National Library of Medicine (NLM) at ImageCLEF2015: Medical Clustering Task

Szilárd Vajda, Sameer Antani, and George Thoma

Lister Hill National Center for Biomedical Communications,
National Library of Medicine, National Institutes of Health
8600 Rockville Pike, Bethesda, MD 20894, USA
{szilard.vajda,sameer.antani,george.thoma}@nih.gov
http://lhncbc.nlm.nih.gov/

Abstract. Besides recognizing medical image modalities, such as X-rays, MRIs, histology images, fluorescence microscopy images, endoscopy images, photos, illustrations, etc., the detection of visual content is equally important. Once the main modality class is detected, a modality such as X-ray can be broken to different sub-classes representing different body parts such as arms, legs, neck, torso, etc. Such a classification can further help the current image-based search engines to return appropriate results based on visual content similarity. For our participation in the ImageCLEF2015 Medical image clustering task, we implemented a classification scheme based on a neural network using two different feature collections – which proved their value in object recognition and chest X-ray analysis.

Keywords: body parts x-ray, classification, shape features, texture features, modality detection, ImageCLEF

1 Introduction

Medical image retrieval in the context of large collections is a challenging and demanding task [16,17]. Increased research interest has resulted in years long systematic evaluation efforts[1]. One outcome of the evaluation was that knowing the modality of an image, i.e., whether it is an X-ray, CT, MRI or a photograph, radically improves the performance of image retrieval [11,19].

2 Methods

This section describes the motivation for the particular feature sets in use, provides a brief description of the different features, and finally has a short section describing the classifier.

[1] http://www.imageclef.org

2.1 Description of features

To characterize the different body parts in the X-ray images, we considered two different feature sets. Feature Set A is inspired from object detection [7,12], and was used with success in a previous work [9] to detect pulmonary abnormalities in frontal chest X-ray images. Feature Set B has been utilized with success in [14] for a medical CBIR system. These features cover a large number of properties, such as color distribution, edginess, texture, curvatures, pixel densities, shape, and other measures necessary to describe images such as in Figure 1. We note that the body parts in the X-ray images appear in in different size, rotation, shape, etc. Therefore, features invariant to size and rotation or shape are most appropriate.

Set A: Is a versatile and compact feature set combining shape, edge and texture descriptors. The final feature representation is built by concatenating the different descriptors (histograms) extracted from the segmented lung regions. In particular, in Set A, the following shape and texture descriptors were considered: Intensity Histogram (IH), Gradient Magnitude Histogram (GM), Shape Descriptor Histogram (SD), Curvature Descriptor Histogram (CD), Histogram of Oriented Gradient (HOG) [5], Local Binary Pattern (LBP) [13]. A modified multiscale approach proposed by Frangi et al. [6] is considered to compute the eigenvalues of Hessian matrix needed for the shape and curvature descriptors. The Hessian describes the second-order surface curvature properties of the local image intensity surface. The normalization makes these descriptors intensity invariant. In [9] we determined that quantizing these features into 32 bins provides good discrimination performance. The size of the feature descriptor is 192.

Set B: Is a diversified, low-level feature collection involving intensity, edge, texture, color and shape moment features. The feature representation is built by concatenating the different descriptors (histograms) extracted from the segmented lung regions. In particular, the following descriptors were considered: Color Layout Descriptor (CLD), Edge Histogram Descriptor (EHD) from MPEG-7 standard [10], Color and Edge Direction Descriptor (CEDD) [3], Fuzzy Color and Texture Histogram (FCTH) [4] Tamura texture descriptor, Gabor texture feature [8], and other texture features such as primitive length, edge frequency, and autocorrelation [15]. This feature set comprises 595 dimensions.

Set C: Is a union of set A and set B. Even though some of the features are similar or similar characteristics, this extended feature collection can be a powerful descriptor for such particular type of X-ray images of different body parts, as discussed in [9]. Practically the features were stitched together to form a larger feature descriptor. The dimension of this feature descriptor is 787.

2.2 Feature classification

For classification a neural network-based classifier was used. Neural networks in particular are known for their capability of estimating complex decision surfaces [2] and handling multi-class problems. Due to the large numbers of features to be handled (up to 787 dimensions), and the lack of information about the possible correlations among the different feature components, a fully connected multi-layer perceptron network was utilized. The number of neurons in the input layer was selected based on the dimensionality of the input feature vector. The number of output neurons was also set based on the possible outcomes: head-neck, upper-limb, body, and lower-limb, while the number of neurons in the hidden layer was estimated based on several trial runs. Finally, for the experiments 15 neurons were considered as being optimal in the hidden layer. For training error-backpropagation strategy was considered, while for learning rate = 0.004, and momentum = 0.3 were used. The different parameters were established based on several trial runs. For the number of hidden neurons in the hidden layer, we considered the criteria to have as less possible neuron to keep the complexity low. Therefore, the recognition time became faster.

