Textural Analysis of Liver Focal Lesions with Co-occurrence Matrix and Wavelet Transform on CT: A Feasible Study in FNH, HEM and HCC

Jia-jun QIU¹, Yue WU¹, Bei HUI², Jia CHEN² and Min WANG³

¹School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610000, China
²School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610000, China
³Department of Radiology West China Hospital, Sichuan University, Chengdu 610000, China

Keywords: Texture analysis, GLCM, Wavelet transform, Liver focal lesion, FNH, HEM, HCC.

Abstract. Focal nodular hyperplasia (FNH), hepatocellular carcinoma (HCC) and cavernous hemangiomia (HEM) are three types of solid focal liver lesions. Using CT images to identify these three types of lesions is the most commonly method. However, this method usually mainly depends on experiences. Whereas, more objective and quantitative image information could be explored with texture analysis method. This research aims to discuss the appreciations of texture analysis based on CT images for identifying FNH, HEM and HCC.

This paper retrospectively analyzed 81 clinically or pathologically diagnosed cases, each of which contains contrast non-enhanced and contrast two-phasic enhanced CT images. The texture analysis was based on gray level co-occurrence matrix (GLCM) and wavelet transform. The results shows that the misclassification rates of texture classification were low to 2.73% between FNH and HEM (between benign lesions), 3.19% between FNH and HCC (between benign lesions and malignant lesions), and 1.67% between HEM and HCC (between benign lesions and malignant lesions) respectively. The effect of texture classification based on contrast two-phasic enhanced CT images was better than contrast non-enhanced CT images.

Introduction

CT images were commonly used to identify various solid focal lesions in liver, which was very important for the making of treatment method in later. However, it still depends on experiences greatly. In the last 10 years, quantitative analysis of distribution patterns of gray level in CT images were generally used to the auxiliary diagnosis [1]. In these auxiliary diagnosis, plain scanning CT images were more used, then contrast enhanced CT images were come to the focus in recent years because it could provide some other information about blood-supply [2].

In this paper, contrast non-enhanced and enhanced (arterial phase and portal phase) CT images were both analyzed based on texture features of FNH, HEM and HCC. 81 cases of solid focal liver lesions in West China Hospital of Sichuan University from 2015-1-1 to 2015-12-31 were analyzed retrospectively, where 21 cases were FNH, 34 case were HEM and 26 cases were HCC. Two slice images were selected in each case. In these cases, FNH and HCC were confirmed by pathology, HEM was confirmed by classical symptom of CT images and follow-up visits. Thus, a total of 162 CT images were analyzed.

Methods of Texture Analysis

First, ROIs that represent the lesions were manually selected. This paper used gray level co-occurrence matrix and wavelet transform to extract texture features of ROIs. After that, Fisher coefficient method was used to select and reduce features. Then, raw data analysis (the scatter plot of 3 first features), linear discriminant analysis (LDA) and K-NN classifier were applied. The whole process was shown in Figure 1.
Extraction of Texture Features

As for 2-D images, GLCM [3] is defined to a 4-D matrix, which describes the spatial distribution of gray-level co-occurring values at a given offset (offset of distance and direction). An element of GLCM could be expressed with $m(a, b, d, \theta)$. The value of the element is calculated how often eligible pixel pair occurs. Eligible pixel pair $(i, j)$ is at distance $d$ and direction $\theta$. An illustrative example is shown in Figure 2.

