Clustering of Medical X-ray Images by Merging Outputs of Different Classification Techniques

Ibrahim Zeiadan[1], Amr Zamel[2] ,and Ahmed Al Zohairy[3]

[1] Professor, Computers and Systems Engineering Department, Zagazig University, Egypt.
ieziedan@gmail.com

[2] Teaching Assistance, Computers and Systems Engineering Department, Zagazig University, Egypt.
eng.amrzml@gmail.com

[3]Assistant Professor, Genetics Department, Faculty of Agriculture, Zagazig University, Egypt
alzohairy@yahoo.com

Abstract. Clustering x-ray images is a complex task, due to the great variations within each class including orientation, alignment and deformation. In this paper, an automatic medical x-ray image clustering is developed by merging the outputs from five different neural networks classifiers. Each classifier employs a set of features derived through different feature-extraction techniques. Such techniques are based on (i) pixel-value, (ii) local binary patterns, (iii) global means of rows and columns, (iv) local means of rows and columns, and (v) local histogram. A test accuracy of 86.2 % was achieved from merged output of the five NN classifiers using the ImageCLEF 2015 database. A somewhat higher accuracy of 87.2% was obtained when merging outputs of only three classifiers.

Keywords: classification, image processing, x-ray, neural network

1 Introduction

The problem of x-ray image classification is gaining a growing interest of many researchers now-a-days. In recent years medical images have increased and so demands for indexing, storing and analyzing such images have also increased. The problem of automatic medical image clustering involves mainly three steps, preprocessing, feature extraction and a classification technique. The key point in the clustering task is classifier features. Features are mainly generated from two levels (i) low-level representation and (ii) patch based image representation [ref]. Low-level image representation is such as edge histogram, local binary patterns, SIFT histogram, Gray Level Co-occurrence Matrix (GLCM), and Canny edge operator. The local binary patterns (LBP) have been used by several researchers in various domains [1,2,3]. Combination of block based local binary patterns with edge histogram was used as a medical image representation for the task of automatic medical image annotation in ImageCLEF 2007 [4].

Local patch-based image representation has been developed for use in feature extraction by M. Zare et al. [5]. A local feature is obtained by dividing the image into sub images (regions or partitions) or selecting interesting points from the image. Bag of words (BoW) is one of the intersection point techniques used to represent images using histograms of quantized appearances of local patches [6].

Merging neural network-classifiers' outputs obtained by using different features extracted through different extraction techniques is proposed in this paper. The paper is organized as follows. In section 2, preprocessing techniques are used to improve x-ray images and remove the noise. Section 3, presents the different feature-extraction techniques generated from the image. In section 4, a clustering frame work using a merging technique of different classifiers is described. Section 5, test results and test accuracy are presented. Section 6, presents conclusions and future work.

2 Preprocessing

X-ray images are usually given with different resolutions and usually contain some boundary and image noise. Image enhancement methods are needed to adjust digital images so that they are more suitable for display or further analysis. Such enhancement includes resizing and removing the boundary or image noise as well as enhancement of the contrast or image intensity. However the image is resized to 512 x 512 pixels after converting it into gray-level one.

Histogram equalization, as one of image enhancement techniques, is applied to improve the quality of the image as well as its contrast [7]. Such contrast adjustment provides better gray intensity distribution.

Generally noise in an image may follow three possibilities a Gaussian distribution, Poisson distribution or a combination of both. To remove such noise, two types of filters, namely linear and median filter may be used [7] .Median filtering [8] is similar to an averaging filter, in which each output pixel is set to an average of the pixel values in the neighborhood of the corresponding input pixel. However, with median filtering, the value of an output pixel is determined by the median of the neighborhood pixels, rather than the mean. The median is less sensitive than the mean to extreme values (called outliers). Median filtering is therefore better at removing these outliers without reducing the sharpness of the image.

3 Feature Extraction

This section describes the features that may be obtained from x-ray images. Features from different feature-extraction techniques are used to train different classifiers. Such techniques are namely pixel – value of the gray image, Local Binary Patterns (LBP), local and global means of pixel value of the rows and columns, and local histogram.

3.1 Pixel Level (Value)

Pixel value is the simplest form of image representation techniques. It carries only intensity information. The intensity of a gray image pixel is expressed within a given range between a minimum and a maximum, inclusive. This range is represented in an abstract way as a range from 0 (total absence, black) and 1 (total presence, white), with any fractional values in between [9].

