Supervised Classification Approaches for Brain Tumour Classification Using Fused Wavelet Features

Shailendra Kumar Mishra[1], Hiran Kumar Singh[2]

[1]Research Scholar, [2]Associate professor, Department of Electronics and Communication Engineerin , VelTech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

Abstract

In this study, an efficient pattern recognition technique is developed for Brain Image Classification (BIC) into normal or abnormal. Wavelet transform features with supervised classification have a potential role to play in bringing Magnetic resonance Images (MRI) of the brain into practical clinical use.The developed pattern recognition technique uses Discrete Wavelet Transform (DWT), Dual tree M-band Wavelet Transform (DMWT), and Stationary Wavelet Transform (SWT) for feature extraction, *k*-Nearest Neighbour (*k*NN) and Naive Bayes (NB) for classification and is considered as an effective and accurate tool for brain image analysis for cancer classification. Also, the predominant coefficients are chosen from the combined feature space by rank features of statistical feature selection approach. Results show that the proposed system acts as a pre-treatment predictor for BIC with an accuracy of 88.5% for *k*NN and 95.5% for NB classification.

Keyword: *Brain tumour, wavelet transforms, supervised classification, k-nearest neighbourand naive Bayes.*

Introduction

MRI is emerging as one of the most powerful tools for the study of the human brain. Recent advances in MRI technology have enabled the acquisition of brain images in high resolution and thus providing a more clinically relevant utility to study brain tumour in the context of pattern recognition. The potential of MRI images to diagnose and classify the brain tumour images has been reviewed.

DWT with three families of filters; 'bio3.7', 'db8', and 'sym8' are well analyzed for BIC[1]. It uses energy features of sub-bands extracted from each level of decomposition up to 5th level. Then, the features are analyzed independently by Support Vector Machine (SVM) classifier. A hybrid approach is described in for MRI-BIC system. The DWT sub-bands are fed to Principal Component Analysis (PCA) algorithm to reduce the dimension of feature space. Then, kNN classifier and Neural Network (NN) are designed to classify the MRI images.

Texture features from the run-length matrix and intensity features based BIC system is discussed[2]. It extracts features from the clustered tumour region after preprocessing. Then the features are reduced by PCA and singular value decomposition techniques, and a simple NN is used for classification. Feature extraction by DWT and feature reduction by PCA is employed in for BIC[3]. It uses extreme learning machine and back propagation NN for classification.

Topographic sparse coding based BIC system is implemented[4]. The gray matter of the brain is obtained after a series of preprocessing stages such as segmentation and smoothing. Then, unsupervised feature learning is employed for classification and comparison is made with PCA and self-learning NN also. A deep learning NN for BIC is described[5]. It classifies the abnormal brain image classes only. The main disadvantage of NN is high computational cost due to the increased number of layers, and also fine-tuning of NN is a time-consuming process.

Corresponding author:
Shailendra Kumar Mishra
E-mail: sant10287@gmail.com

Stockwell transform-based BIC system is described[6]. After taking the Stockwell transform, Sammon mapping is employed to reduce the feature space. NB classifier is applied to classify the MRI images. Wavelet and statistical-based features are used for BIC system[7]. It uses wavelet coefficients along with the first or second-order statistical features as features and multi-layer perceptron, random forest and NB as classifiers for brain image analysis.

Automatic BIC system using salp swarm algorithm is described[8]. It uses decision tree, kNN and NB as learning algorithms for BIC with the help of three levels vortex wavelet transform. Feature reduction is also employed by PCA. Features from the Fourier transform spectrum are employed for BIC[9]. It uses the SVM classification technique for classification after feature reduction by PCA.

In this paper, MRI-BIC system using fused wavelet features by two supervised classification approaches; *k*NN and NB is developed. For feature reduction, the system uses statistical *t*-test instead of PCA, and thus the wavelet coefficients are directly used for the classification. The organization of the paper is as follows: Section 2 discusses the methods used for MRI-BIC are described briefly. The performances of MRI-BIC system using 200 images of MRI brain images using *k*NN and NB are discussed, and the conclusion is made in the final section.

Methods and Materials

The workflow of MRI-BIC system consists of three stages; feature extraction, selection and then classification. To represent the input MRI images in multi-resolution, which provides different sub-band coefficients, they are given to DWT, DMWT, and SWT. After representation, the coefficients are selected using the ranking approach by a statistical t-test. A predefined number of predominant rank features are selected from the sub-bands of each transform coefficients according to their rank. Then the selected features are classified by KNN and NB classifier.

