ISSN 2229-5518

An Efficient Technique of MRI Image Classification using PSO-SVM

Priti Sahu, Prof. Sunita Gond

Abstract— With the rapid development of the medical science more and more medical images are generated rapidly like MRI, CT scan, X-ray etc. Due to that an efficient system is essential for the indexing, storing and analyzing such images. The analyzing cost of such images is very high. The analysis quality also differs and highly prone to errors. The classification of such images is a quite harder job. Hence it is essential to develop a strong system for classification of such images. The analysis of the medical images provides a way of detecting and predicting diseases in the images. Since various techniques are implemented for the analysis of images containing diseases. The existing technique implemented for the disease classification using manifold learning provides efficient detection and classification of diseases in MRI images [1]. But the technique implemented for the disease classification is based support vector machine based classifier which is less efficient and contains more error rate. Hence an efficient technique is implemented for the disease classification by the optimization of support vector machine using particle swarm optimization. The proposed technique implemented here provides efficient classification and detection of diseases in the image.

Index Terms— Minimum 7 keywords are mandatory, Keywords should closely reflect the topic and should optimally characterize the paper. Use about four key words or phrases in alphabetical order, separated by commas.

1 Introduction

Many remote sensing systems record brightness values at different wavelengths that commonly include not only portions of the visible light spectrum, but also photo infrared and, in some cases, middle infrared bands. The brightness values for each of these bands are typically stored in a separate grayscale image (raster). Each ground-resolution cell in an image therefore has a set of brightness values which in effect represent the "color" of that patch of the ground surface, if we extend our concept of color to include bands beyond the visible light range.

The Automatic Classification process in TNTmips uses the "colors", or **spectral patterns**, of raster cells in a multispectral image to automatically categorize all cells into a specified number of spectral classes. The relationship between spectral classes and different surface materials or land cover types may be known beforehand, or determined after classification by analysis of the spectral properties of each class. The Automatic Classification process offers a variety of classification methods as well as tools to aid in the analysis of the classification results.

The spectral pattern of a cell in a multispectral image can be quantified by plotting the raster value from each band on a separate coordinate axis to locate a point in a hypothetical "spectral space". This spectral space has one dimension for each band in the image. Most classification methods use some measure of the distance between points in this spectral space to assess the similarity of spectral patterns. Cells that are close together in spectral space have similar spectral properties and have a high likelihood of imaging the same surface features.

MRI Images

Magnetic resonance imaging (MRI), nuclear magnetic resonance imaging (NMRI), or magnetic resonance tomography (MRT) is a medical imaging technique used in radiology to investigate the anatomy and physiology of the body in both health and disease. MRI scanners use strong magnetic fields and radiowaves to form images of the body. The technique is widely used in hospitals for medical diagnosis, staging of disease and for follow-up without exposure to ionizing radiation.

MRI has a wide range of applications in medical diagnosis and there are estimated to be over 25,000 scanners in use worldwide [1]. MRI has an impact on diagnosis and treatment in many specialties although the effect on improved health outcomes is uncertain [2]. Since MRI does not use any ionizing radiation its use is recommended in preference to CT when either modality could yield the same information [3]. MRI is in general a safe technique but the numbers of incidents causing patient harm have risen [4]. Contraindications to MRI include most cochlear implants and cardiac pacemakers, shrapnel and metallic foreign bodies in the orbits, and some ferromagnetic surgical implants. The safety of MRI during the first trimester of pregnancy is uncertain, but it may be preferable to alternative options [5]. The sustained increase in demand for MRI within the healthcare industry has led to concerns about cost effectiveness and over diagnosis [6, 7].

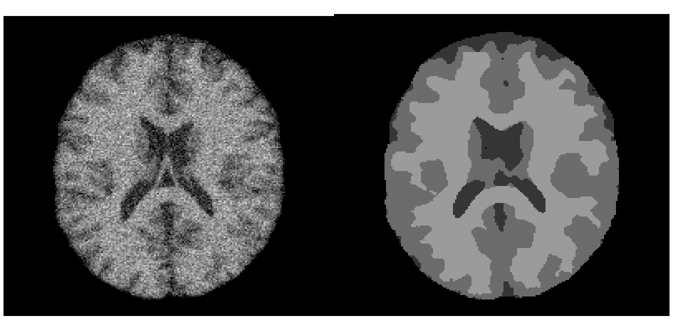

Figure 1. MRI Images

ISSN 2229-5518

2 LITERATURE SURVEY

A novel approach was proposed to increase the number of classes with a higher accuracy rate by iterative filtering on the training dataset by Zare et al [1]. Filtering is done according to their classification performance. They presented a novel method to achieve classification of class of Image CLEF 2007 medical database [6]. In this scheme they have four iterations steps. These steps hold different classification models. Within the iteration generation process was performed in two steps. The construction of a model from the entire dataset was the first step. This was used to assess filter high accuracy classes (HAC). This will achieve accuracy of 80 % and using this process they can train 20% of the data set. The classes under HAC were only used to construct the classification model under second step. These steps are also continued to next iterations [1].

