

Content-based Medical Image Retrieval for Liver CT Annotation

¹Imane Nedjar, ²Saïd Mahmoudi, ³Mohammed Amine Chikh

^{1,3}Biomedical Engineering Laboratory, Tlemcen University, Tlemcen, Algeria

²Computer Science Department, Faculty of Engineering, University of Mons, Mons, Belgium

imane.nedjar@mail.univ-tlemcen.dz, said.mahmoudi@umons.ac.be, ma_chikh@mail.univ-tlemcen.dz

ABSTRACT

The increase number of medical image stored and saved every day presents a unique opportunity for content-based medical image retrieval (CBMIR) systems. In this paper, we propose content-based medical image retrieval for annotating liver CT scans images in order to generate a structured report. For that, we have used the Bidimensional Empirical Mode Decomposition (BEMD), and then we have applied Gabor wavelet transform to extract the mean and the standard deviation as features descriptors. Finally, a proposed similarity distance was employed to retrieve the most similar training images to the image query, and a majority voting scheme was used to select the answers for an unannotated image. We have used the IMAGECLEF 2015 annotation dataset and the obtained score was 88.9%.

Keywords: Medical image, Liver annotation, Image Retrieval, BEMD, Gabor wavelet.

1 Introduction

The huge amount of digital images produced in hospitals and health care centers needs the development of automatic mechanisms for their storage and retrieval from databases. The structured reports are highly valuable in medical contexts due to the processing opportunities that they provide. However, structured reports are time consuming, since they need a lot of time to be created. Furthermore, their creation requires high domain expertise. Consequently, such structured medical reports are often not found or are incomplete in practice [1].

The goal of this work is to improve computer-aided automatic annotation of liver CT volumes by filling in a structured radiology report. Therefore, the major part of this work is devoted to describe the semantic features of the liver, its vascularity, and the types of lesions in the liver.

2 Related works

Several schemes and algorithms have been presented in the literature for annotation of liver CT scans.

In 2015 the authors of [2], proposed two methods for annotating the liver image. The first one uses a classification approach composed of two main phases. The first step consists of a pre-processing allowing the extraction of texture and shape based features vector from both liver and lesion. Afterward, a classification process is achieved by using random forest classifier. The second method uses a specific signature of the liver. This method is applied to a set of slices from 3D liver CT scans which are normalized into rectangular blocks. After applying the 1D Log-Gabor filters transformation dominant phase data of each block is extracted and quantized to four levels in order to encode the unique pattern of the liver into

a bit-wise template. The Hamming distance is employed for retrieval. The proposed methods were evaluated on ImageCLEF2015 liver CT annotation dataset, all these methods achieved scores (>90), and a completeness level of 99%.The best score was 91% given by the second method.

In 2014, Kumar et al [3] proposed two strategies for annotating the liver images. The first method uses multi-class classification scheme where each label has a classifier trained to separate it from other labels. The second method uses the similarity scores from an image retrieval algorithm as weights for a majority voting scheme, thereby, reducing the inherent bias towards labels that have a high number of samples. The proposed methods were evaluated by using ImageCLEF2014 liver CT annotation dataset, they achieved the highest score of 94.7% out of all the submission done to imageCLEF2014.

On the other side, the authors in [4] tried four different classifiers in the learning phase: linear discriminant analysis (LDA), logistic regression (LR), K-nearest neighbors (KNN), and finally SVM, to predict labels. An exhaustive search of every combination of image features is done by using leave-one-out cross validation method on training data for every label and classifier. Python scikit-Learn Machine learning toolbox was used for implementing each classifier with default parameters. As result, for most of the labels they got almost the same performance by using any classifier and any combination of image features. The proposed method was evaluated with ImageCLEF2014 liver CT annotation dataset, and they obtained a score of 93%.

In [5] the authors proposed an approach based of a probabilistic interpretation of tensor factorization models, i.e. Generalized Coupled Tensor Factorization (GCTF). This method can simultaneously fit a large class of matrix/tensor models to higher-order matrices/tensors with common latent factors using different loss functions. The authors considered the dataset as heterogeneous data, and the GCTF approach was applied to predict labels. They considered KLdivergence and Euclidean-distance based cost functions as well as the coupled matrix factorization models by using the GCTF framework. The proposed method was evaluated with ImageCLEF2014 liver CT annotation dataset, and their highest score was 67.7%.

