# Computer-Aided Diagnosis for Feature Selection and Classification of Liver Tumors in Computed Tomography Images

Wen-Jia Kuo

Department of Information Management, Yuan Ze University
Taoyuan, Taiwan, R.O.C.
Innovation Center for Big Data and Digital Convergence, Yuan Ze University
Taoyuan, Taiwan, R.O.C.

wjkuo@saturn.yzu.edu.tw

**Abstract**

We propose a computer-aided diagnosis (CAD) system to classify liver tumors in non-enhanced computed tomography (CT) images. There are three parts in our proposed system. First, the feature extraction module extracts 102 statistical texture features. Second, the feature selection module acquired the combination of the best features by integrating the particle swarm optimization (PSO) algorithm with support vector machine (SVM) to reduce the complexity of computation. Finally, a SVM based classification model was constructed to identify benign and malignant liver tumors. Experiments show the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the proposed CAD system for classifying liver tumors was 84.53%, 80%, 88.09%, 88.77%, and 84.01%, respectively. The accurate rate is up to 80 % both on benign and malignant tumors of CT images. We can find that the proposed method can achieve the purpose of enhancing the accuracy of automatic identify effectively to assist further diagnosis.

**Key words:** Computed Tomography, Particle Swarm Optimization, Support Vector Machine, Fuzzy Logic System

## Introduction

Currently, the most common imaging method to detect liver tumor is Computed tomography (CT). In order to achieve the purpose of early detection and early treatment, this paper presents a computer-aided diagnosis system based on statistics and texture-based features. To reduce the dimension and the complexity, optimization algorithm was used to select the feasible features. We also use support vector machine to identify the classification work to assist the doctor for further differential diagnosis.

According to the relevant literatures, texture features are roughly divided into statistical methods such as the First Order Statistics [1] and the autocorrelation coefficient [2]. The other uses the inter-pixel texture features, such as Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run-Length Textural Matrix (GLRLM), and so on. To select the feasible features, optimization algorithms such as Genetic Algorithm (GA) [1], Sequential Forward Selection (SFS), Sequential Floating Forward Selection (SFFS), Dipolar Decision Tree, Simulated Annealing (SA), Ant colony Optimization (ACO) and Particle Swarm Optimization (PSO) [3] were used. For the classifier, Decision Tree, Voting Scheme [4], Neural Network [5], [6], Support Vector Machine [7], k-nearest neighbor (k-NNR) [8] and Bayes Classifiers [9] are commonly used methods.

To establish an effective computer aided diagnosis system for CT scan images, we will calculate the image features based on texture and statistics. Combined with BoltzMann probability function, an improved particle swarm optimization algorithm was used to find the best combination of different features. Finally, support vector machine use the selected features to construct the classification model. Combined with the fuzzy logic theory and weighted values, we can find the classification results of benign and malignant liver tumors.

## Proposed Method

In this section, we will introduce the details for the classification of liver tumors in CT images.

### *A. Features*

Many kinds of features can be used to represent an image. In this paper, features based on texture and statistics were mainly used to differentiate different images. In the ROI of CT liver tumor images, 102 image features were extracted. The details are described as follows.

(a) Statistical Feature Matrix (SFM)

The Statistical Feature Matrix (SFM) is a statistical feature matrix used to calculate the pixel-to-pixel distance over a fixed range in the image [10]. We use the dissimilarity matrix to find the Dissimilarity ($F_1$) between adjacent pixels with a distance of 5, and the formula is as follows:

$$Dissimilarity = \frac{\sum_{i=0}^{\Delta x-1}\sum_{j=0}^{\Delta y-1} a_{ij}}{\Delta x \Delta y}$$

(b) Gray Level Co-occurrence Matrix (GLCM)

Haralick et al proposed the Gray Level Co-occurrence Matrix (GLCM) in 1973 [11]. Using the GLCM method, we can get a matrix G. In this paper, we take the average of gray level co-occurrence matrix $G_1$~$G_4$ with distance 1 and 4 different degrees as the feature matrix $G$. We can calculate 13 feature values ($F_2$ ~ $F_{14}$) from the grayscale co-occurrence

*ISBN 978-1-5386-4342-6*

matrix *G* as follows.