In this research, 11 texture features were extracted from GLCM as given in Table 1. Where the distance $d$ equaled to 1 and the angle $\theta$ were respectively $0^\circ$, $45^\circ$, $90^\circ$ and $135^\circ$. A total of 44 features were extracted with GLCM.

| Methods               | Features                                                                 |
|-----------------------|--------------------------------------------------------------------------|
| GLCM                  | angular second moment; contrast; correlation; sum of squares; inverse difference moment; sum average; sum variance; sum entropy; entropy; difference variance; difference entropy; |
| Wavelet transform     | entropy                                                                  |

An 2-D image could be expressed by a two-dimensional function $f(x, y)$. Therefore, 2-D wavelets are required from the vector product of $\psi(x)$ and $\varphi(x)$, which are defined from Eq. 1:

$$
\begin{align*}
\varphi(x, y) &= \varphi(x) \varphi(y) \\
\psi^H(x, y) &= \psi(x) \varphi(y) \\
\psi^V(x, y) &= \varphi(x) \psi(y) \\
\psi^D(x, y) &= \psi(x) \psi(y)
\end{align*}
$$

(1)

where $\varphi(x, y)$ is a 2-D scaling function, $\psi^H$, $\psi^V$ and $\psi^D$ are 2-D wavelets, $H$, $V$ and $D$ represent the horizontal, vertical and diagonal directions respectively. It means that an image could be discomposed to 4 sub-band images, where the original information is remained in one sub-band image and the other 3 sub-band images represent the details of horizontal, vertical and diagonal directions respectively [4,5]. These 3 sub-band images are the high frequency parts of the original image with craggy variances and some certain patterns. These variances can generally indicate the texture features. 3 texture features could be extracted from the coefficient matrix of wavelet transform [6].

Haar [7] wavelet was used in this research. It is a sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. As a special case of the Daubechies wavelet, the Haar wavelet is also known as Db1. The Haar wavelet's mother wavelet function can be described as Eq. 2:
\[
\psi(x) = \begin{cases} 1 & 0 \leq x < 0.5 \\ -1 & 0.5 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)
\]

Its scaling function can be described as Eq. 3:

\[
\phi(x) = \begin{cases} 1 & 0 \leq x < 1, \\ 0 & \text{otherwise} \end{cases}
\quad (3)
\]

Energy of coefficient matrix was used in this paper. It was computed at 5-scales wavelet within 4 frequency bands. So, a total of 20 features were extracted with wavelet transform.

**Feature Selection**

Fisher coefficient [8] is calculated for the extracted texture features. It is defined as a ratio of between-class variance to within-class variance. In this paper, top 10 features were selected based on Fisher coefficient value from larger to small.

**Data Analysis**

The selected features were standardized and saved. After that, raw data analysis (the scatter plot of 3 first features) and LDA were applied. Then, the k-nearest neighbors (K-NN) classifier would be applied based on raw data and LDA. The result would be discussed in next section.

**Result**

**MaZda**

This research used the MaZda [9-11] software to complete the texture analysis. MaZda is a computer program for calculation of texture parameters (features) in digitized images. It has been under development since 1998, to satisfy the needs of the participants of COST B11 European project "Quantitative Analysis of Magnetic Resonance Image Texture" (1998-2002) and COST B21 European project "Physiological modelling of MR Image formation".

**Group**

The data set were divided into four groups as show in Table 2. Group 1 is between benign lesions (HEM and FNH). FNH is a rare benign liver lesion, and HEM is a very common benign liver lesion.

| Number | 1 | 2 | 3 | 4 |
|--------|---|---|---|---|
| Category | FNH vs HEM | FNH vs HCC | HEM vs HCC | FNH vs HEM vs HCC |

Group 2 is between benign lesion and malignant lesion (FNH and HCC). HCC is a very common malignant liver lesion. Group 3 is between benign lesion and malignant lesion (HEM and HCC). Group 4 is between benign lesion and malignant lesion (FNH, HEM and HCC).