3 Experiments

This section gives a brief description of the data followed by the description of the evaluation protocols, and the different results obtained by the neural classifier utilizing the different feature collections.

3.1 Data description

The data provided by the ImageCLEF2015 [18] Medical image clustering task organizers contains 500 X-ray images of variable sizes [1]. The content of the image is equally distributed among X-rays containing head-neck, upper-limb, body, and lower-limb images. An equal number of 100 images of true negatives - containing completely different images, were also provided to help the researchers providing negative examples to their systems. Some images from the data collection are shown in Figure 1. For more details about the data, please refer to [1].

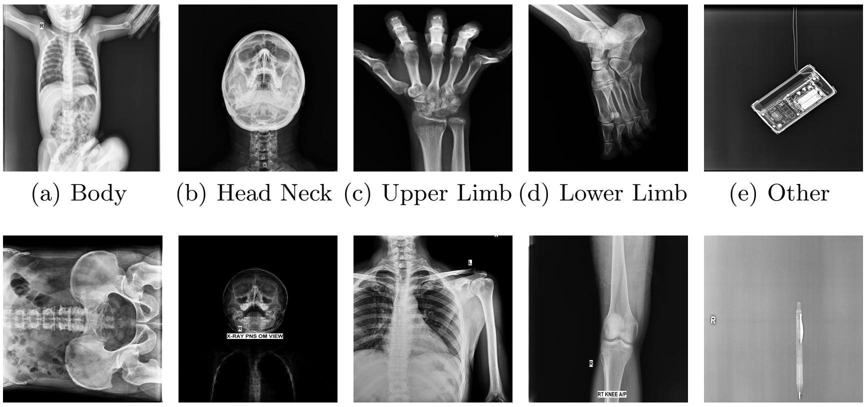
3.2 Evaluation protocols

The accuracy (ACC) was measured to properly evaluate the performance of the method. Each of our experiments follows a 10x cross-validation protocol, and the reported results are the average scores of the different folds. However, the competition organizers used three different measures, namely the exact match, any match and the Hamming distance [1].

3.3 Results

Using the previously mentioned feature set A, B and the C, three different experiments were conducted. The first results are presented in Table 1.

One can observe the superiority of the Set C, which contains both features from set A and Set B, respectively. It is quite important to mention the fact, that color features such as CEDD, FCTH also contribute to the higher accuracy of the system, whilst, features such as intensity histogram, LBP and HOG, which



(a) Body (b) Head Neck (c) Upper Limb (d) Lower Limb (e) Other

(f) Body (g) Head Neck (h) Upper Limb (i) Lower Limb (j) Other

Fig. 1. Image samples from the training collection.

Feature	Accuracy (%)
Set A	75.20
Set B	80.80
Set C	81.60

Table 1. Results for the different features on the training set using 10x cross-validation.

describe the edginess and texture of the image, perform the worse; even though they are rotation invariant, except HOG. The results shown in Table 1. are reported using the 10x cross-validation protocol on the training samples (labels available) provided originally by the organizers.

The results reported in Table 2. were generated based on the test set (no labels available) provided by the competition organizers. This image collection contains 250 samples, similar to the training material, equally distributed among the 5 classes. To train the neural network all data (500 samples) available in the training set were considered. Despite the usage of different metrics for evaluating the performance of this experiment, the rank among the features is preserved, as expected.

Feature	Exact match	Any match	Hamming similarity
Set A	0.543	0.656	0.810
Set B	0.593	0.716	0.842
Set C	0.613	0.740	0.849

Table 2. Results for the different features on the test sets using other metrics [1].

4 Conclusion

In this paper we introduced three different feature sets (A, B and C) applied with sucess in pulmonary disease detection in chest X-ray images, and image retrieval and modality detection. It is interesting to note that features such as intensity histogram or histogram of oriented gradient or localy binary patterns among others belonging to set A were less effective than generic features such as Fuzzy Color and Texture histogram, Tamura texture feature, Gabor texture features, Edge histogram descriptor and Color Layout descriptor coming from feature set B. The combination of these features into a larger set C allowed us to classify quite effectively the different body parts from the X-ray images.