3 Material

In our experimentation, we have used the ImageCLEF2015 liver CT annotation dataset. The training dataset includes 50 cases, each consisting of:

- A cropped CT image of the liver, a 3D matrix. The volumes had varied resolutions (x: 190-308 pixels, y: 213-387 pixels, slices: 41-588) and spacing (x, y: 0.674-1.007mm, slice: 0.399-2.5mm),
- A liver mask that specifies the part corresponding to the liver, A bounding box (ROI) corresponding to the selected region of the lesion within the liver,
- An RDF file generated using Liver Case Ontology (LiCO) representing manually entered imaging observations by a radiologist. In total, there are 73 User Expressed Features (Use). These features clinically characterize the liver, hepatic vascularity, and liver's lesions, an example of Use is shown in the following table.

Table 1. Exemple of UsE features [6]

Group	Concept	Properties	Possible values (assigned indices)
Liver	Liver	Liver Placement	downward displacement(0), normal placement(1), leftward displacement(2), upward displacement(3), other(4)

The test dataset contained 10 CT volumes, with varied resolutions and pixel spacing, cropped to the region around the liver. The test data also included a mask of the liver pixels and a bounding box (ROI) (see Fig. 1). But, there is no RDF file and imaging observations are missing.

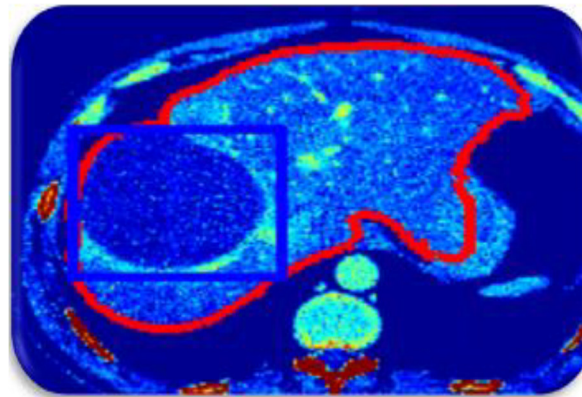


Figure 1. Image test of the liver, the contour in red and ROI in blue

4 The proposed method

Our goal in this work was to investigate a new annotation process using content based image retrieval. The proposed approach is based on a spectral images textural analysis that uses a combination of Bidimensional Empirical Mode Decomposition (BEMD) and Gabor wavelet transform. The features vector includes the mean and the standard deviation of each filtered image.

Thereafter, based on a similarity distance we selected the five most similar images, and the label that has the majority voting will be assigned to the User Expressed Features (UsE).

We used the proposed similarity score calculation between the unannotated image (U) and a training image (T) as the distance between their respective features vectors is given as bellow in equation (1):

$$Sim(U, T) = \sum_{i=0}^d 1 - \frac{|u_i - t_i|}{v_i} \quad (1)$$

Where v_i is the i -th maximum feature in dataset, u_i is the i -th feature in the feature vector of U, t_i is the i -th feature in the feature vector of T, and d presents the dimensionality of the feature set.

In the following section, we present a general overview of the different methods used in this paper.

4.1 Bidimensional Empirical Mode Decomposition (BEMD)

The Bidimensional Empirical Mode Decomposition (BEMD) approach is a highly adaptive decomposition [7]. It is mainly based on the characterization of image by Intrinsic Mode Function (IMF) decomposition, where the image can be decomposed into a redundant set of composite images called IMF and a residue. Adding all the IMFs with the residue allow reconstructing the original image without distortion or loss of information [8]. This method, derived from image data and which is fully unsupervised, permits to analyze nonlinear and non-stationary data as texture images.

4.2 Gabor Wavelet Transform

The Gabor wavelet representation allows good recognition without correspondence, because it captures the local structure corresponding to spatial frequency (scale), spatial localization, and orientation selectivity [9]; it is defined as follows in equation (2):

$$\psi_{u,v}(z) = \frac{\|K_{u,v}\|^2}{\sigma^2} e^{\left(\frac{\|K_{u,v}\|^2 \|z\|^2}{2\sigma^2} \right)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right] \quad (2)$$

Where u and v define the orientation and the scale of the Gabor kernels, and $\|\cdot\|$ denotes the norm operator, and the wave vector K_{uv} is defined as follows:

$$K_{u,v} = K_v e^{i\theta_u} \quad (3)$$

Where $K_v = K_{\max}/f_v$ and $\theta_u = \pi u / 8$, $Z=(x,y)$ and K_{\max} is the maximum frequency, f is the spacing factor between kernels in the frequency domain. We have used $\sigma = 2\pi$, $K_{\max} = \pi / 2$, and $f = 2^{1/2}$.