$F_2$: Angular Second Moment (Energy)

$$Energy = \sum_{i=0}^{N_g}\sum_{j=0}^{N_g}(p(i,j))^2$$

$F_3$: Contrast

$$Contrast = \sum_{n=0}^{N_g} n^2\left[\sum_{i=0}^{N_g}\sum_{j=0}^{N_g} p(i,j)\right],\ |i-j| = n$$

$F_4$: Correlation

$$Correlation = \frac{\sum_{i=0}^{N_g}\sum_{j=0}^{N_g}(ij)p(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y}$$

$F_5$: Variance

$$Variance = \sum_{i=0}^{N_g}\sum_{j=0}^{N_g}(i-\mu)^2 p(i,j)$$

$F_6$: Sum Average

$$SumAverage = \sum_{i=2}^{2N_g} iP_{x+y}(i)$$

$F_7$: Sum Variance

$$SumVariance = \sum_{i=2}^{2N_g}(i - Varuance)^2 P_{x+y}(i)$$

$F_8$: Entropy

$$Entropy = -\sum_{i=0}^{N_g}\sum_{j=0}^{N_g} p(i,j)\log(p(i,j))$$

$F_9$: Sum Entropy

$$SumEntropy = -\sum_{i=2}^{2N_g} P_{x+y}(i)\log(P_{x+y}(i))$$

$F_{10}$: Difference Variance

$$DifferenceVariance = \sum_{i=0}^{N_g}[i - \sum_{j=0}^{N_g} jP_{x-y}(j)]^2 P_{x-y}(i)$$

$F_{11}$: Difference Entropy

$$DifferenceEntropy = -\sum_{i=0}^{N_g} P_{x-y}(i) log[P_{x-y}(i)]$$

$F_{12}$: Local Homogeneity

$$LocalHomogeneity = \sum_{i=0}^{N_g}\sum_{j=0}^{N_g}\frac{p(i,j)}{1+(i-j)^2}$$

$F_{13}$: Cluster Shade

$$ClusterShade = \sum_{i=0}^{N_g}\sum_{j=0}^{N_g}(i - M_x + j - M_y)^3 p(i,j)$$

$F_{14}$: Cluster Prominence

$$ClusterProminence = \sum_{i=0}^{N_g}\sum_{j=0}^{N_g}(i - M_x + j - M_y)^4 p(i,j)$$

(c) Gray Level Run-Length Textural Matrix (GLRLM)

Gray Level Run-Length Textural Matrix (GLRLM) was proposed by Haralick et al. in 1975, [12]. With the GLRLM method, we can get a matrix *M*. We can calculate 5 feature values ($F_{15}$~$F_{19}$) from matrix *M* as follows.

$F_{15}$: Short Run Emphasis

$$ShortRunEmphasis = \frac{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1}\frac{d(i,j)}{(j+1)^2}}{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1} d(i,j)}$$

$F_{16}$: Long Run Emphasis

$$LongRunEmphasis = \frac{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1}(j+1)^2 d(i,j)}{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1} d(i,j)}$$

$F_{17}$: Gray Level Non-uniformity

$$GrayLevelNonuniformity = \frac{\sum_{i=0}^{N_g}(\sum_{j=0}^{l-1} d(i,j))^2}{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1} d(i,j)}$$

$F_{18}$: Run Length Non-uniformity

$$RunLengthNonuniformity = \frac{\sum_{j=0}^{l-1}(\sum_{i=0}^{N_g} d(i,j))^2}{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1} d(i,j)}$$

$F_{19}$: Run Percentage

$$RunPercentage = \frac{\sum_{i=0}^{N_g}\sum_{j=0}^{l-1} d(i,j)}{H \times W}$$

(d) Law's Texture Energy Matrix

Law's Textural Energy Matrix (TEM) [13]mainly use three one-dimensional core vector to perform convolution with each other to obtain three two-dimensional core vectors. Then, it uses the two-dimensional core vector and its transpose vector to perform convolution to get five $5\times5$ masks. With image *I* and the five $5\times5$ masks, we can achieve 5 averaging features ($F_{20}$~$F_{24}$) and 5 variation features ($F_{25}$~$F_{29}$) as follows:

