ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Review on Medical Image Classification

Venkatesh N[1], A Rafega Beham[2], Dr.Jitendranath Mungara[3]

B. E Student, Department of ISE, New Horizon College of Engineering, Karnataka, India[1]

Senior Assistant Professor, Department of ISE, New Horizon College of Engineering, Karnataka, India[2]

HOD, Department of ISE, New Horizon College of Engineering, Karnataka, India[3]

ABSTRACT: Image Classification in Image processing field is used for analysis and also to extract information. Image Classification methodologies is used in various real world applications. In the same application field, different classification methods are used. In this paper we discuss the various classification methods used in the the medical field and we discuss which classification technique has the higher recognition rate. Various methods like PCA(Principle Component Analysis),SVM(Support Vector Machine),KNN(K-Nearest Neighbour) are used.In each classification method we discuss which features are chosen to classify the images.

KEYWORDS: SVM, PCA, KNN, Edge based, Sparse, ODL.

I.INTRODUCTION

Image Classification is one of the disciplines in Image processing which is having various aspects and features. Image Classification is used in various fields like vehicle image classification,soil type classification, biomedical imaging, remote sensing and many more.In each of these fields various classification methodologies are applied like ANN(Artificial Neural Networks),SVM(Support Vector Machines),PCA(Principle Component Analysis) etc.By applying any one of the method in the particular fields will classify images completely. In recent times,with the advent of high processing capabilities of computers, high and low priced video cameras has generated an intrest in object classification algorithms. In Classification flow diagram shown below we have the following steps, Image Aquisition(Input image),pre-processing, Object detection,feature extraction and object classification.

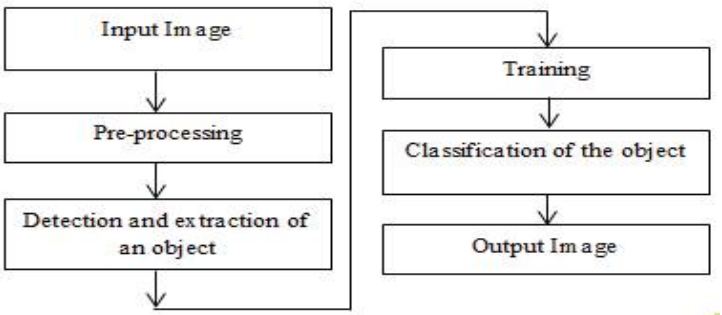
1)Image Acquisition: Input Image is actually obtained from source.Ususally in this step,usually we obtain the image but we do not perform any operations to remove noise in the image.

2)Pre-Processing:The aim of pre-processing the image is to enhance the feature of the image,so unwanted distortions like noise is removed.

3)Object detection and extraction: It includes recognizing the position or shape of moving object obtained from the camera,and in extraction it will extract the important feature of the detected object in the image.

4)Training:In image processing, training is used for classifying pixels in order to segment different objects.

5)Classification of the Object :Classification steps classify revealed objects into predefined modules by means of image classification techniques that evaluates the image patterns with the target patterns.The following is a classification diagram overflow.



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

In this paper we will discuss the various classification techniques used in medical image classification and we will analyze which technique has higher recognition rate.

II. RELATED WORK

MEDICAL IMAGE CLASSIFICATION

The remarkable progress in digital image processing and the retrival techniques, is used in one of the field like medicine for practical process.There is also rapid development in the medical devices.The information found in these images has to be effectively organized for efficient search and for further retrival and analysis of data. After retrival of data in these images it is used for diagnosis purposes.In the following section we will discuss the various methods used in the applied field in medical image classification.

a)SVM and PCA classifiers:

In medical image classification method usually we use ImageCLEF which is the open data platform and the database used is ImageCLEF 2005.This database is a standard one for medical image retrival which has 9,000 training images.These training images corresponds to 57 classes.1,000 unlabeled radiographs is the testing dataset which is used in this method .In SVM and PCA classifiers we need to construct a multi-class classifier for classifying the test images and label the test images using training dataset.But there are few problems regarding the classification of the medical images.They are listed below:

1)Great unbalance between classes:One particular class say class 1 may have 500 samples and other class say class 2 may have 2000 samples.So there is a great unbalance between these classes.

2)Visual similarities between some classes:In some cases the medical images have the need to compare from different sources.Even some of the experts may not find the visual differences among the images.

In order to face these problems SVM and PCA are used as classifier methods.