The size of the image is scaled down to (32*32) pixels to reduce the number of features. The feature vector of pixel intensity value is then obtained as a single column of size 1024 elements for each image.

3.2 Local Binary Patterns

Local-binary patterns extract the texture in the gray image [10]. Normally, LBP labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number that is converted to decimal value. Firstly, the image is divided into non-overlapping square image blocks with the same size of neighbor set. Secondly, the square image is converted into binary by finding the center pixel value and using it as a threshold. If the value of a neighborhood is greater than the threshold it is represented by a one otherwise it is represented by a zero. Thirdly, the decimal value for each square block in the image is calculated. Finally, the binary pattern distribution for each square block included in the image can be represented as a histogram having 59 bins. To reduce the effect of position variation a local binary pattern at different sub levels (L0, L1, L2) from the image is obtained [ref], The levels of LBP is obtained by dividing the image into 2x2 or 4x4 non-overlapping regions for levels L1 and L2. Level L0 is obtained for the whole image. The final LBP features are generated by combining local histogram features. Since there are 21 sub regions (whole image, 2x2 sub image, 4x4 sub images) a total of 59×21= 1239 histogram bins are generated for the feature vector.

3.3 Global Means of Rows and Columns

The image is represented as a matrix of pixel intensity values, so we can calculate some statistical properties such as the mean of rows and columns. The mean value for the pixel intensity of each row and column is calculated and then combined to form a feature vector of the global image. So, the final length of the global feature vector is the number of rows plus the number of columns. Since the image size is 512 x 512 a total of 1024 value (feature) is generated for the global means-feature vector.

3.4 Local Means of Rows and Columns

Local features are more robust to occlusion and clutter. An average per region technique was used in face recognition task and produced high accuracy rate [11]. Therefore, the image was divided into 4×4 non-overlapping sub-images and then for each sub-image the mean value of each row and each column was obtained. A final mean feature vector is generated by combining the mean of each row and column of each sub image. Since the image size is 512x512 pixels, each sub image size is 128 x128. The mean feature vector length of each sub image is 128+128 = 256. The final local-mean feature vector of the whole image is 16x256 = 4096.

3.5 Local Histogram

As is well-known a histogram measures the distribution of intensity level in gray images. Image histogram features were used for face recognition based on minimum distance between a test image and a gallery of database of images **[**12**]**. Local features are extracted from small sub-images that are generated by partitioning the original image into a number of segments. In this paper, the gray level of the image is divided into 30 pins. Local histograms are obtained for the x-ray 4x4 non-overlapping sub-images. Final histogram features vector was generated by combining features of each sub-image. Since there are 16 sub-images, a total of 30×16=480 feature vector elements are generated.

4 Image Clustering Framework

Clustering x-ray images is a complex task, due to the great variations within each class including orientation, alignment and deformation. In this paper, the image clustering framework is consisting of three different phases namely feature extraction, training neural network (NN) classifiers, and a merging technique as shown in Fig.1.