The classification process consists training phase and testing phase. The training phase was used to selected features that are extracted from all the training images and the classifier is trained on the extracted features to create a classification model. The training phase was also responsible for feature extraction and model generation. In feature extraction they used bag of visual word (BoW). BoW was used for extracting feature using detect and extract local features, codebook construction and represent an image in a histogram using the codebook. Extraction of BoW started with detecting and extracting local features. Feature detection is the process in which a spatially limited image region that is salient or prominent must be identified. The next step in implementation of BoW is codebook construction where it is built using clustering or the vector quantization algorithm. This step usually uses k-means clustering method, and uses cluster centre as visual vocabulary term. The goal of codebook construction is to identify a set of visual patterns that reflects the image collection contents. Upon identification of cluster centres, each image is represented as histograms of these cluster centres by simply counting the frequency of the words appearing in an image [1].

Upon extraction of BoW representation from the training dataset, it was then used as input to support vector machine (SVM) classifier to construct the model. This can evaluate in the training phase itself to ensure that the best possible accuracy rate is attained for every individual class in the database. Eighty percent of it was used to construct the classification models, and the remaining 20% of the training data were taken for test images for evaluation purpose of the generated model. Only one of the classification models was constructed in training phase for testing. This will be applied on the test image in order to classify it into the predefined category. Classification process on the unseen test image starts with identifying the respective classification model for that particular image. If right model was not chosen than it may cause decreasing accuracy rate. For those classes which are visually similar, SVM or any other classification technique would be biased to the classes with a larger number of training images. This examination shows that depending on only one technique to gain high accuracy for every individual class of such a database with the said complexity is unreliable [1].

Automatic classification of medical X-ray images: hybrid generative-discriminative approach was also proposed by Zare et al [2]. Along with rapid progress in the application of local descriptor in pattern recognition, computer vision and image retrieval, the bag of word (BoW) approach has appeared promising for object classification and image retrieval. The classification task begins with extracting appropriate features of the image. It is one of the most important factors in design process of such system. Moreover, the feature extraction step affects all other subsequent processes. The probabilistic latent semantic analysis (PLSA) [7] has been proposed to learn co-occurrence information between elements in the vector space in an unsupervised manner to disambiguate the BoW representation. PLSA can help to disambiguate visual words because of the ability of the PLSA model to generate a robust, high level representation and low-dimensional image representation since PLSA introduces a latent,

In recently Selvi and Kavitha [3] developed an efficient and powerful medical search engine to classify and search the radiographic medical images. They used differnet feature extraction method to develop this system. The images that are used was categorised in two groups one is Low-level image representation and second is local patch-based image representation. In first category Gray Level Co-occurrence Matrix (GLCM), Canny Edge Operator, Local Binary Pattern (LBP) and pixel value type's images were used. While in second category Bag of Words (BoW) was used. They also offered a novel image indexing and retrieval algorithm using local tetra patterns (LTrPs) for content-based image retrieval (CBIR). The standard local binary pattern (LBP) and local ternary pattern (LTP) encode the relationship between the referenced pixel and its surrounding neighbors by computing gray-level difference. This method encodes the relationship between the referenced pixel and its neighbors, based on the directions that are calculated using the first-order derivatives in vertical and horizontal directions. For this they chosen a generic strategy to compute th-order LTrP using th-order horizontal and vertical derivatives for efficient CBIR and analyze the effectiveness of our proposed algorithm by combining it with the Gabor transform [3].

In year 2010, Dan et al [1] presented A Synthesized Data Mining Algorithm Based on Clustering and Decision Tree. They improve the traditional algorithms like CURE and C4.5 appropriately, and present a new synthesized algorithm CA for mining large-scale high dimensional datasets. The basic idea of CA is shown as follows: first introduce PCA to analyze the relevancy between features and replace the whole dataset with several countable composite features; then improve CURE to part the set into several clusters which can be the pretreatment of other algorithms and achieve the reduction of sample scale; finally introduce parallel processing into C4.5 to enhance the efficiency of building decision tree. C4.5 only can covenant with the dataset that has the classification feature. The dataset treated is a little small which will impact the final output of algorithms [4].