In each group, there are 3 phases: contrast non-enhanced phase, arterial phase of contrast enhanced, portal phase of contrast enhanced. To evaluating the effect of texture analysis, misclassification rate (MCR) was used in each group for 3 phases. According the value of MCR, it was divided into 5 levels as demonstrated in Table 3.

| Level | Excellent | Good | Medium | Ordinary | Bad |
|-------|-----------|------|--------|----------|-----|
| MCR   | MCR ≤ 10% | 10%<MCR ≤ 20% | 20%<MCR ≤ 30% | 30%<MCR ≤ 40% | MCR>40% |
Analysis

This paper did not divide the data into a training and a test dataset, because of its relatively not big scale of size, and because the main interest in this paper is the general feasibility of texture analysis based on GLCM and wavelet transform of various phases of CT images in FNH, HEM and HCC, rather than specifically in the discrimination of liver lesions in clinical diagnosis. In addition, the K-NN classifier implemented in the MaZda program uses the leave one out testing technique, which does not require a separate training data set [9]. The experimental results are illustrated in Table 4.

Table 4. Experimental results.

| Group | Phrase       | MCR (Raw data analysis) | MCR (LDA)    |
|-------|--------------|-------------------------|--------------|
| 1:    | FNH vs HEM   | Non-enhanced 12/110=11.91% Good 26/110=23.64% Medium |
|       |              | Arterial 3/110=2.73% Excellent 5/110=4.55% Excellent |
|       |              | Portal 3/110=2.73% Excellent 4/110=3.64% Excellent |
| 2:    | FNH vs HCC   | Non-enhanced 15/94=15.96% Good 14/94=14.89% Good |
|       |              | Arterial 4/94=4.26% Excellent 6/94=6.38% Excellent |
|       |              | Portal 5/94=5.32% Excellent 3/94=3.19% Excellent |
| 3:    | HEM vs HCC   | Non-enhanced 17/120=14.17% Good 26/120=21.67% Medium |
|       |              | Arterial 11/120=9.17% Excellent 16/120=13.33% Good |
|       |              | Portal 2/120=1.67% Excellent 12/120=10.00% Excellent |
| 4:    | FNH vs HEM vs HCC | Non-enhanced 22/162=13.58% Good 42/162=25.93% Medium |
|       |              | FNH+HEM vs HCC: 15/162=9.26% Excellent 23/162=14.2% Good |
|       |              | Arterial 27/162=16.67% Good 42/162=25.93% Medium |
|       |              | FNH+HEM vs HCC: 20/162=12.3% Good 34/162=21% Medium |
|       |              | Portal 14/162=8.64% Excellent 30/162=18.52% Good |
|       |              | FNH+HEM vs HCC: 11/162=6.79% Excellent 25/162=15.43% Good |

In group 4 of Table 4, FNH+HEM vs HCC means the misclassification between FNH and HCC was eliminated, because the objective of the fourth group experiments was to observe the MCR between benign and malignant lesions. In all these experiments, K-NN classifier based on raw data analysis were generally better than based on LDA. They are shown in Figure 3 and Figure 4.

HEM vs FNH: This was the identification between benign lesions. They were excellent or good in MCR. The classification based on arterial phase and portal phase had the lower MCR, they were both 2.73%.

FNH vs HCC: This was the identification between benign lesions and malignant lesions. The MCR of identification based on enhanced CT images were all excellent. The classification based on portal phase was the lowest MCR, it was low to 3.19%.

HEM vs HCC: This was the identification between the most common benign liver lesions and the most common malignant liver lesions. The MCR of identification based on enhanced CT images were good or excellent. The classification based on portal phase was the lowest MCR, it was low to 1.67%.

FNH+HEM vs HCC: These experiments was classification of three classes. But the objective of the fourth group experiments was to observe the MCR between benign and malignant lesions. The MCR of identification based on enhanced CT images were good or excellent. The classification based on portal phase was the lowest MCR, it was low to 6.79%.

Besides, the MCR from group 1 to group 4 in non-enhanced phase were 11.91%, 14.89%, 14.17% and 9.26%. They were all good or excellent.
Summary

This paper shows that texture analysis, based on the contrast enhanced CT images to identify the solid focal liver lesions, was more efficient than the contrast non-enhanced CT images. Enhanced CT images provide easier distinguished texture information (mainly refers to blood-supply) between solid focal liver lesions.

FNH and HEM are benign lesions that both have rich blood supply relatively to HCC. It means the manual MCR for these two lesions with CT images will be higher.