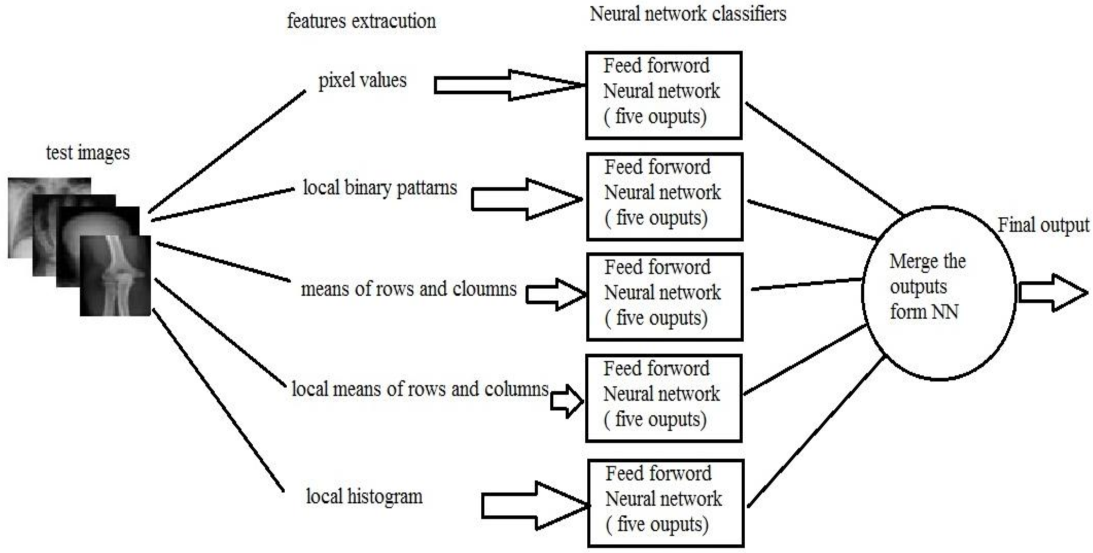

Fig. 1. Image clustering framework

In the first phase, features are extracted from the preprocessed images as mentioned before (section 3). Pixel level and local binary patterns, global means of rows and columns, local means of rows and column, and local histogram are the basic means of feature extraction in this work, and a feature vector is derived for each.

In the second phase, a NN is developed for each feature vector. Each neural network consists of two hidden layers and five outputs for the five clusters. The number of inputs for each network equals the number of elements of its corresponding feature vector as derived in section 3.

In the third phase, the outputs from five NN's are merged to get the final output. The merging processing is done in four steps. Step 1, Normalize the output of each classifier by dividing each output by the total sum from the classifier outputs. Step 2, a weighted sum of the outputs from different classifiers constitutes the merging operation. The weight of each classifier is taken equal to the accuracy obtained from it. The accuracy may be defined as follows.

$$accuracy = \frac{TC}{TC + FC} \tag{1}$$

Where, TC is the number of true classifications. FC is the number of false classifications. TC + FC is the total number of the test images. It is should be noted that a true classification means a test image belongs to only its cluster. Step 3, the final outputs are normalized by dividing each output by the total sum. Step 4, the final clustering of the test image is obtained by selecting the group that has the maximum value in the merged outputs.

5 Experimental Results

In this section, a set of five classifiers is developed, one for each feature vector using a neural network. The database used in this research work is the ImageCLEF 2015 [13,14,15]. This database contains 500 x-ray images belonging to five main groups (Body, Head-Neck, Lower Limb, Upper Limb, and True Negative). Each group has 100 images. 25 % of the 500 images were randomly selected and used in the testing phase. The remaining 75% are used in the training phase of NN's. After training, the accuracy of each classifier is calculated first and then used as a weight in the merging process as mention before (section 4). Each NN consists of two hidden layers of size 200 and 100 neurons. These were selected after trying different numbers of hidden layer neurons.

Testing was carried out using 25 images from each group first and then it was performed for the 125 testing images. Results are shown in table 1 where the accuracy of each NN classifiers is listed. The pixel-value classifier shows a total accuracy of 79.2% with a highest accuracy of 92% for the true negative cluster and a lowest one of 64% for the upper limb cluster. The LBP classifier gives 80% as a total accuracy and 84% for the Body and Head-Neck clusters and a lower accuracy rate of 76% for the upper-limb and lower-limb. The global means of rows and columns classifier shows a higher total accuracy of 83.2% with a highest accuracy of 96% for the true negative cluster and a lowest one of 60% for the lower-limb cluster. The local means classifier shows a higher accuracy rate for the true-negative and body clusters of 88% and a lower accuracy rate for the lower limb of 52% with a total accuracy of 70.4%. The local histogram classifier gives a total accuracy of 71.2% and shows a higher accuracy rate of 88% for the body cluster and a lower one of 48% for the lower-limb cluster.

Merged outputs from different classifiers are then used to improve the overall accuracy. The weight of each classifier is equal to its accuracy. Two experiments were performed in this respect.

First, merging all classifier outputs gave a total accuracy of 86.4%. A higher accuracy of 92% was obtained in three cluster and a lower accuracy of 72% occurred in the lower-limb cluster. This merged classifier was submitted to the "Medical Clustering task of ImageCLEF 2015" [14]. In this task result was the tenth among the 29 participating groups with a total accuracy of 78% and hamming similarity of 86.8% for 250 test images.

Second, merging the first three higher accuracy classifiers (pixel level, LBP and Global mean of rows and columns) raised the total accuracy to 87.2% .obviously this is due to the increase in the accuracy of the true negative cluster (96%).