Manish Verma et al [5] proposed A Comparative Study of Various Clustering Algorithms in Data Mining. They provide a comparative study among various clustering. They compared six types of clustering techniques- k-Means Clustering, Optics, DBSCAN clustering, Hierarchical Clustering, Density Based Clustering and EM Algorithm. Such clustering methods are implemented and analyzed using a clustering tool WEKA. Performances of the 6 techniques are presented and compared. Running

ISSN 2229-5518

the clustering algorithm using any software produces almost the same result even when changing any of the factors because most of the clustering software uses the same procedure in implementing any algorithm [5].

In year 2012, Bhaskar N. Patel et al [8] presented an Efficient Classification of Data Using Decision Tree. They survey many techniques related to data mining and data classification techniques. They also select clustering algorithm k-means to improve the training phase of Classification. The Learning classification techniques in data mining can be classified into three fundamental types; first one is supervised second one is unsupervised and finally third one is reinforced.

It can handle hi-dimensional data. The classification and learning steps of decision tree induction are simple and fast. Their representation of acquired knowledge in tree form is easy to assimilate by users. The decision tree algorithm is a top-down induction algorithm. The mean of this algorithm is to construct a tree that has leaves that are harmonized as potential. The most important step of this algorithm is to carry on dividing leaves that are not homogeneous into leaves that are as homogeneous as possible. Once the result obtained, it can be reused for next research. This research depicts on compares reformulated decision tree with standard decision tree for dataset. This comparison is from threshold (complexity) from low to high with reference to the testing accuracy [8].

In same year Gothai, E. and P. Balasubramanie proposed An Efficient Way for Clustering Using Alternative Decision Tree [9]. Their study proposes a Multi-Level Clustering mechanism using alternative decision tree algorithm that combines the advantage of partition clustering, hierarchical clustering and incremental clustering technique for rearranging the most closely related object. The clustering initiation should happen based on the short name value, each short name pointing to the appropriate whole record object.

ADTree divide the data based on short name; if cluster is already available with the short name then insert a record into the same cluster else create a new cluster with the new name of short name then insert into a new cluster. In every cluster, sub-set diminutive name points to the whole record. The cluster formation method mainly focus on form a similarity value in single group, for this purpose they are using different method and result of each method is different cluster based on data and spread condition.

A novel feature extraction framework for medical x-ray images classification is proposed by Ghofrani et al [10]. As per this scheme, extract centre symmetric local binary patterns from local part of shape and directional information extracted from images to achieve a set of capable features after some preprocessing. This method worked in three: preprocessing, feature extraction and classification process. Preprocessing is used to eliminate the effects of noises and also manage grey level variation along with set of capable features. After this feature were extracted local parts of each image. This can be done in three parts.

Avni et al [11] presented an efficient image categorization and retrieval system for medical image database especially for radiographic images. They presented a patch-based classification system that has demonstrated very strong classification rates while also providing efficiency in the retrieval process. This scheme was composed of a feature extraction phase, a dictionary construction based on the training archive, an image representation phase and a classification phase.

In year 2010, a learning-based algorithm for automatic medical image annotation based on sparse aggregation of learned local appearance cues was suggested by Tao et al [12]. They adopted a hybrid approach based on robust aggregation of learned local appearance findings, followed by the exemplar-based global appearance filtering. This scheme is used to detect multiple focal anatomical structures within the medical image. It detects multiple focal anatomical structures within the medical image. This is achieved via learning-by-example landmark detection algorithm.

3 Proposed Methodology

1. Take an input dataset of disease images.
2. Apply DCT transformation to divide into regions.
3. On each of the region perform disease sensitive descriptor.
4. Now apply PSO-SVM classification approach.
5. Fusion is performed for each of the region and provides labeling.

The following algorithm is used for the optimization of SVM.

1. Initialize max-iterations and number of particle and dimensions.
2. for i= 1:no_of_particles
3. for j= 1:dimensions
4. particle_position(i,j) = rand*10;
5. particle_velocity(i,j) = rand*1000;
6. p_best(i,j) = particle_position(i,j);
7. end
8. end
9. for count = 1:no_of_particles
10. p_best_fitness(count) = -1000;
11. end
12. for count = 1:max_iterations
13. for count_x = 1:no_of_particles
14. x = particle_position(count_x,1);
15. y = particle_position(count_x,2);
16. ker = '@linearKernel';
17. global p1 ;
18. p1 = x;
19. C = y;
20. trnX=X;
21. trnY=Y;
22. tstX=X';
23. tstY=Y';
24. [nsv,alpha,bias] = svmTrain(trnX,trnY,C);
25. actfunc = 0;
26. predictedY = svcoutput(trnX,trnY,tstX,ker,alpha,bias,actfunc);
27. Result = ~abs(predictedY)
28. Percent = sum(Result)/length(Result)

ISSN 2229-5518

```
soln = 1-Percent
    if soln~=0
  current_fitness(count_x) = 1/abs(soln)+0.0001;
else
  current_fitness(count_x) =1000;
end
End
```

4 Result Analysis

The table shown below is the analysis and comparison of existing and the proposed methodology on the basis of three images.