$F_{20}$~$F_{24}$:

$$LE_{mean} = \frac{\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]}{(H-1)(W-1)},\ EL_{mean} = \frac{\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times EL]}{(H-1)(W-1)},$$

$$SL_{mean} = \frac{\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times SL]}{(H-1)(W-1)},\ EE_{mean} = \frac{\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times EE]}{(H-1)(W-1)},$$

$$LS_{mean} = \frac{\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LS]}{(H-1)(W-1)}$$

$F_{25}$~$F_{29}$:

$$LE_{var} = \frac{\left(\left(\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]\right) - LE_{mean}\right)^2}{(H-1)(W-1)},\ LE_{var} = \frac{\left(\left(\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]\right) - LE_{mean}\right)^2}{(H-1)(W-1)},$$

$$SL_{var} = \frac{\left(\left(\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]\right) - SL_{mean}\right)^2}{(H-1)(W-1)},\ EE_{var} = \frac{\left(\left(\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]\right) - EE_{mean}\right)^2}{(H-1)(W-1)},$$

$$LS_{var} = \frac{\left(\left(\sum_{i=2}^{H-2}\sum_{j=2}^{W-2}[mask(i,j)\times LE]\right) - LS_{mean}\right)^2}{(H-1)(W-1)}$$

(e) Gray Level Dependence Textural Matrix (GLDM)

Gray level dependent texture matrix [13] is the number of times the adjacent pixels with the same pixel value as their own pixels are counted in the distance *d*, and thus the brightness relationship of pixels is obtained. According to the gray level dependence textural matrix Q, five features can be calculated as $F_{30}$~ $F_{34}$, respectively.

$F_{30}$: Small Number Emphasis (*SNE*)

$$SNE = \frac{\sum_K\sum_S\{\frac{Q(K,S)}{S^2}\}}{R}$$

$F_{31}$: Large Number Emphasis *(LNE)*



*ISBN 978-1-5386-4342-6*

$$LNE = \frac{\sum_K \sum_S \{S^2 Q(K,S)\}}{R}$$

$F_{32}$: Number Non-uniformity (*NNU*)

$$NNU = \frac{\sum_S [\sum_K Q(K,S)]^2}{R}$$

$F_{33}$: Second Moment (*SM*)

$$SM = \frac{\sum_K [\sum_S Q(K,S)]^2}{R}$$

$F_{34}$: Entropy

$$Entropy = -\frac{\sum_K \sum_S \{Q(K,S)\log[Q(K,S)]\}}{R}$$

(f) Wavelet Features

Wavelet was proposed by Morlet and Grossman in 1987 [2]. In this paper, we use the *LL* sub-band of Haar Transform to achieve 12 feature values ($F_{35} \sim F_{46}$).

$F_{35}$: Mean

$$Mean = \frac{\sum_{i=0}^{h} \sum_{j=0}^{w} LL(i,j)}{h \times w}$$

$F_{36}$: Standard Deviation

$$StandardDeviation = \sqrt{\frac{\left(\sum_{i=0}^{h} \sum_{j=0}^{w} LL(i,j) - Mean\right)^2}{h \times w}}$$

$F_{37}$~$F_{41}$: mean with Law's masks

$$LL\_LE_{mean} = \frac{\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]}{(h-1)(w-1)}, \quad LL\_EL_{mean} = \frac{\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times EL]}{(h-1)(w-1)},$$

$$LL\_SL_{mean} = \frac{\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times SL]}{(h-1)(w-1)}, \quad LL\_EE_{mean} = \frac{\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times EE]}{(h-1)(w-1)},$$

$$LL\_LS_{mean} = \frac{\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LS]}{(h-1)(w-1)}$$

