on abstracted descriptions. The first problem is due to an image's original representation – an array of pixel values - corresponding poorly to visual response, an image's semantic understanding. For retrieval purposes an image's mathematical description is called its signature. Signature extraction and image similarity calculation cannot be separated from a design perspective. Signature formulation determines definitions of similarity measures. But, intuitions are usually early motivating factors to design similarity measures in a specific way. This puts requirements on signatures construction. When compared with pre-2000 CBIR work a marked difference recently is image signatures increased diversity. Advances in new features derivation and the features based signature construction has resulted in progress in the latter type being more emphatic. Signatures mathematical formulations richness grows along with the invention of new similarity measuring methods [6].

This study suggests a method to classify brain MRI images related to stroke. Stroke is sudden neurological damage development. Strokes are of 2 types: hemorrhagic (15%) and ischemia (85%). Blood clots in strokes are dangerous as blood flow is cut off due to arteries being blocked, which is called Ischemia. An ischemic stroke is of two types; embolic and thrombotic [7].

Embolic Stroke: Here, a blood clot formed somewhere in the body (usually the heart) travels to the brain via the bloodstream. When it reaches the brain, it enters a blood vessel which is small enough to block blood flow. The clot lodges there and a stroke is the result due to the blood vessel being blocked. Embolus is the medical term for this type of clot.

Thrombotic Stroke: Hereproblems are due to one or more arteries supplying blood to the brain being blocked. Thrombosis is the process which leads to this blockage. Strokes due to this are called thrombotic strokes due to the medical word for a clot in a blood vessel deposit is thrombus.

Hemorrhagic: Strokes caused by blood vessel breakage or "blowout" in the brain are known as hemorrhagic strokes as the medical word for such breakage is hemorrhage. Hemorrhages are caused by various disorders affecting blood vessels and can include long-standing high blood pressure and cerebral aneurysms. Hemorrhagic stroke are of 2 types: subarachnoid and intra cerebral.

Image classification/categorization is often a preprocessing step to expedite image retrieval in large databases and improve accuracy for automatic image annotation. Unsupervised clustering ensures retrieval speedup and also improved result visualization when labeled data is absent. Though image clustering is dependent on similarity measure, various methods can perform image categorization that needs no similarity metrics. Image categorization is succeeded by similarity measurement, restricted to a large database images belonging to the same visual class as predicted for query.

In this paper, the MRI images for stroke classification use Gabor filters and Histograms to extract features from the images. The extracted features are classified using Support Vector Machine (SVM) with various kernels.

II. RELATED WORKS

Huang et al [8] explored SVM (support vector machine) for infarct prediction on individual pixel basis using acute cerebral blood flow (CBF), apparent diffusion coefficient (ADC) MRI data. The SVM prediction model’s efficacy was tested on 3 stroke groups: 30-min, 60-min, and permanent middle-cerebral-artery occlusion (n=12 rats for each group).The acute phase led to the acquisition of CBF, ADC and relaxation-time constant (T2) up to 3hrs and again at 24hrs. Usage of acute (30-min) stroke data Infarct led to the prediction. Receiver-operating characteristic (ROC) analysis quantified prediction accuracy. Receiver-operating curves were 86±2.7%, 89±1.4%, and 93±0.8% using ADC+CBF data for the 30-min, 60-min and permanent middle cerebral artery occlusion (MCAO) group, respectively. Addition of neighbouring pixel information and spatial infarction incidence improved performance to 88±2.8%, 94±0.8%, and 97±0.9%, respectively. SVM prediction compared favourably to earlier publish artificial neural network (ANN) prediction algorithm operating similar data sets. SVM prediction model can provide quantitative frameworks to help clinical decision-making in acute stroke treatment.

A classification approach to differentiate between proteomic stroke patient’s samples and controls, and another second novel predictive model is described by Reddy et al [9] and developed to predict stroke severity as measured by the National Institutes of Health Stroke Scale (NIHSS). Logical Analysis of Data (LAD) methodology aided model construction to mass peak profiles of 48 stroke patients with 32 controls. This model 75% accuracy when tested on 35 stroke patients and 25 controls independent validation set. Superior performance was exhibited by the predictive model as compared to alternative algorithms. Both models, despite high accuracy are very simple, being developed through the use of a common 3 peak set.