Feature Extraction

In this section, the three transformations; DWT[10], SWT[11] and DMWT[12] used for the extraction of features for MRI-BIC system are discussed.The wavelet analysis leads to a signal representation in time and frequency domains simultaneously. The basic idea is to use a cutting window that is shifted along with the signal and calculate the spectrum for each of these shifted positions (translation). This process is then iterated, changing the window size (dilation). The result is a collection of time-frequency representations with different resolutions, hence the definition of multi-resolution analysis. The goal of the multi-resolution analysis is to apply the divide and conquer strategy: separate the signal in multiple components that can be independently analyzed and processed by different algorithms. Let‘s define the function $\psi(t)$ called mother wavelet as follow:

$$\psi_{\tau,\sigma}(t) = \frac{1}{\sigma}\psi\left(\frac{t-\tau}{\sigma}\right) \quad (1)$$

where σ is the scale factor (dilation), τ the translation factor and ψ represents a basis function that can be designed to taste as long as it fulfils general wavelet properties such as admissibility, regularity and vanishing moments.

The main drawback of conventional wavelet is that it has shift-invariant against the input samples. Thus the size of the sub-bands obtained from the conventional wavelet is not equal to the size of the input data. This is due to the downsampler in the decomposition filter. To remove the shift-invariant of conventional wavelet, SWT[11] is designed by just removing the downsampler. Thus, the resolution of the sub-bands does not change. Though the wavelets are multi-resolution in nature, it provides information in three directions; vertical, horizontal and diagonal only. To get more information from other directions, DMWT is used. The decomposition is obtained by using *the M*-band filter bank and dual *M*-band filter bank, and then the linear combinations of the sub-bands are obtained with more information[12]. After decomposition by DWT, DMWT and SWT, the coefficients in each sub-band are considered as features. In the later stage, dominant coefficients are selected for the further classification process.

Feature Selection

After representation, the feature space of each transform becomes high (except DWT) as they provide more information in different sub-bands. It is well known that not all the wavelet coefficients carry

significant information, and thus a feature selection approach is needed. Without feature selection, the use of all coefficients reduces system performance. After representation, the coefficients are selected using statistical *t*-test. The *t*-test is based on the average of the features extracted from normal and abnormal MRI images which is given by Eqn. (2)

$$t(x) = \left(\bar{y}_1(x) - \bar{y}_2(x)\right) / \sqrt{\left(s_1^2(x)/n_1 + s_2^2(x)/n_2\right)} \quad (2)$$

where $\bar{y}_1(x)$ is the means of normal, $\bar{y}_2(x)$ is the means of abnormal, $s_1^2(x)$ is the standard deviations of normal and $s_2^2(x)$ is standard deviations of abnormal features.

A predefined number (N) of predominant rank features are selected from the sub-bands of each transform coefficients according to their rank. In this study, three different sets of features are computed and named as FV-I-25, FV-II-50, and FV-III-75. The first variable FV indicates the feature vector and the second variable denotes the feature set number, and the last variable indicates the number of features selected from each transform. These set of features are the inputs to the classification system for BIC.

Classification

Machine learning is often applied to problems where it is difficult to construct a physical model of the processes generating the data but where a reasonable number of examples are available. In this study, two classifiers are such as *k*NN[13] and NB[14] are used for classification. *k*NN is a similarity-based approach based on the minimum distance measure between the testing samples and the training samples. For *k*NN, *k* numbers of training samples (i.e. nearest neighbours) are used to label unknown samples according to the distance measure. The distance metric can be any similarity measure statistics. This study uses Euclidean distance in Eqn. 3 as a distance measure with *k*=1. The common advantages of KNN are that it is computationally simple, less prone to noise and bias. The Euclidean distance between two points $u = (x_1, y_1)$ and $v = (x_2, y_2)$ is given by

Euclidean distance

$$(u, v) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (3)$$

An NB classifier is a statistical-based approach based on the elementary Bayes Theorem with strong independence assumptions, i.e. all samples in the data set are unrelated to each other. Given a *N*umber of classes in the data set and *f* is an observation (features) in the data set with a known probability density of the class $p(f_N)$. If the class prior probability $p(N)$ is known, then the posterior probability $p(N|f_N)$ of the observation can be expressed as:

$$p(N|f_N) = \frac{p(f_N)}{p(N)} . \frac{L}{p(f_N)} \quad (4)$$

Where, *L* is the predefined number of labeled observations used to measure the likelihood of the observation, irrespective of their class labels.

Results and Discussions

To simulate the proposed BIC system discussed in section 2 and analyzes the performance of the system, REpository of Molecular BRAin Neoplasia DaTa (REMBRANDT) database[15] images are used. MRI images are acquired from 130 subjects with a size of 256x256 pixels[16]. From the database, 100 images per category (normal and abnormal) are selected. Figure2 shows sample normal images (top row) and abnormal images (bottom row) in the REMBRANDT database[17].