$F_{42}$~$F_{46}$: variance with Law's masks

$$LL\_LE_{var} = \frac{\left(\left(\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]\right) - LE_{mean}\right)^2}{(h-1)(w-1)}, \quad LL\_EL_{var} = \frac{\left(\left(\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]\right) - EL_{mean}\right)^2}{(h-1)(w-1)},$$

$$LL\_SL_{variance} = \frac{\left(\left(\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]\right) - SL_{mean}\right)^2}{(h-1)(w-1)}, \quad LL\_EE_{variance} = \frac{\left(\left(\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]\right) - EE_{mean}\right)^2}{(h-1)(w-1)},$$

$$LL\_LS_{variance} = \frac{\left(\left(\sum_{i=2}^{h-2} \sum_{j=2}^{w-2} [LL(i,j) \times LE]\right) - LS_{mean}\right)^2}{(h-1)(w-1)}$$

(g) Local Features of Fourier Coefficients

Consider the Fourier transformation [14] with center pixel $f(x,y)$, we can find the magnitude and phase angle used to evaluate the mean and variance values.

$F_{47}$~$F_{54}$: Magnitude Mean

$$M(s,t) = \frac{\sum_{x=1}^{H-1} \sum_{y=1}^{W-1} Magnitude(u+s, v+t)}{(H-1)(W-1)}$$

$F_{55}$~$F_{62}$: Magnitude Variance

$$SD\_M(s,t) = \sqrt{\frac{\left(\left(\sum_{i=1}^{H-1} \sum_{j=1}^{W-1} Magnitude(u+s, v+t)\right) - M(s,t)\right)^2}{(H-1)(W-1)}}$$

$F_{63}$~$F_{70}$: Phase Angle Mean

$$PA(s,t) = \frac{\sum_{x=1}^{H-1} \sum_{y=1}^{W-1} Phase\_angle(u+s, v+t)}{(H-1)(W-1)}$$

$F_{71}$~$F_{78}$: Phase Angle Variance

$$SD\_PA(s,t) = \sqrt{\frac{\left(\left(\sum_{i=1}^{H-1} \sum_{j=1}^{W-1} Phase\_angle(u+s, v+t)\right) - PA(s,t)\right)^2}{(H-1)(W-1)}}$$

(h) Autocorrelation Coefficient ($F_{79}$~$F_{102}$)

The mean-removed autocorrelation coefficients were used to represent interpixel correlations. It was defined as

$$\gamma(\Delta m, \Delta n) = \frac{A'(\Delta m, \Delta n)}{A(0,0)},$$

where

$$A'(\Delta m, \Delta n) = \frac{1}{(m-\Delta m)(n-\Delta n)} \times \sum_{x=0}^{m-1-\Delta m} \sum_{y=0}^{n-1-\Delta n} \left| \begin{matrix} [f(x,y) - \bar{f}] \\ \times [f(x+\Delta m, y+\Delta n) - \bar{f}] \end{matrix} \right|,$$

and $\bar{f}$ represents the mean value of $f(x,y)$ [2].

## *B. Best Features Selection*

The best feature combination for classification can be found in two stages. In the first stage, the modified optimal particle swarm optimization algorithm is combined with support vector machine. The main purpose of this stage is to reduce the number of features to less than half. Thus, the performance of the classification model constructed by the selected features can reached a certain requirement. Next , the features selected in the first stage are further combined through the vote mechanism so that the number of features is less than 20. The criterion of voting in the voting mechanism is to evaluate the 5 evaluation value of classification result for the constructed support vector machine model.

## *C. Classification with Support Vector Machine Model and Fuzzy Logic*

After the previous two-stage best features selection steps, the best feature combination used to construct the classification will be achieved. Then, we combined the fuzzy logics weighted values with voting mechanism to obtain the classification result. There are two stages in the classification procedures. First, several classification models will be constructed by using support vector machines. In the second stage, several classification results are obtained in the first stage. The classification results are further voted with fuzzy logic weights to get the final classification results. The main purpose of constructing several classification models in the first stage is to cooperate with the next stage to have more than two groups to carry out weighted voting so that the final classification result is more objectivity.

# **Results**

Sensitivity ($Sn$), specificity ($Spc$), accuracy ($Acc$), positive predictive value ($PPV$), and negative predictive value ($NPV$) were used to measure the performance of the proposed method. They are defined as follows:

$Sn = TP / (TP + FN)$ ,

$Spc = TN / (TN + FP)$ ,

$Acc = (TP + TN) / (TP + FP + TN + FN)$ ,

$TP$ , $TN$ , $FP$ , and $FN$ are true positives, true negatives, false positives, and false negatives, respectively.