The Feature sets of an image are color,texture,shape, layout. Layout feature has low-resolution-pixel-map (LRPM).LRPM is the perfect or corresponding spatial position of the color.Middle-level regional features like Blob is used in medical image retrieval.Blob parameters are color, texture, area, length of long and short axes, rotation angle, Fourier decomposition parameters, etc.Since there are various features for medical images retrival to be considered the best way to select the feature sets is by conducting simulation experiments which will lead to better classification results.SVM is one of the statistical learning algorithm and it is used for binary classification problem.In SVM when it is used for multiple classes classification two kinds of techniques are followed. First technique is to include multiple class labels rightly to the quadratic solving algorithm. Another type of approach is to incorporate several binary classifiers where we use SVM^{Tourch}, which is in close connection to the Kernel.The Kernel selection is an important issue for SVM because operation of the SVM depends on the suitable option to the kernel.Different kernels will accomodate various nonlinear mappings.Four types of kernel are used in SVM^{Tourch}: linear, polynomial, radial basis function (RBF), sigmoid tanh. We are using RBF method.PCA process will start from creating matrix . First consider a vector data set which includes n images in association with training data as shown in following equation.

$$X=(x_1,x_2,.....x_n)^T$$

$\mathbf{x_i}(i = 1,2,\ldots,n)$ can be considered as an image vector with the length p where p represents the number of features.PAC matrix can be constructed using the follpwing equation:

$$C_1=X^TX \text{or } C_2=XX^T$$

In this equation C1 is the matrix of nxn, and C2is the matrix of*pxp*. If n(number of images) is decreased with C1after applying PCA method will result in creating PCA templates which constitutes the entire dataset.In other case when C2is applied and p(nnumber of features)is decreased , PCA component will reduce the feature vectors measurments..

The above method used will be based on the recognition rate of the system and the Average Accuracy(AA).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

AA=number_of_all_right_classified_images/size_of_whole_data_set.
When PCA is used with Blob feature has high image retrival rate but the performance rate of AA is decreased.PCA when it is used with any other features such as texture is not so successful. The reason is because of over-fitting problem which is caused by the unbalance of 57 classes.In these 57 classes some classes have huge amount of training data while the others do not have such large data.So in the testing stage many images are falsely classified.SVM when it is used with different features likeSVM with Blob,SVM with LRPM,SVM with texture,SVM with LRPM and with texture,and finally SVM with Blob with LRPM and with texture.In each of these feaures it is found that SVM with Blob with LRPM and with texture has highest AA of 82% where SVM uses LRPM and Blob as feature sets.

b)SVM and KNN classifier:
We use the same database called ImageCLEFmed2005 database with the same training set and test set images. Two main techniques,the SVM and KNN techniques, will be separately employed to classify these images.In image extraction local,pixel, global, and Surf features are extracted. In addition,to this both texture and shape features is also extracted with the help of global features. Local featurs are extracted by dividing each image into four blocks and extracting features from each block. The gray-level co-occurrence matrix and wavelet techniques are used to extract the texture features.The histogram method is used to extract the shape features.By combining the values of local,pixel,global,Surf values it will result in 1,785 feature vector.The image classification rate is improved by combining the large number of features with the feature vector. This resultant feature vector is too large, so PCA is applied thrice. By using PCA it will result in length reduction of the feature vector from 1,785 to 25, 50, and 100 features, which will help in improving accuracy of the classification system. Furthermore, these features are organized into feature vectors to enhance the accuracy of the classification system. Two classifiers will be used during the classification stage, and these are SVM and KNN.A threshold value of 200 images at most for each category is used through these stages. The reason for applying threshold because this value is related to the unbalanced image numbers within 57 categories.For the SVM classifier, the achieved accuracy results were improved to 95.36% for the 90% training and 10% testing partition; 92.20% for the 80% training, 20% testing database partition without using a threshold; and 91.92% and 90.45% with a threshold used. For the KNN classifier, the results were improved to 92.712% for the 90% training and 10% testing case; and 89.32% for the 80% training, 20% testing case without using the threshold.

III. PROPOSED METHOD

Edge based features and Sparse Representation
We use edge based features and sparse method representation for medical image classification.as the proposedmethod.In this approach we use the database called IRMA(Image Retrival in Medical Application) along with ODL(Online Dictionary Learning).In IRMA database the datasets has to be in sparse models so we use learned dictionary.Content based medical image classification and retrieval techniques uses features like images color, shape and texture. By using these features directly, we may not be able to retrieve similar images easily. In the proposed method, we consider two type of feature extraction methods to represent the content of medical images.
1)Edge based feature extraction is used to extract edge information of the medical images.Medical images of different body parts contains different shapes and different edgeinformation, which is classified easily based on the edge features.Canny edge detection method is used for finding the edges of the images.
2)Patch based feature extraction method is used on edge images. An edge image is divided into equal size of patches.Each patch of the image is partitioned into concentric circular regions of equal area .
The mean and variance of pixel intensity in each circular region become a component of the feature vector.We are using equations shown below, where P is the number of pixels in each region, m is the mean of pixels intensity values and S is the variance of pixels intensity values in each region.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

$$m = \frac{1}{P}\sum_{k=1}^{P}(y_k)$$

$$S = \sum_{k=1}^{P}(y_k - m)(y_k - m)^t,$$

In the feature extraction process,Edge information is extracted from medical images,Then we patch every image by partitioning equally into concentric circular regions where each circular region has same number of pixels .Finally the mean variance of each circular region is calculated.