To predict the abnormality, a training set composed of both normal and abnormal MRI images (the inputs) and the corresponding label to the images (the outputs) is required. These are prepared with the help of *k*-fold cross-validation with *k* equal to 10. This approach validates the classifier performance with different training, and testing samples and also all the samples undergo into training and testing phase as well. At first, the images in each category (class) are divided into 10-folds with an equal number of images in each group. Each iteration images in one fold are used to test the classifiers, and the images in remaining folds are used to train the classifiers[18]. This procedure is repeated for 10-iterations so that all the images are used to test the classifiers.

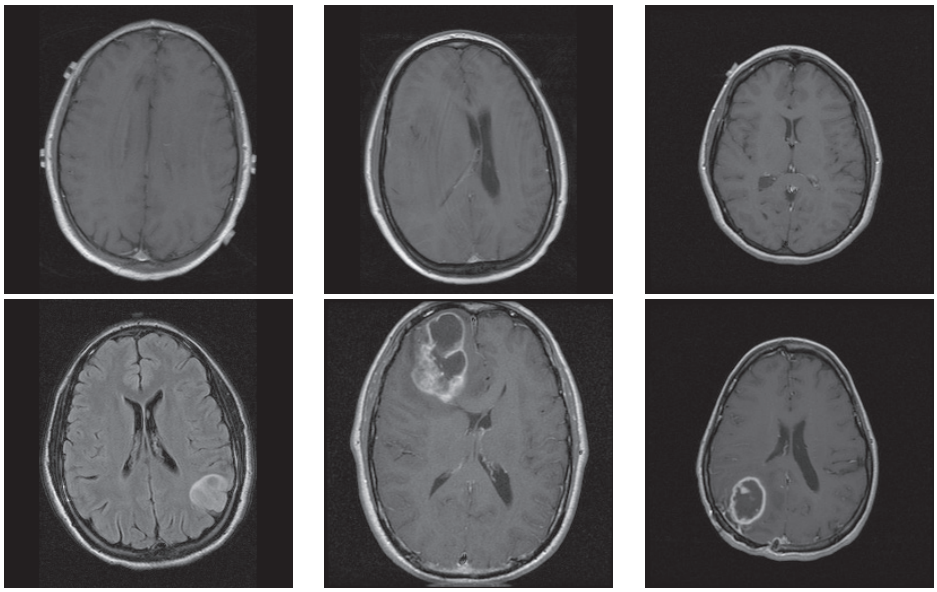

Figure 2 REMBRANDT database images

In each iteration, the correctly and misclassified number of images are noted and finally, the total number of correctly and misclassified images at each category are calculated, and the following performance metrics such as sensitivity, specificity and accuracy are computed to analyze the performance of the MRI-BIC system. The definitions are as follows:

- **Sensitivity:** It gives the correct classification rate of abnormal images and is given below:

$$Sensitivity = \frac{P}{P + N} \quad (4.1)$$

Where, True Positive (TP) and False Negative (FN) are the correct and misclassified abnormal images respectively.

- **Specificity:** It gives the correct classification rate of normal images and is given below:

$$Specificity = \frac{N}{N + P} \quad (4.2)$$

Where, True Negative (TN) and False Positive (FP) are the correct and misclassified normal images respectively.

- **Accuracy:** It gives the overall classification rate of the system and is given below:

$$Accuracy = \frac{P + N}{P + N + N + P} \quad (4.3)$$

Tables 1 to 3 display the evaluations of BIC system in terms of accuracy, sensitivity and specificity which are computed using the above formula.

Table 1 MRI-BIC system performance in terms of accuracy

Different wavelet transform (DWT, SWT and DMWT) Levels	Accuracy (%)					
	FV-I-25		FV-II-50		FV-III-75	
	kNN	NB	kNN	NB	kNN	NB
1	73.5	82.5	71	78.5	76	84.5
2	80	88	88.5	95.5	83.5	90
3	76	85	68	75	77	83
4	65.5	73.5	66.5	74	74.5	82.5
5	78	85	72.5	80.5	80.5	89

Table 2 MRI-BIC system performance in terms of sensitivity

Different wavelet transform (DWT, SWT and DMWT) Levels	Sensitivity (%)					
	FV-I-25		FV-II-50		FV-III-75	
	kNN	NB	kNN	NB	kNN	NB
1	80	90	75	82	76	84
2	80	89	89	96	84	90
3	80	88	84	92	85	90
4	79	87	78	86	80	88
5	82	90	82	90	80	90