In our experiments, 164 images (84 benign and 80 malignant) were used to evaluate the overall performance of the proposed model. Table 1 shows the summary of classification results.

Table 1 Classification results.

| | Proposed method | |
|---|---|---|
| | Without Fuzzy | With Fuzzy |
| Accuracy | 50.00% | 84.31% |
| Sensitivity | 38.10% | 80% |
| Specificity | 61.91% | 88.09% |
| PPV | 50.66% | 88.77% |
| NPV | 48.88% | 84.01% |

The accuracy, sensitivity, specificity, PPV, and NPV of computer-aided diagnosis system proposed in this study were 84.31%, 80%, 88.09%, 88.77%, and 84.01%, respectively. All the five evaluation results were higher than 80%.

## Conclusions

In this paper, we propose a computer-aided diagnosis system with texture analysis for liver tumors in CT images. Combined with the improved PSO algorithm and SVM, the number of the best features is reduced to less than 20 by the proposed two-stage selection method. Meanwhile, fuzzy logic weight value is adopted to determine the final classification result. The five evaluation results were all above 80%. Therefore, the proposed method provides a liver tumor computer-aided diagnosis system with high differential diagnosis accuracy.

## References

[1] S. G. Mougiakakou, I. K. Valavanisa, A. Nikitab and K. S. Nikitaa, "Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers," *Artificial Intelligence in Med.*, vol. 41, no. 1, pp. 25-37, Jul., 2007.

[2] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", 3rd ed., Prentice Hall, 2007.

[3] J. Kennddy and R. Eberhart, "Particle Swarm Algorithm," in Proc, *IEEE Int. Conf. on Neural Networks*, pp. 1942-1948, Perth, 1995.

[4] L. Lam and S. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," *IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 27, no. 5, pp. 553-568, Sep., 1997.

[5] S. Haykin, "Neural Networks: A Comprehensive Foundation," Prentice Hall, Jul., 1998.

[6] M. G. Penedo, M. J. Carreira, A. Mosquera and D. Cabello, "Computer-aided diagnosis: a neural-network-based approach to lung nodule detection," *IEEE Trans. on Medical Imaging*, vol. 17, no. 6, pp. 872-880, Dec., 1998.

[7] B. E. Boser, I. M. Guyon and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. Annual Workshop on Computational Learning Theory*, pp. 144-152, New York, 1992.

[8] Y. Zeng, B. Wang, L. Zhao and Y. Yang, "The extended nearest neighbor classification," in *Proc. Chinese Control Conference, CCC*, pp. 559-563, China, 2008.

[9] M. J. Islam, Q. M. J. Wu, M. Ahmadi and M. A. Sid-Ahmed, "Investigating the Performance of Naive-Bayes Classifiers and K- Nearest Neighbor Classifiers," in *Proc. Int. Conf. on Convergence Information Technology, ICCIT,* pp. 1541-1546, Korea, 2007.

[10] C.-M. Wu and Y.-C. Chen, "Statistical feature matrix for texture analysis," in *Proc. CVGIP: Graphical Models and Image Processing,* vol. 54, no. 5, pp.407-419, Sep., 1992.

[11] R. M. Haralick, K. Shanmugam and I. H. Dinstein, "Textural Features for Image Classification," *IEEE Trans. on* Systems, *Man, and Cybernetics*, vol. 3, no. 6, pp.610-621, Nov., 1973.

[12] T. Xiaoou, "Texture information in run-length matrices," *IEEE Trans. on Image Processing,* vol. 7, no. 11, pp.1602-1609, Nov., 1998.

[13] C. Sun and W. G. Wee, "Neighboring Gray Level Dependence Matrix for Texture Classification," *Computer Vision, Graphics and Image Processing*, vol. 23, no. 3, pp.235-245, Sep., 1982.

[14] Z. Feng, F. Ju Fu and S. Qing Yun, "Texture feature based on local Fourier transform," in *Proc. Int. Conf. Image Processing,* Thessaloniki, pp. 610-613, Greece, 2001.



*ISBN 978-1-5386-4342-6*