5 Images	Existing Work	Proposed Work
1	0.89	0.96
2	0.92	0.963
3	0.91	0.97

Table 1. Analysis & Comparison of Accuracy

The table shown below is the analysis and comparison of existing and the proposed methodology on the basis of three images on elapsed time in sec.

Images	Existing Work	Proposed Work
1	2.56	1.265
2	3.67	1.428
3	3.41	1.732

Table 2. Analysis & Comparison of Elapsed Time

The figure shown below is the analysis and comparison of existing and the proposed methodology on the basis of three images.

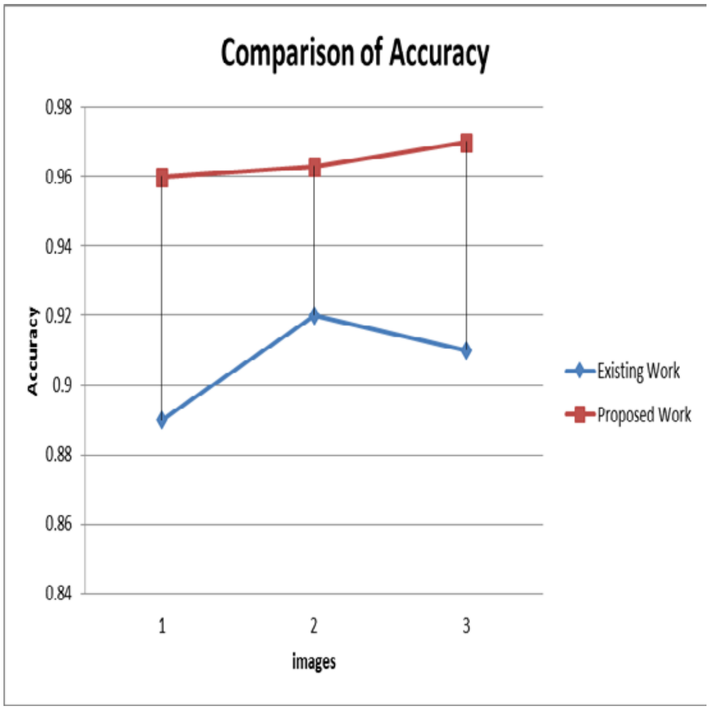

Figure 2. Comparison of Accuracy

The figure shown below is the analysis and comparison of existing and the proposed methodology on the basis of three images on elapsed time in sec.

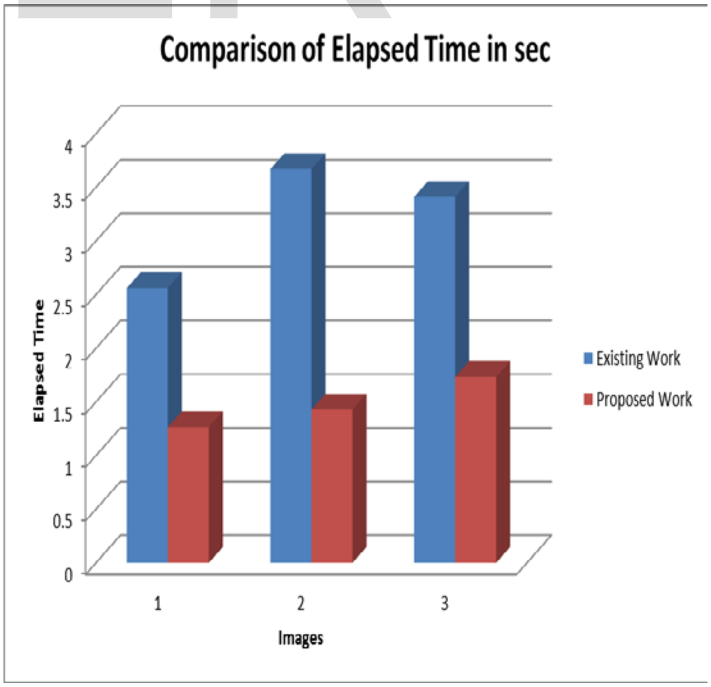

Figure 3. Comparison of Elapsed Time

4 Conclusion

The proposed methodology implemented here for the classification of MRI images using particle swarm optimization based support vector machine is efficient as compared to the existing technique implemented for the MRI Images Classification. The experimental results performed on various MRI Images shows the improvement of the proposed methodology.