Functional magnetic resonance imaging (MRI) scans usefulness from an auditory language comprehension experiment to predict individual language recovery in 21 aphasic stroke patients was explored by Saur et al [10]. Subjects having least moderate language impairment underwent extensive language testing 2 weeks and 6 months following a left-hemispheric stroke. Six months after a stroke, multivariate machine learning predicted language outcomes. The aim was to predict language improvement degree over 6 months. 76% patients were correctly separated as having good and bad language performance 6 months after stroke, based on functional MRI data from relevant language areas. Accuracy further improved (86% correct assignments) When age and language scores were entered alongside functional MRI data accuracy further improved to 86 %. The same accuracy was arrived at when predicting language improvement degree, based on imaging, age/language performance. Only chance level prediction was possible when exploring diffusion weighted imaging usefulness as well as functional MRI acquired two days after stroke. Current machine learning techniques high potential to predict system-specific clinical outcome for a disease as heterogeneous as stroke was demonstrated by this study. Best language recovery prediction is possible when brain activation potential post system-specific stimulation is assessed in the second week after a stroke. Intensive early rehabilitation is possible for those with predicted poor recovery and extension to other systems like motor and attention, is possible.

III. METHODOLOGY

In this paper it is proposed to investigate stroke classification by fusing Gabor filter and Histogram features. The Gabor filter with histogram feature alternating between mean and standard deviation combination for feature selection is used. The feature set so obtained is classified using SVM with various kernels.

Feature Extraction

Gabor wavelet and derivatives attempt feature extraction at a different frequency signal components to provide the signal's basic functions [11]. Texture is the two-dimensional signal with different frequency bands at multiple descriptions. As Gabor energy captures the signal's local features, texture features are computed from this energy distribution group. Gabor filter's tunable property at various scale and orientation ensure its usefulness for texture analysis. Gabor filters act as local band pass filter having some optimal joint localization properties in both spatial and spatial frequency domains. A schematic diagram to extract texture features [3] is revealed in Figure 1.

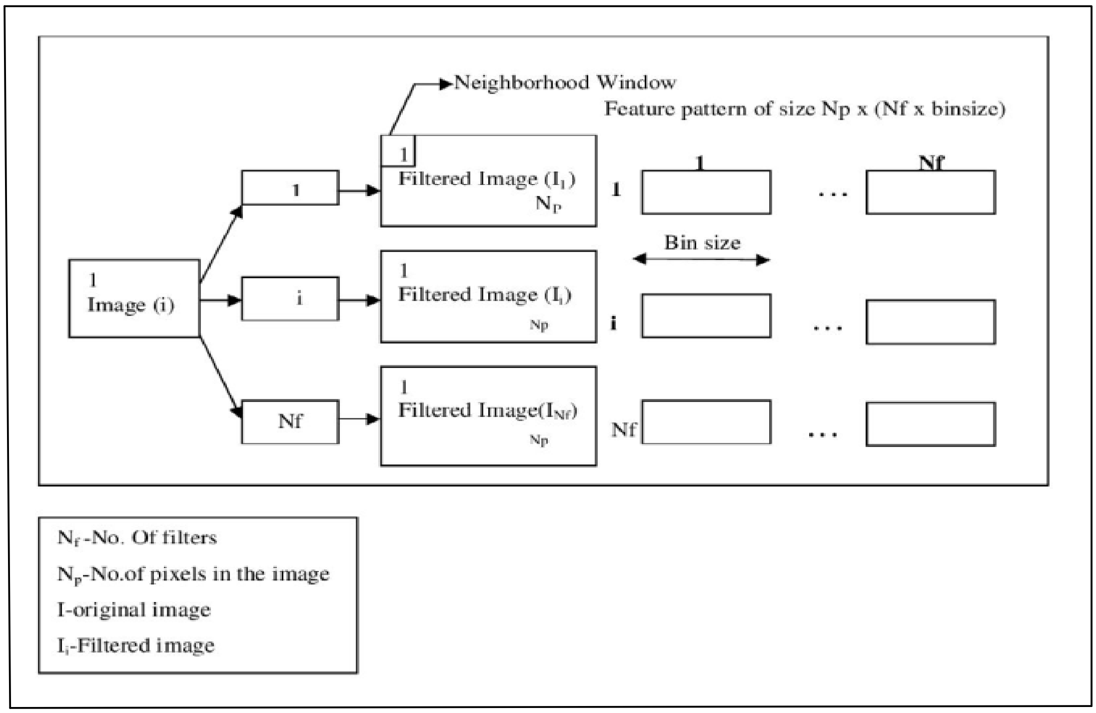

Figure. 1: A Schematic diagram of extracting texture features

A Gabor functions are complex, sinusoidal functions modulated by a rotated Gaussian [12]. Gabor function's spatial domain equation is given in Eq.(1)

$$g(x, y) = g_a(x', y)e^{(2\pi\, jfx)} \quad (1)$$

where the rotated co-ordinates (x`,y`) are shown in Eq.(2)

$$(x', y') = (x\cos\phi + y\sin\phi, -x\sin\phi + y\cos\phi) \quad (2)$$

Here ga is the Gaussian function.