HCC have different information of blood supply from FNH and HEM. The research by Gletsos [12] etc., showed that the MCR for identifying normal liver organism, HEM, HCC and hepatic cyst were low to 15.79% based on texture analysis of non-enhanced CT images. Yu-Len Huang et al [13] extracted auto-covariance features for classifying malignancies with an accuracy of 81.7% by an SVM classifier. This paper showed that the MCR for identifying based on non-enhanced CT images between HEM and HCC were low to 14.17%, which is more efficient than the previous.

In order to get more complete diagnostic conclusions, physicians often need to consider comprehensively lesion characteristics of plain and enhanced scans. Although, the researches of texture analysis about the ability of identifying solid focal liver lesions, which combined with
contrast non-enhanced CT images and contrast triphasic enhanced CT images, were uncommon, but texture analysis could become the way to improve the accuracy of identification. Besides, texture features in this paper were only 2-D features. If the extraction of texture features were based on 3-D model, the result will be much better [14].

Acknowledgement

The research work was supported by the Fundamental Research Funds for the Central Universities under Grant No. ZYGX2015J068.

References

[1] Van Ginneken B., Romeny B.M.T.H., Viergever M.A. Computer-aided diagnosis in chest radiography: a survey [J]. IEEE Transactions on Medical Imaging, 2001, 20(12): 1228-1241.
[2] Militzer A., Hager T., Jager F., et al. Automatic detection and segmentation of focal liver lesions in contrast enhanced CT images [C]//Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010: 2524-2527.
[3] Z Haralick R.M., Shanmugam K. Textural features for image classification [J]. IEEE Transactions on systems, man, and cybernetics, 1973 (6): 610-621.
[4] Arivazhagan S., Ganesan L. Texture classification using wavelet transform[J]. Pattern Recognition Letters, 2003, 24(9): 1513-1521.
[5] Nanthagopal A.P., Rajamony R.S. Classification of benign and malignant brain tumor CT images using wavelet texture parameters and neural network classifier [J]. Journal of Visualization, 2013, 16(1): 19-28.
[6] Hu S., Xu C., Guan W.Q., et al. Texture feature extraction based on wavelet transform and gray-level co-occurrence matrices applied to osteosarcoma diagnosis [J]. Bio-medical Materials and Engineering, 2014, 24(1): 129-143.A.
[7] Haar A. Zur theorie der orthogonalen funktionensysteme [J]. Mathematische Annalen, 1910, 69(3): 331-371.
[8] Fisher R.A. The use of multiple measurements in taxonomic problems [J]. Annals of Eugenics, 1936, 7(2): 179-188.
[9] M. Strzelecki, P. Szczypinski, A. Materka, A. Klepaczko, A software tool for automatic classification and segmentation of 2D/3D medical images, Nuclear Instruments & Methods In Physics Research A, 702, 2013, pp. 137-140.
[10] P. Szczypinski, M. Strzelecki, A. Materka, A. Klepaczko, MaZda-A software package for image texture analysis, Computer Methods and Programs in Biomedicine, 94(1), 2009, pp 66-76.
[11] P. Szczypinski, M. Strzelecki, A. Materka, MaZda - a Software for Texture Analysis, Proc. of ISITC 2007, November 23-23, 2007, Republic of Korea, pp. 245-249.
[12] Gletsos M., Mougiakakou S.G., Matsopoulos G.K., et al. A computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier [J]. IEEE Transactions on Information Technology in Biomedicine, 2003, 7(3): 153-162.
[13] Yu-Len Huang, Jeon-Hor Chen, Wu-Chung Shen, "Diagnosis of hepatic tumors with texture analysis in non enhanced computed tomography images", Academic Radiology, Vol. 13, pp 713-720, 2006.
[14] Gillies R.J., Kinahan P.E., Hricak H. Radiomics: images are more than pictures, they are data [J]. Radiology, 2015, 278(2): 563-577.