1998 International Computer Symposium
Workshop on Image Processing and Character Recognition
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

Automatic Recognition and Identification of Abdominal Organs in CT Images

*Chien-Cheng Lee, Pau-Choo Chung and Hong-Ming Tsai**

Department of Electrical Engineering,
National Cheng Kung University, Tainan, Taiwan 70101, Republic of China
Email: pcchung@eembox.ncku.edu.tw

*Department of Medicine
National Cheng Kung University, Tainan, Taiwan 70101, Republic of China

ABSTRACT

Recognizing abdominal organs is one of essential steps in visualizing organ structure, for providing assistant in teaching, clinic training, and diagnosis. This paper describes a framework for automatic abdominal organ recognition from a series of CT image slices, designed based on shape analysis, image contextual constraint, and between-slice relationship. Two processing phases, feature extraction (object segmentation) and recognition, are included in this framework. In the phase of object segmentation, a multi-module contextual neural network is applied to segment each image slice into disconnected regions. For each region, its shape features, including relative location, relative distance, tissue, area, compactness, and elongatedness are calculated, along with its spatial relationships with respect to spine. Further, according to the knowledge of anatomy, these features are constructed to form fuzzy rules for organ recognition. Followed in the recognition phase, the obtained features and the overlapping information between adjacent slices are used for identifying each organ. This proposed framework has been tested on many clinic patient cases. Results indicate that this framework can successfully recognize abdominal organs, not being affected by partial volume effects.

1. INTRODUCTION

Computerized Tomography (CT) imaging is one of the most widely used radiographic techniques. Identifying organs from a series of CT images is very important which not only could provide doctors with patients anatomy, but is also one essential step of constructing 3D display of organs. Research in this area can be seen in [1]-[4]. However, most of them focused on identifying brain objects. As for the recognition of abdominal organs, the research is relatively unexplored.

This work was supported by National Science Council, Taiwan, ROC. Under Grant NSC87-2213-E006-023

This is because the structure of abdominal images are much complicated and its gray level distribution has higher variation. Therefore, many difficulties may be encountered during the recognition of abdominal organs from CT image series.

The first difficulty results from the partial volume effects, caused from spatial averaging, patient movement, beam hardening, and reconstruction artifacts. This causes the edges of the images blurred and contrast degraded. The second difficulty is due to the very similar gray values of different organs, resulting that the adjacent organs overlap. Furthermore, the same organ may exhibit different gray-level values in different slices or patients. Besides these dilemmas, the location and the shape of abdominal organs vary from top slices to bottom slices. Even for the same slice position, different patients may have slightly different organ shapes and relative positions. Because of the forging reasons, the edge-based technique, or a simple gray-level-dependent thresholding technique cannot accomplish our purpose.

In this paper, a contextual neural network combined with spatial fuzzy relations are proposed for the recognition of abdominal organs such as spine, spleen, kidneys and liver from a sequence of CT images. The contextual neural network is a multi-module one-layer neural network. Each neuron of this network estimates the possibility of a pixel belonging to a particular label. This estimation is based on the pixels gray level and the contextual information around this pixel. This contextual information is carried by the synaptic weights connecting this pixel and its surrounding neighborhood pixels. Because of the incorporation of contextual information, the effects caused from the interference of noises and the gray level variation of a same organ can be significantly reduced.

After the regions are extracted by using the contextual neural network, a number of descriptors were applied to describe these regions including their relative locations, relative distances, tissue intensities, area sizes, compactness, and elongatedness. These descriptors are described by fuzzy variables to express the imprecision of organ properties and applied to a fuzzy-rule-based recognition system for the identification of abdominal organs. The fuzzy-rule-based recognition system is constructed based on the knowledge of abdominal anatomy including each organ's shape descriptors and its spatial relations with other organs. With the adoption of fuzzy approach, the delimma in the recognition, resulted from the imprecision of each characteristic, can be resolved.