By selecting a Gabor wavelet sensitive to this narrow frequency band and orientations, it is possible to know the presence of a specific texture. Channels filter images at a specific frequency tuned to an image's textures. Channel output magnitude is high appropriate frequency regions.

The histogram method can represent, analyze, and recognize images due to its easy calculation and efficiency. It is also robust to noise and local image transformations [13, 14]. Identification of 3-D objects [15] is an initial histogram application due to Swain and Ballard's work. Later, many histogram based recognition systems came into being. But, histograms do not suit many applications, suffering as it does from losing object information [16]. Multi-resolution histograms were meant to encode structure information through convolution of Gaussian filters, difference of Gaussians, Gabor filters and others. Histogram is a good texture image description tool. Nearest-neighbor classifier with histogram intersection recognition ensures similarity measurement performance.

A histogram of size m*n*2 with each bin alternating between representing mean μ_{mn} and standard deviation σ_{mn} is defined in Eq.(3) and Eq.(4) as [17]

$$\mu_{mn} = \iint |w_{mn}|(x, y) / \, dxdy \tag{3}$$

$$\sigma_{mn} = \sqrt{\iint (|w_{mn}(x, y) - \mu_{mn})^2 \, dxdy} \tag{4}$$

In the proposed feature extraction method, the feature vectors obtained by Gabor filter and the histogram are fused. The Gabor filter with histogram feature alternating between mean and standard deviation combination is used. Let $g_1 \in R^{n_1}$ and $h_1 \in R^{n_2}$ be the Gabor and Histogram features extracted from the MRI image. The fused feature vector $f \in R^{n_1+n_2}$ is obtained by the z-score normalization

Classifier

SVM is a linear and nonlinear data [18] classification method, using a nonlinear map to transform original training data into a higher dimension. It searches for a linear optimal separating hyper plane within the new dimension. SVM performs similar to an ANN (Artificial Neural Network) and can obtain global optimum with over fitting being controlled. SVM can also find nonlinear decision boundaries in input space [19].

SVM tries to locate a decision hyper plane which is shown in Eq.(5)

$$w.\phi(x_i) + b = 0 \tag{5}$$

where w and b are classification model parameters and Φ maps to a specific higher dimensional space where xi undergoes linear separation. The training task for the model as an optimization task is formalized as

$$\min_w \left(\|w\|^2 / 2 \right) \text{ subject to } y_i \left(w.\phi(x_i) + b \right) \geq 1$$

The convex optimization problem task is rewritten as an optimization formula to a Lagrangian function L (w, b, λ).

L (w, b, λ) and its dual form L (λ) are shown in Eq.(6) and (7).

$$L(w, b, \lambda) = \left(\|w\|^2 / 2 \right) - \sum_{i=1}^{N} \lambda_i \left(y_i \left(w.\phi(x_i) + b \right) - 1 \right) \tag{6}$$

$$L(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \lambda_I \lambda_J Y_I Y_J \phi(X_I).\phi(X_J) \tag{7}$$

subject to the Karush–Kuhn–Tucker conditions

$$\lambda_i \geq 0, \lambda_i \left\{ y_i \left(w.\phi(x_i) + b \right) - 1 \right\} = 0$$

Where λi Lagrangian multipliers are calculated by exploiting quadratic programming techniques or faster heuristic algorithms**.**

Linear support vector machine

Linear support vector machine is a popular supervised learning MVPA technique. Linear SVM ensures good classification and hence is chosen. Linear SVM is intuitive, easy to visualize and usually provides good classification, but it also has disadvantages which include SVMs being unable to give a probabilistic output. Also standard linear SVM cannot regress**.**

Polynomial support vector machine

Non-linear SVMs and linear SVMs are similar except in attempting identification of decision boundaries described by non-linear functions like polynomial function, radial basis function (RBF) or Gaussian function instead of a linear function. Hence, non-linear SVMs have one or more additional parameters connected to non-linear functions. The ‘kernel width’ parameter is an additional Gaussian function parameter affects non-linear function’s scaling.