Although the technique implemented here is efficient in terms of accuracy but further enhancements can be done for the improvement of the proposed methodology such as using some other supervised learning approach.

References

[1] Zare, Mohammad Reza, Ahmed Mueen, and Woo Chaw Seng. "Automatic classification of medical X-ray images using a bag of visual words." IET Computer Vision, vol. 7, no. 2, pp. 105-114, 2013.

[2] Zare, Mohammad Reza, Ahmed Mueen, Mohammad Awedh, and Woo Chaw Seng. "Automatic classification of medical X-ray images: hybrid generative-discriminative approach." IET Image Processing, vol. 7, no. 5, pp. 523-532, 2013.

[3] Selvi, S. Malar, and C. Kavitha. "Radiographic medical image retrieval system for both organ and pathology level using bag of visual words." International Journal of Engineering Sciences & Emerging Technologies, ISSN: 22316604, vol. 6, no. 4, pp. 410 – 416, 2014.

[4] Ji Dan, Qiu Jianlin, Gu Xiang, Chen Li, and He Peng “A Synthesized Data Mining Algorithm Based on Clustering and Decision Tree”, IEEE 10th International Conference on Computer and Information Technology (CIT), pp. 2722 – 2728, 2010.

[5] Manish Verma, Mauly Srivastava, Neha Chack, Atul Kumar Diswar, Nidhi Gupta “ A Comparative Study of Various Clustering Algorithms in Data Mining”, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, Vol. 2, pp.1379-1384, Issue 3, May-Jun 2012.

[6] Müller, Henning, Thomas Deselaers, Thomas Deserno, Paul Clough, Eugene Kim, and William Hersh. "Overview of the ImageCLEFmed 2006 medical retrieval and medical annotation tasks." In Evaluation of Multilingual and Multi-modal Information Retrieval, pp. 595-608, Springer Berlin Heidelberg, 2007.

[7] Hofmann, Thomas. "Unsupervised learning by probabilistic latent semantic analysis." Machine learning 42, no. 1-2 (2001): 177-196.

[8] Bhaskar N. Patel, Satish G. Prajapati and Dr. Kamaljit I. Lakhtaria “Efficient Classification of Data Using Decision Tree”, Bonfring International Journal of Data Mining, ISSN 2277 – 5048, Vol. 2, No. 1, pp. 6- 12, March 2012.

[9] Gothai, E. and P. Balasubramanie “An Efficient Way for Clustering Using Alternative Decision Tree”, American Journal of Applied Sciences, ISSN 1546-9239, vol. 9, no. 4, pp. 531-534, 2012.

[10] Ghofrani, Fatemeh, Mohammad Sadegh Helfroush, Habibollah Danyali, and Kamran Kazemi. "Medical X-ray image classification using Gabor-based CS-local binary patterns." In Proceedings International Conference on Electronics, Biomedical Engineering and its Applications, pp. 7-8, 2012.

[11] Avni, Uri, Hayit Greenspan, Eli Konen, Michal Sharon, and Jacob Goldberger. "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words." IEEE Transactions on Medical Imaging,vol. 30, no. 3, pp. 733-746, 2011.

[12] Tao, Yimo, Zhigang Peng, Bing Jian, Jianhua Xuan, Arun Krishnan, and Xiang Sean Zhou. "Robust learning-based annotation of medical radiographs." In Medical Content-Based Retrieval for Clinical Decision Support, pp. 77-88, Springer Berlin Heidelberg, 2010.

[13] Deselaers, Thomas, and Hermann Ney. "Deformations, patches, and discriminative models for automatic annotation of medical radiographs." Pattern Recognition Letters, vol. 29, no. 15, pp. 2003-2010, 2008.

[14] Selvi, S. Malar, and Mrs C. Kavitha. "Content Based Medical Image Retrieval System (CBMIRS) Using Patch Based Representation." IOSR Journal of Computer Engineering (IOSR-JCE), ISSN: 2278-0661,pp. 24-36, 2013.

[15] Foncubierta-Rodríguez, Antonio, Alba García Seco de Herrera, and Henning Müller. "Medical image retrieval using bag of meaningful visual words: unsupervised visual vocabulary pruning with PLSA." In Proceedings of the 1st ACM international workshop on Multimedia indexing and information retrieval for healthcare, pp. 75-82, ACM, 2013.