Let $I(x,y)$ be the gray level distribution of an image, the convolution of the image I and a Gabor kernel $\psi_{u,v}$ is defined as follows:

$$O_{u,v} = I(z) * \psi_{u,v}(z) \quad (4)$$

Where $*$ denotes the convolution operator, and $O_{u,v}(z)$ is the convolution result corresponding to the Gabor kernel at orientation u and scale v .

The energy information of the image $I(x,y)$ can be expressed as follows[10]:

$$E(u,v) = \sum_x \sum_y |O_{u,v}| \quad (5)$$

In this work, we propose to use the mean and the standard deviation as texture features. Suppose that $I(x,y)$ denotes an image of $M \times N$ pixels, $\mu_{uv}(x)$ and $\sigma_{uv}(x)$ denote its mean and standard deviation computed from the scale v and the direction u , respectively, $\mu_{uv}(x)$ and $\sigma_{uv}(x)$ can be computed as follows:

$$\mu_{u,v} = E_{u,v} / MN \quad (6)$$

$$\sigma_{u,v} = \sqrt{\frac{\sum_x \sum_y (|O_{u,v}(z) - \mu_{u,v}|)^2}{MN}} \quad (7)$$

5 The proposed method

In order to evaluate our approach, we calculated two descriptors:

Descriptor n°1: computed by applying Gabor wavelet transform with sixteen orientations and ten scales we have extracted 160 elements of means and 160 elements of deviation from the first BIMF. Therefore, the total dimension of the first descriptor is 320.

Descriptor n°2: calculated from the first, the second and the third BIMF, we have extracted 160 elements of means and 160 elements of deviation for each one of them. Therefore, the total dimension of the second descriptor is 960.

5.1 Evaluation methodology

The evaluation of our experiments is performed on the basis of completeness and accuracy of the predicted annotations, with reference to the manual annotations of the test dataset. Completeness is defined as the number of predicted features divided by the total number of features, equation (8), while accuracy is the number of correctly predicted features divided by the total number of predicted features, equation (9).

$$\text{Completeness} = \frac{\text{number of predicted UsE features}}{\text{Total number of UsE features}} \quad (8)$$

$$\text{Accuracy} = \frac{\text{number of corrected predicted UsE features}}{\text{Number of predicted UsE features}} \quad (9)$$

$$\text{TotalScore} = \sqrt{\text{Completeness} \times \text{Accuracy}} \quad (10)$$

The following table shows the score results obtained by the two proposed descriptors applied for predicting different groups of UsE features.

Table 2. The score results obtained for each group

	Descriptor n°1		Descriptor n°2	
	<i>Cmp</i>	<i>Acc</i>	<i>Cmp</i>	<i>Acc</i>
Liver	1.00	0.92	1.00	0.92
Vessel	1.00	1.00	1.00	1.00
Lesion Area	1.00	0.65	1.00	0.67
Lesion Lesion	1.00	0.48	1.00	0.51
Lesion component	0.96	0.86	0.96	0.89
Total score	0.880		0.889	

The results presented in table 2 show that the scores obtained by the two descriptors for Liver group and Vessel group are the same. This can be explained by the fact that there were some instances having the same annotation (unbalanced dataset) in all the training samples, and this is the major challenge in this dataset. We notice that the second descriptor outperforms the first descriptor in the other groups. The

best result is obtained by using three BIMFS with the score of 88.9% compared to the first descriptor with the score of 88%. This suggests that, using more details improves the accuracy of the annotation.

Table 3 shows that all authors have completed the vessel UsE features with high accuracy, also they completed the liver UsE features with accuracy more than 80% except the method proposed in [5].

Concepts related to lesion-lesion are annotated completely by [2] [3] and our proposed approach. However, [3] has got the accuracy of 83% by using the second feature set of size 1446, while we achieved 51%. We attribute this difference to the size of feature set which has significantly improved the accuracy of the annotation.