Proposed system:

In medical image classification a sparse model is applied which constitutes the test data as a sparse linear in conjunction with training data.This training data is taken from dictionary.In this paper, class C = [C1, . . . ,CN] composed of training samples which accomodates directly with regard to images.In the proposed sparsity model, images associated to the same class are supposed to be relative to the low dimensional subspace.Given$N$training classes, the *pth*class has Kptraining images $\{y_i^N\}$ i=1.....,Kp. Let b be an image belonging to the *pth*class, then it is represented as a linear combination of these training samples:

$$b = \mathbf{D}p\Phi p,$$

Dp= m ×Kpa dictionary whose columns are the training samples in the *pth*class and $\Phi p$is a sparse vector.Proposed system has two steps.

In the first step we have to construct the dictionary for each class which has training images using ODL algorithm . We use the following equation to calculate and to construct the dictionary

$$(\hat{\mathbf{D}}_i, \hat{\Phi}_i) = \arg\min_{\mathbf{D}_i,\Phi_i} \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\|\mathbf{C}_i - \mathbf{D}_i\Phi_i\|_2^2 + \lambda\|\Phi_i\|_1$$

$$\text{satisfying } \mathbf{C}_i = \hat{\mathbf{D}}_i\hat{\Phi}_i, \quad i = 1, 2, \ldots, N.$$

satisfying$\mathbf{C}i$= ˆ$\mathbf{D}i$ˆΦi, i= 1, 2, . . . , N.

In the second step we have to classify the images.We choose a parameter for classification called sparse vectorΦwhich corresponds to test image and test dataset B=[b1, . . . , bl]. In order to solve optiization problem we have to use the dictionaries of training samples D =[D1, . . . ,DN], and the sparse representation Φ which satisfies the equation $D\Phi$=B, following equation is used

$$\Phi^j = \arg\min_{\Phi} \tfrac{1}{2}\|\mathbf{b_j} - \mathbf{D}\Phi_j\|_2^2 \quad \text{subject to} \|\Phi_j\|_1$$

$$and\ \hat{i} = \arg\min_i \|\mathbf{b_j} - \mathbf{D}\delta_i(\Phi^j)\|_2^2 \quad j = 1, \cdots, t,$$

whereδiis a characteristic function that selects the coefficients.Thenbjis assigned to Ci associated with the ithdictionary.Then, test data is assigned to the class associated with this sparsest dictionary.

The proposed method gives best classification results of 98.5% as compared to other image classification techniques such as Linear Discriminative Analysis (LDA), Kernel SVM, Neural Network (NN), KNN (K-Nearest Neighbor) and Bayes Classifier (BC).

IV. CONCLUSION

In the first classifier method various features were extracted and with the SVM and PCA it was found that SVM had better classification rate.Further SVM when used with KNN the classification rate was improved.In both of these classification methods the same database was used.The edge-based features along with on-line dictionary learning and

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

sparse representation based classification gives the best possible classification performance till date.It had a higher classification rate of 98.5%.The results suggest that the proposed method is better than other well known classification algorithm.

REFERENCES

1)A. R. S. R. Beaula, P. Marikkannu, A. Sungheetha and C. Sahana, "Comparative study of distinctive image classification techniques," *2016* 10th InternationalConference on Intelligent Systems and Control (ISCO), Coimbatore, 2016, pp. 1-8.doi: 10.1109/ISCO.2016.7727002
2) Bo QIU, Chang Sheng XU, Qi TAN "An Automatic Classification System Applied in Medical Images,"
2006 International Conference on Multimedia and Expo Pages: 1045 - 1048, DOI:10.1109/ICME.2006.262713
3)Abdulrazzaq, M.M., Fadhil, M.A., Mohd, S.A." Medical Image Annotation and Retrieval by Using Classification Techniques"Advanced Computer Science Applications and Technologies (ACSAT), 2014 3rd International Conference on, pp. 32-36, 2014
4)Aboul Ella Hassanien, Ajith Abraham, James F. Peters, Gerald Schaefer, Christopher Henry, "Rough Sets and Near Sets in Medical Imaging: A Review", Information Technology in Biomedicine IEEE Transactions on, vol. 13, pp. 955-968, 2009, ISSN 1089-7771.
5)M.M. Abdulrazzaq, ShahrulAzmanMohd, Muayad A. Fadhil, "Medical Image Annotation and Retrieval by Using Classification Techniques", Advanced Computer Science Applications and Technologies (ACSAT) 2014 3rd International Conference on, pp. 32-36, 2014.
6).M. Srinivas; C. Krishna Mohan "Classification of medical images using edge-based features and sparse representation "2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) Pages: 912 - 916, DOI: 10.1109/ICASSP.2016.7471808