The method developed in this paper has been tested on more than forty sets of CT image series. Results indicate that we have effectively overcome the difficulties discussed above. Not only all the organs are successfully recognized, the dissected regions, such as the left and right liver lobes, are also correctly identified.

The remaining of this paper is organized as follows. The contextual neural network used for region segmentation is introduced in Section 2. Section 3 describes the applied shape descriptors associated with abdominal anatomy. The fuzzy-rule-based recognition system is also addressed in this section. Section 4 describes the organ recognition process. Our experimental results will be shown in Section 5. Finally, the conclusions are drawn in Section 6.

2. IMAGE SEGMENTATION USING CONTEXTUAL NEURAL NETWORK

The first essential step of organ recognition is the extraction of possible organ regions. In this section, a contextual neural network is proposed for this purpose [5], [6]. The contextual neural network uses the gray level and contextual information supported by neighboring pixels to label image pixels into disjointed regions. Fig. 1 shows the architecture of the contextual neural network, from which we can see that the network consists of several modules. Each module of the network estimates the possibilities that image pixels belong to the corresponding label associated with this module. In other words, module k calculates the probabilities of the image pixels belonging to label k.

Let $X=\{X_1, X_2, \ldots, X_n\}$ contain the probability vectors that the image pixels are assigned to each label, where $X_i=[x_{i1}, x_{i2}, \ldots, x_{im}]^T$, $1\le i\le n$, and each element x_{ik} in X_i represents the probability of pixel i assigned as label k. The n here is the image size and m is the number of labels. With this approach, these probability values satisfy the following condition:

$$\sum_{k=1}^{m} x_{ik} = 1 \quad for \ all \ pixel \ vector \ \mathbf{X}_i, \qquad and \quad 0 \le x_{ik} \le 1 \tag{1}$$

The connection weights between the neurons in module i and module j are set to a compatibility matrix $W=[W_1, W_2, \ldots, W_m]^T=[w_{ij}]$ in which each element w_{ij} $1\le i,j\le m$ indicates the compatibility measure, positive or negative, between label i and label j. If label i and label j have a high tendency of co-occurrence on two neighboring pixels, leading high support to each other, then w_{ij} is positive. If they were not to co-occur, then w_{ij} is negative. If their occurrences are independent of each other, then w_{ij} equals zero. The magnitude of w_{ij} is proportional to the strength of the constraints; in particular, w_{ij} would range in [-1, 1]. One typical situation is to set W as the correlation matrix. By doing such, W would be symmetric. Alternatively, if we assign the labels following the order of gray level values, the interconnection weights can be assigned, for simplicity, as follows:

$$w_{ij} = 1 - \frac{2|i-j|}{m-1} \qquad \forall \ \ i,j=1,\ldots,m \tag{2}$$

In this network, the ith neurons at modules k, $1\le k\le m$, receive the neighborhood supports from all modules, which are defined as follows:

$$\begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{im} \end{bmatrix} = \frac{1}{p} \begin{bmatrix} \sum_{\delta\in\Delta} \mathbf{W}_1^T \mathbf{X}_{i+\delta} \\ \sum_{\delta\in\Delta} \mathbf{W}_2^T \mathbf{X}_{i+\delta} \\ \vdots \\ \sum_{\delta\in\Delta} \mathbf{W}_m^T \mathbf{X}_{i+\delta} \end{bmatrix} \tag{3}$$

The v_{ik} here is the total input received by the ith neuron of module k, also denoting the total supports for label k at pixel i, and the degree of supports depends on the compatibility matrix W. $X_{i+\delta}$ contains the neighborhood probability values of the ith pixel, Δ defines the neighborhood around the particular pixel i, and p is the size of the neighborhood. W_k, $1\le k\le m$, here is the kth row vector of W.