Radial Basis Function

RBFs which are characterized by the fact that each basis function depends only on the radial distance from a centre, so θ j(x) = (‖ x- μ j ‖). RBF was introduced for exact function interpolation.

Linear K(xi,xj) = xiTxj

Polynomial K(xi,xj) = (γxiTxj + r)d, γ > 0

RBF K(xi,xj) = exp(-γ ||xi - xj||2), γ > 0

where:

γ is the gamma term in the kernel function for all kernel types except linear.

d is the polynomial degree term in the kernel function for the polynomial kernel.

γ, d, and r are user-controlled parameters, as their correct definition significantly increases the accuracy of the SVM solution.

IV. RESULT AND DISCUSSION

A program was developed to evaluate the performance efficiency of the various SVM kernels for classifying the MRI medical images for stroke classification. The images were classified as stroke and non-stroke images. Features are extracted from the MRI images using Gabor filter with histograms. The evaluation is based on the classification accuracy, Root Mean Square Error (RMSE), precision and recall. The classification accuracy is the ratio of the total number of predictions that are accurate.

Classification Accuracy (%) = (TN + TP) / (TN + FN + FP + TP)

Precision is the ratio of the predicted valid instances that were accurate.

Precision (%) = TP / (FP + TP)

Recall is the ratio of the valid instances pages that were properly identified.

Recall (%) = TP/ (FN+TP)

where TP (True Positive) = Number of correct predictions that an instance is valid

TN (True Negative) = Number of correct predictions that an instance is invalid

FP (False Positive) = Number of incorrect predictions that an instance is valid

FN (False Negative) = Number of incorrect predictions that an instance is invalid

The Root Mean Square Error (RMSE) is a measure of the difference between the values predicted by a model and the values actually observed. The RMSE of a prediction with respect to the estimated variable l is defined as the square root of the mean squared error and is shown in Eq.(8)

$$RMSE = \sqrt{\frac{I = \sum_{I=1}^{N}\left(P_{obs,i} - P_{\mathrm{mod}\,el,i}\right)^2}{n}} \quad (8)$$

Where P_{obs} is observed values and $P_{\mathrm{mod}el}$ is modeled values at time/place i.

The following Table 1 shows the summary of the experimental results for Linear, Polynomial, and RBF kernel for different feature extraction methods. Fig. 2 shows the classification accuracy obtained by various kernels.

Table 1: Classification Accuracy for Various SVM Kernels and Feature Extraction Techniques

Technique Used	**Gabor Features**	**Histogram Feature**	**Proposed Feature Extraction**
SVM with RBF kernel	55	47.5	50
SVM with linear kernel	62.5	57.5	82.50
SVM with polynomial kernel	77.5	72.5	80

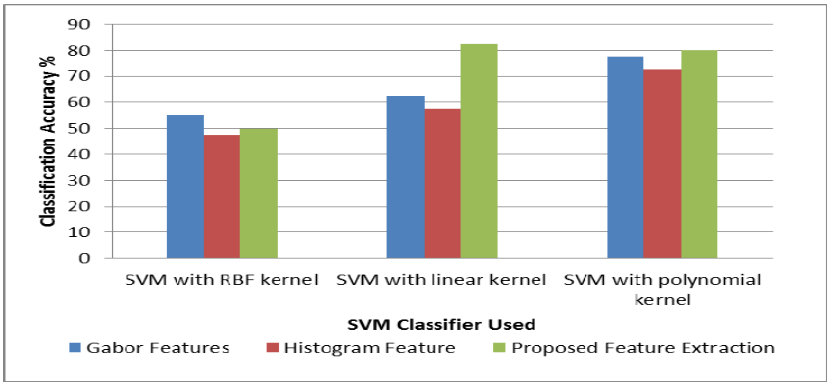

Fig.2: Classification Accuracy Obtained by Various Kernels and Feature Extraction Techniques

It is observed from Table 1 and Fig. 2 that the linear function achieves the maximum classification accuracy of 82.50%. It is also observed that the proposed feature extraction improves the performance of the classifiers with the exception of SVM-RBF kernel. The proposed feature extraction improves the performance by 3.23% to 43.48%.