Among the 7 participants in the contest (see detailed description in [1]) we ended up on 4^{th} place, while among the 29 runs submitted by the different research groups, our runs finished at 11^{th},17^{th}, and 25^{th}, respectively.

To further improve the classification scores, we envision an exhaustive feature selection mechanism applied to the set C, to eliminate those features which are rotation variant such as HOG, and other feature components which do not contribute much to the final scores.

Acknowledgment

This research is supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine, and Lister Hill National Center for Biomedical Communications (LHNCBC).

References

1. Amin, M.A., Mohammed, M.K.: Overview of the ImageCLEF 2015 medical clustering task. In: CLEF2015 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org, Toulouse, France (September 8-11 2015)
2. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA (1995)
3. Chatzichristofis, S.A., Boutalis, Y.S.: Cedd: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In: Proceedings of the 6th International Conference on Computer Vision Systems. pp. 312–322. ICVS'08, Springer-Verlag, Berlin, Heidelberg (2008)
4. Chatzichristofis, S.A., Boutalis, Y.S.: Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In: Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services. pp. 191–196. WIAMIS '08, IEEE Computer Society, Washington, DC, USA (2008)
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA. pp. 886–893 (2005)
6. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Muliscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI'98, First International Conference, Cambridge, MA, USA, October 11-13, 1998. pp. 130–137 (1998)

7. Gonzalez, R.C., Woods, R.E.: Digital Image Processing (3rd Edition). Prentice-Hall, Inc., Upper Saddle River, NJ, USA (2006)
8. Howarth, P., Yavlinsky, A., Heesch, D., Ruger, S.: Medical image retrieval using texture, locality and colour. In: Peters, C., Clough, P., Gonzalo, J., Jones, G., Kluck, M., Magnini, B. (eds.) Multilingual Information Access for Text, Speech and Images. Lecture Notes in Computer Science, vol. 3491, pp. 740–749. Springer Berlin Heidelberg (2005)
9. Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F.M., Xue, Z., Palaniappan, K., Singh, R.K., Antani, S., Thoma, G.R., Wang, Y., Lu, P., McDonald, C.J.: Automatic tuberculosis screening using chest radiographs. IEEE Trans. Med. Imaging 33(2), 233–245 (2014)
10. Lux, M.: Caliph & emir: Mpeg-7 photo annotation and retrieval. In: Proceedings of the 17th ACM International Conference on Multimedia. pp. 925–926. MM '09, ACM, New York, NY, USA (2009)
11. Müller, H., Kalpathy-Cramer, J., Demner-Fushman, D., Antani, S.: Creating a classification of image types in the medical literature for visual categorization. In: SPIE medical imaging (2012)
12. Murphy, K.P., Torralba, A., Eaton, D., Freeman, W.T.: Object detection and localization using local and global features. In: Toward Category-Level Object Recognition. pp. 382–400 (2006)
13. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recognition 29(1), 51–59 (1996)
14. Rahman, M.M., You, D., Simpson, M.S., Antani, S., Demner-Fushman, D., Thoma, G.R.: Interactive cross and multimodal biomedical image retrieval based on automatic region-of-interest (ROI) identification and classification. IJMIR 3(3), 131–146 (2014)
15. Singh, S., Sharma, M.: Texture analysis experiments with meastex and vistex benchmarks. In: Singh, S., Murshed, N., Kropatsch, W. (eds.) Advances in Pattern Recognition ICAPR 2001. Lecture Notes in Computer Science, vol. 2013, pp. 419–426. Springer Berlin Heidelberg (2001)
16. Vajda, S., You, D., Antani, S., Thoma, G.R.: Label the many with a few: Semi-automatic medical image modality discovery in a large image collection. In: 2014 IEEE Symposium on Computational Intelligence in Healthcare and e-health, CICARE 2014, Orlando, FL, USA, December 9-12, 2014. pp. 167–173 (2014)
17. Vajda, S., You, D., Antani, S., Thoma, G.: Large image modality labeling initiative using semi-supervised and optimized clustering. International Journal of Multimedia Information Retrieval 4(2), 143–151 (2015)
18. Villegas, M., Müller, 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 Roldán García, M.: General Overview of ImageCLEF at CLEF2015 Labs. Lecture Notes in Computer Science, Springer International Publishing (2015)
19. You, D., Rahman, M.M., Antani, S., Demner-Fushman, D., Thoma, G.R.: Text- and content-based biomedical image modality classification. In: Proc. SPIE Medical Imaging. pp. 86740L–86740L–8 (2013)