Table 1. Accuracy of each NN classifiers

	Body	Head-Neck	Upper-Limb	Lower-Limb	True-Negative	accuracy
Pixel Value classifier	80%	80%	80%	64%	92%	79.2%
LBP classifier	84%	84%	76%	76%	80%	80%
Global mean classifier	80%	92%	88%	60%	96%	83.2%
Local mean classifier	88%	68%	56%	52%	88%	70.4%
Local histogram classifier	88%	84%	72%	48%	64%	71.2%
Merged the five classifiers outputs	92%	92%	84%	72%	92%	86.4%
Merged the upper three classifiers	92%	92%	84%	72%	96%	87.2%

6 Conclusion and Future work

In this paper an automatic medical x-ray images clustering system was developed by merging the outputs from different neural-network classifiers with different feature extraction techniques. These techniques are based on pixel-value, local binary pattern, global means of rows and columns, local-partition means of rows and columns, and local histogram features. Merged outputs from different classifiers show improvement in overall accuracy than individual classifiers. The best individual classifier is the global means of rows and columns classifier with 83.2% accuracy rate. The merged outputs from the five classifiers gave an accuracy of 86.2 %. The merged outputs from the top three classifiers produced an accuracy of 87.2%.

The proposed approach can be easily extended with new feature extraction methods, and can thus be applied to other domains. The proposed approach for merged classifier outputs can be easily applied to arbitrary domains with different feature extraction techniques with different sizes. Also, matching techniques may be employed for classification other than using NN.

Reference

1. Dimitrovski, I., Kocev, D., Loskovska, S., Dzeroski, S.: ImageCLEF 2009 Medical Image Annotation Task: PCTs forHierarchical Multi-Label Classification. In : CLEF 2009 (2010)
2. Kim, S., Lee, J.-H., Ko, B., Nam, J.: X-ray image classification using Random Forests with Local Binary Patterns. In : International Conference on Machine Learning and Cybernetics (2010)
3. Mustaffa, M., Ahmad, F., Rahmat, R., Mahmod, R.: Content-Based Image Retrieval Based On Color-Spatial Features. Malaysian Journal of Computer Science (2008)
4. Tian, G., Fu, H., Feng, D.: Automatic medical image categorization and annotation using LBP and MPEG-7 edge. In : International Conference on Information Technology and Applications in Biomedicine (2008)
5. Zare, M., Seng, W., Mueen, A.: AUTOMATIC CLASSIFICATION OF MEDICAL X-RAY IMAGES. Malaysian Journal of Computer Science 26 (2013)
6. Feng, J., Jiao, L., Zhang, X., Yang, D.: Bag-of-Visual-Words Based on Clonal Selection Algorithm for SAR Image Classification. IEEE Geoscience and Remote Sensing Letters (2011)
7. Gonzalez, R., Woods, R.: Digital Image Processing 2nd edn. Prentice Hall (2002)
8. Lim, J.: Two-Dimensional Signal and Image Processing. Englewood Cliffs, NJ, Prentice Hall (1990)

9. Gonzalez, R., Woods, R., Eddins, S.: Digital image processing Using MATLAB. Pearson Perntice Hall (2004)
10. Ojala, T., Pietikäinen, M., Harwood, D.: A Comparative Study of Texture Measures with Classification Based on Feature Distributions. Pattern Recognition 29 (1994)
11. Zieadan, I., Nasef, B.: A New Trend for face Recognition Features. international jouranl of advanced image processing Techniques 1 (2014)
12. Singh, S., Sharma, M., Suresh, N.: Robust & accurate face recognation using histograms. Int.J.of Mangment It, and Engineering (IJMIA) 2 (2012)
13. Villegas, M., Muller, H., Gilbert, A., Piras, L.,Wang, J., Mikolajczyk, K., de Herrera, A.G.S., Bromuri, S.,Amin, M.A., Mohammed, M.K., Acar, B., Uskudarli, S., Marvasti, N.B., Aldana, J.F., del Mar Roldan Garcia,M.: General Overview of ImageCLEF at CLEF2015 Labs. Lecture Notes in Computer Science, SpringerInternational Publishing (2015)
14. Amin, M.A., Mohammed, M.K.: Overview of the ImageCLEF 2015 medical clustering task. In: CLEF2015 Working Notes. CEUR Workshop Proceedings, CEURWS.org, Toulouse, France (September 8-11 2015)
15. Faruque, M. S. S., Banik, S., Mohammed, M. K., Hasan, M., Amin, M. A.: Teaching & Learning System for Diagnostic Imaging; Phase I: X-Ray Image Analysis & Retrieval. In Proceedings of the 6th International Conference on Computer Supported Education (2015). https://www.dropbox.com/s/7334oyhgyezxz7q/Teaching%20%26%20Learning%20System%20for%20Diagnostic%20Imaging%20Camera%20Ready%202.pdf?dl=0