Similar to classification accuracy achieved, the RMSE is also the lowest for linear kernel which is shown in Table 2 and Fig. 3

Table 2: RMSE for Various SVM Kernels and Feature Extraction Techniques

Technique Used	**Gabor Features**	**Histogram Feature**	**Proposed Feature Extraction**
SVM with RBF kernel	0.6842	0.6922	0.7071
SVM with linear kernel	0.5614	0.5842	0.4183
SVM with polynomial kernel	0.4872	0.5228	0.4472

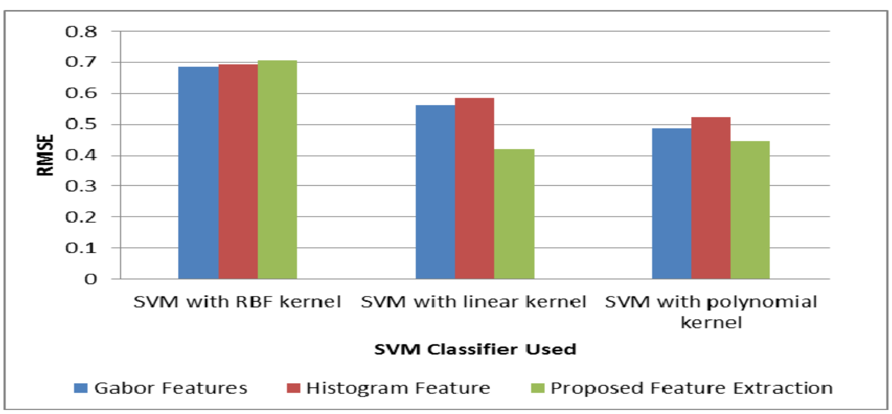

Fig. 3: RMSE Obtained by Various Kernels and Feature Extraction Techniques

From Table 3, 4 and Fig. 4,5 it is observed that the best precision and recall was achieved for linear kernel. Though, the classification accuracy achieved is satisfactory, further investigations are required to improve the classification accuracy.

Table 3: Precision of Various Kernels and Feature Extraction Techniques

Technique Used	Gabor Features	Histogram Feature	Proposed Feature Extraction
SVM with RBF kernel	0.5505	0.4749	0.5
SVM with linear kernel	0.6253	0.5752	0.826
SVM with polynomial kernel	0.7813	0.7256	0.8

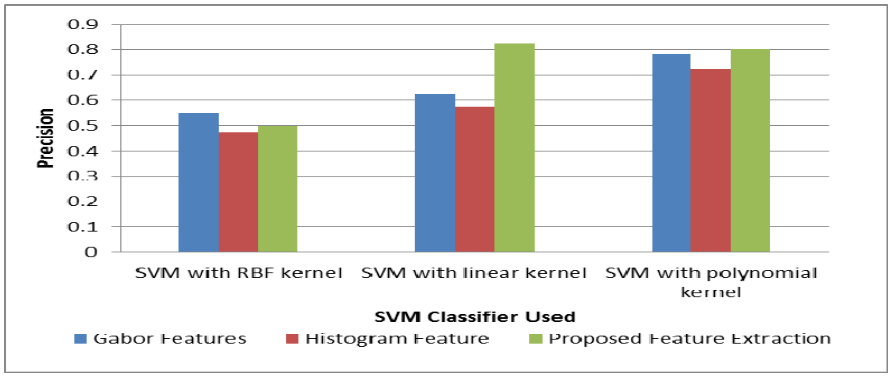

Fig. 4: Precision of Various Kernels and Feature Extraction Techniques

Table 4: Recall of Various Kernels and Feature Extraction Techniques

Technique Used	Gabor Features	Histogram Feature	Proposed Feature Extraction
SVM with RBF kernel	0.5778	0.4631	0.5
SVM with linear kernel	0.6095	0.5895	0.825
SVM with polynomial kernel	0.7196	0.7071	0.8