Table 3 MRI-BIC system performance in terms of specificity

Different wavelet transform (DWT, SWT and DMWT) Levels	Specificity (%)					
	FV-I-25		FV-II-50		FV-III-75	
	kNN	NB	kNN	NB	kNN	NB
1	67	75	67	75	76	85
2	80	87	88	95	83	90
3	72	82	52	58	69	76
4	52	60	55	62	69	77
5	74	80	63	71	81	88

It can be seen from Table 1 that the FV-II-50 set produces the highest accuracy of 95.5% while using the features of different wavelets from 2ndlevel of decomposition by NB classifier when compared with FV-I-25 and FV-III-75 set. For the same set of features, kNN gives an accuracy of 88.5%. It is noted that the performance of MRI-BIC system is improved while increasing the predominant coefficients from each wavelet transform from 25 to 50 features. Further increase in feature dimension reduces the performance as it adds redundant features in the feature space.

Conclusion

A pattern recognition technique for BIC system is presented using three different wavelet transforms; DWT, DMWT and SWT. It also highlights the potential role of *k*NN and NB in developing pattern recognition applications for the classification. The automatic feature

selection can infer relationships between normal and abnormal MRI images. The following feature sets such as FV-I-25, FV-II-50 and FV-III-75 are employed for performance evaluation. These sets are extracted for each level of decomposition and each wavelet transform as well. Then they are given to the *k*NN and NB for classification independently. Results show that the FV-II-50 set produces the highest accuracy of 88.5% (*k*NN) and 95.5% (NB) at 2nd level combined wavelet features.

Ethical Clearance- NA

Source of Funding- Self

Conflict of Interest –Nil

References

1. Mohankumar S. Analysis of different wavelets for brain image classification using support vector machine. International Journal of Advances in Signal and Image Sciences. 2016 Jun 30;2(1):1-4.
2. El-Dahshan ES, Hosny T, Salem AB. Hybrid intelligent techniques for MRI brain images classification. Digital signal processing. 2010 Mar 1;20(2):433-41.
3. Al-Saffar ZA, Yildirim T. A Novel Approach to Improving Brain Image Classification Using Mutual Information-Accelerated Singular Value Decomposition. IEEE Access. 2020 Mar 13;8:52575-87.
4. Satapathy P, Pradhan SK, Hota S. An Empirical Performance Analysis of Brain Image Classification Models Using Variants of Neural Networks. In2019 International Conference on Applied Machine Learning (ICAML) 2019 May 25 (pp. 87-91). IEEE.
5. Wang X, Wang W. MRI Brain Image Classification Based on Improved Topographic Sparse Coding. In2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS) 2018 Nov 23 (pp. 1116-1119). IEEE.
6. Hemanth DJ, Anitha J, Naaji A, Geman O, Popescu DE. A modified deep convolutional neural network for abnormal brain image classification. IEEE Access. 2018 Dec 10; 7:4275-83.
7. Saminathan K. MRI Brain Image Classification Based on S-Transform, Sammon Mapping and Naïve Bayes Classifier. International Journal of Innovative Technology and Exploring Engineering, 2019 October; 8 (12): 790-793.
8. Latif G, Iskandar DA, Alghazo JM, Mohammad N. Enhanced MR image classification using hybrid statistical and wavelets features. IEEE Access. 2018 Dec 18;7:9634-44.
9. Nejad MB, Shiri ME. A new enhanced learning approach to automatic image classification based on Salp Swarm Algorithm. Computer Systems Science and Engineering. 2019 Mar 1;34(2):91-100.
10. Zhang YD, Chen S, Wang SH, Yang JF, Phillips P. Magnetic resonance brain image classification based on weighted□type fractional Fourier transform and nonparallel support vector machine. International Journal of Imaging Systems and Technology. 2015 Dec;25(4):317-27.
11. Stephane M. A wavelet tour of signal processing, Academic Press; 2008.
12. Fowler JE. The redundant discrete wavelet transform and additive noise. IEEE Signal Processing Letters. 2005 Aug 15;12(9):629-32.
13. Chaux C, Duval L, Pesquet JC. Image analysis using a dual-tree M-band wavelet transform. IEEE Transactions on Image Processing. 2006 Jul 17;15(8):2397-412.
14. Hall P, Park BU, Samworth RJ. Choice of neighbor order in nearest-neighbor classification. the Annals of Statistics. 2008;36(5):2135-52.
15. Bolstad WM, Curran JM. Introduction to Bayesian statistics. John Wiley & Sons; 2016 Oct 3.
16. Scarpace L, Flanders A, Jain R, Mikkelsen T, Andrews DW. Data from REMBRANDT. The cancer imaging archive, http://doi.org/10.7937/K9/TCIA.2015.588OZUZB.
17. REMBRANDT: https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT
18. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging. 2013 Dec 1;26(6):1045-57.