Lesion-components group are fully completed and annotated with an accuracy higher than 90% by both [3] and [4]. In our proposed approach we achieved a completeness score of 96%, and accuracy of 89%.

It is interesting to note that one of the major considerations in the annotation of the liver is the appropriate combination of features and methods.

Table 3. Comparison test results presented by different authors and summarized in [6]

Methods presented In	Liver		Vessel		Lesion-Area		Lesion-Lesion		Lesion-Component	
	Cmp	Acc	Cmp	Acc	Cmp	Acc	Cmp	Acc	Cmp	Acc
[2]	1.00	0.92	1.00	1.00	1.00	0.75	1.00	0.48	0.96	0.89
[3]	1.00	0.93	1.00	1.00	0.92	0.79	1.00	0.83	1.00	0.94
[4]	1.00	0.93	1.00	1.00	0.85	0.81	0.90	0.82	1.00	0.94
[5]	0.62	0.88	1.00	1.00	0.46	0.77	0.20	1.00	0.12	0.15
Descriptor 2	1.00	0.92	1.00	1.00	1.00	0.67	1.00	0.51	0.96	0.89

6 Conclusion

In this paper, we have presented a new approach aiming to annotate liver CT scans with the use of content-based medical image retrieval. We first extracted a set of features descriptors based on spectral images textural analysis that uses a combination of Bidimensional Empirical Mode Decomposition (BEMD) and Gabor wavelet transform, finally the five most similar training images was used to select the answers for an unannotated image. Our experiments results have been conducted by using the ImageClef 2015 annotation dataset. Our best score of 88.9% was achieved by the second descriptor. In our futures works we plan to use more image features and examining others classifiers.

ACKNOWLEDGMENT

The authors would like to thank Mrs Neda Marvasti from Bogaziçi University, Electrical and Electronics Department, Istanbul, Turkey , for making the database available for the experiments.

REFERENCES

- [1] N.Marvasti et al., "ImageCLEF Liver CT Image Annotation Task 2014. In: CLEF 2014 Evaluation Labs and Workshop", Online Working Notes. (2014).

- [2] I.Nedjar, S.Mahmoudi, A.Chikh, K.Abi-ayad, and Z.Bouafia, "Automatic annotation of liver CT image:ImageCLEFmed 2015," in CLEF2015 Working Note.CEUR Workshop Proceedings,CEUR-WS.org, Toulouse,france, Septembre 8-11 2015.
- [3] A.Kumar, S.Dyer, C.Li, P.H.W.Leong, and J.Kim, "Automatic annotation of liver ct images: the submission of the bmet group to imageclefmed 2014", in CLEF 2014 Labs and Workshops, Notebook Papers. CEUR, Workshop Proceedings (CEUR-WS.org), September 2014.
- [4] A.B.Spanier and L.Joskowicz, "Towards content-based image retrieval: From computer generated features to semantic descriptions of liver ct scans", in CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (CEUR-WS.org), September 2014.
- [5] B.Ermis and A.T.Cemgil, "Liver ct annotation via generalized coupled tensor factorization", in CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings (CEUR-WS.org), September 2014.
- [6] N.Marvasti et al., "Overview of the ImageCLEF 2015 liver CT annotation task," in CLEF2015 Working Notes. Workshop Proceedings.CEUR-WS.org, no. 1613-0073, September 8-11. 2015.
- [7] N.E.Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. London. A 454, 903-995, 1998.
- [8] J.C.Nunes, Y.Bouaoune, E.Delechelle, O.Niang, and P.Bunel, "Image analysis by bidimensional empirical mode decomposition," Image and Vision Computing 21 1019–1026, 2003.
- [9] C.Liu and H.Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," IEEE Transactions on Image Processing, vol. 11, no. 4, pp. 467-476, 2002.
- [10] G.Zhang, ZM.Ma, and L.Deng, "Texture Feature Extraction and Description Using Fuzzy Set of Main Dominant Directions of Variable Scales in Content-Based Medical Image Retrieval," in SAC '08 Proceedings of the ACM symposium on Applied computing, Fortaleza, Ceará, Brazil, March 16-20, 2008, pp. 1760-1761.