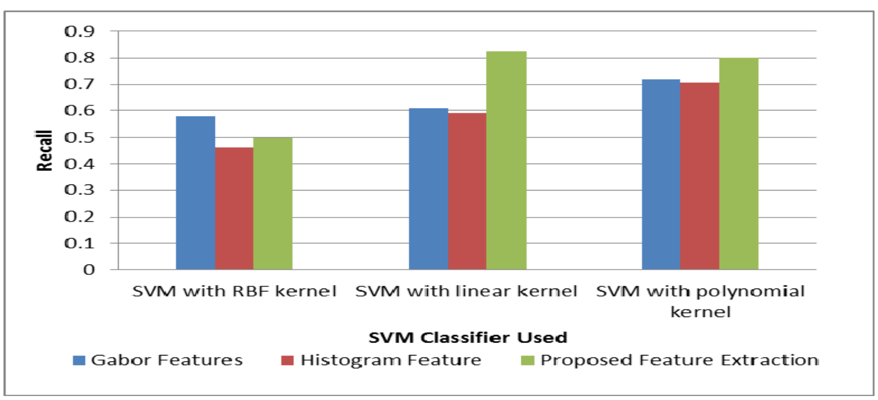

Fig 5: Recall of Various Kernels and Feature Extraction Techniques

V. CONCLUSION

This work investigates the efficiency of various SVM kernels for classifying MRI medical images for stroke classification. The Gabor filter and histograms are used to extract features from MRI medical images. The extracted features are classified using SVM kernels. The best classification accuracy of 82.50% was achieved for linear kernel. Though, the classification accuracy achieved is satisfactory, further investigations are required to improve the classification accuracy. The RBF kernel achieves lower performance due to the random value of cost function and gamma. Further investigation to optimize the parameters of RBF kernel to improve classification accuracy is to be carried out.

REFERENCES

[1] Das, S., Chowdhury, M., &Kundu, M. K. (2013). Brain Mr Image Classification Using Multiscale Geometric Analysis of Ripplet. Progress In Electromagnetics Research, 137, 1-17.
[2] Akgül, C. B., Rubin, D. L., Napel, S., Beaulieu, C. F., Greenspan, H., &Acar, B. (2011). Content-based image retrieval in radiology: current status and future directions. Journal of Digital Imaging, 24(2), 208-222.
[3] Muneeswaran, K., Ganesan, L., Arumugam, S., &RubaSoundar, K. (2006). Texture image segmentation using combined features from spatial and spectral distribution. Pattern Recognition Letters, 27(7), 755-764.
[4] Gudivada, V. N., &Raghavan, V. V. (1995). Content based image retrieval systems. Computer, 28(9), 18-22.
[5] Di Mascio, T., Frigioni, D., & Tarantino, L. (2010). VISTO: A new CBIR system for vector images. Information Systems, 35(7), 709-734.
[6] Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys (CSUR), 40(2), 5.
[7] Ueda, T., Hatakeyama, T., Kumon, Y., Sakaki, S., &Uraoka, T. (1994). Evaluation of risk of hemorrhagic transformation in local intra-arterial thrombolysis in acute ischemic stroke by initial SPECT. Stroke, 25(2), 298-303.
[8] Huang, S., Shen, Q., & Duong, T. Q. (2011). Quantitative prediction of acute ischemic tissue fate using support vector machine. Brain research, 1405, 77-84.
[9] Reddy, A., Wang, H., Yu, H., Bonates, T. O., Gulabani, V., Azok, J., ...& Li, K. C. (2008). Logical Analysis of Data (LAD) model for the early diagnosis of acute ischemic stroke. BMC medical informatics and decision making, 8(1), 30.
[10] Saur, D., Ronneberger, O., Kümmerer, D., Mader, I., Weiller, C., &Klöppel, S. (2010). Early functional magnetic resonance imaging activations predict language outcome after stroke. Brain, 133(4), 1252-1264.
[11] Carmona, R., Hwang, W. L., &Torrésani, B. (1998). Practical Time-Frequency Analysis: Gabor and Wavelet Transforms, with an Implementation in S (Vol. 9). Academic Press.
[12] J.G. Daugman, “Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression,” IEEE Transactions on Acoustics, Speech, Signal Processing, vol. 36, no. 7, pp. 1169-1179, 1988
[13] Zhang, B., Shan, S., Chen, X., &Gao, W. (2007). Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition.Image Processing, IEEE Transactions on, 16(1), 57-68.
[14] C. Schmid and R. Mohr, “Local greyvalue invariants for image retrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 5, pp. 530–534, May 1997.
[15] M. J. Swain and D. H. Ballard, “Color indexing,” Int. J. Comput. Vis., vol. 7, no. 1, pp. 11–32, 1991.
[16] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, “Multiresolution histograms and their use for recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 7, pp. 831–847, Jul. 2004.
[17] Allison, D., &Wintz, M. Related Image Retrieval using Color and Gabor Filter Histogram Featurization.
[18] Suykens, J. A., &Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.
[19] Davy, M., Gretton, A., Doucet, A., &Rayner, P. J. (2002). Optimized support vector machines for nonstationary signal classification. Signal Processing Letters, IEEE, 9(12), 442-445.