Since each element in W would range in [-1, 1] and each element $x_{i+\delta,k}$ in $X_{i+\delta}$ is in the range [0, 1], all the values of v_{ik} in (3) would range in [-1, 1]. The $x_{i+\delta,k}$ is treated as the likelihood (probability) that label k appears in pixel i. If the likelihood is high and the corresponding compatibility indicates a high support, the v_{ik} would take a large and positive quantity. If the likelihood is high and the corresponding compatibility gives significant negative support, it would make a substantial large negative contribution to v_{ik}. Conversely, if the likelihood is very small, the quantity of the support would be negligible.

According to the above observations, the neighborhood support could be regarded as the amount of change to update the probabilities of label assignment. Thus, the network state of the ith neuron at the module k, $x_{ik}(t)$, is updated as follows.

$$x_{ik}(t+1) = \frac{Pos(x_{ik}(t) + v_{ik}(t))}{\sum_{\lambda=1}^{m} Pos(x_{i\lambda} + v_{i\lambda})} \tag{4}$$

where $Pos(X)= X$ if $X>0$ and 0, otherwise.

According to Equation (4), we can see that the neuron i at module k receives the contribution v_{ik} from neighborhoods, and adjusts itself to produce a new output $x_{ik}(t+1)$. The $Pos(\bullet)$ is a nonnegative function, and the denominator is a normalization term to guarantee the summation of all elements in X_i to equal 1. The network adjustment is conducted iteratively so as to reduce local ambiguity and achieve a global convergence.

Before the network starts learning, the initial labeling probabilities of each pixel have to be determined in advance of the network progress. The Kohonen self-organizing algorithm is adopted to analyze the gray level distribution of the input image and define the initial probabilities to each pixel vector. The algorithm comprises the learning phase and the retrieving phase. Let A=$\{I_1, I_2, \ldots, I_n\}$ contain the gray values of the input image, in which the I_i is the gray level of pixel i, and n is the image size. Let J_k, $1 \le k \le m$, be the weight associated with the kth neuron of the Kohonen neural network. In the learning phase, upon receiving a gray level, the network locates a winning neuron based on the following distance function

$$\|I_i - J_c\| = \min_{1 \le k \le m} \{\|I_i - J_k\|\} \tag{5}$$

where the index c refers to the winning neuron. Then, the weights around the winning neuron are updated as follows:

$$J_k(t+1) = \begin{cases} J_k(t) + \alpha(t)(I_i - J_k(t)), k \in N_c(t) \\ J_k(t), \qquad otherwise \end{cases} \tag{6}$$

where $N_c(t)$ is the neighborhood function around the winning neuron c at time t, and $\alpha(t)$ is the corresponding value of the learning-rate parameter. Both $N_c(t)$ and $\alpha(t)$ vary dynamically during learning. The behavior of the update rule (6) is to move the weight vector of the winning neuron c toward the input I_i. Once the training has progressed sufficiently, the weight vector will tend to follow the distribution of the input data (the image gray level).

In the retrieving phase, when we apply the gray level of pixel i, a winning neuron c will be determined by (5). Then, the initial labeling probabilities in X_i are assigned using the Gaussian distribution N(c, σ) with its mean located at c and standard deviation equal to σ. Thus, the initial values of the labeling assignments for pixel i are defined as follows:

$$\mathbf{X}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{im} \end{bmatrix} = \frac{1}{\sum_{k=1}^{m} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(k-c)^2}{2\sigma^2}}} \begin{bmatrix} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(1-c)^2}{2\sigma^2}} \\ \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(2-c)^2}{2\sigma^2}} \\ \vdots \\ \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(m-c)^2}{2\sigma^2}} \end{bmatrix} \tag{7}$$

Here, the standard deviation σ is a spread factor to control the degree of dispersion of initial probabilities. Based on (7), we can see that a large probability value will be assigned to the element x_{ic} in pixel vector X_i. Conversely, smaller values will be assigned to the remaining elements x_{ik}, where $k \ne c$, following the probability density function of Gaussian distribution N(c, σ) and the distance from k to c. The denominator is a normalization term to guarantee the assigned probabilities satisfy the constraint defined in (1).

After setup the initial probabilities and network interconnection weights, the network proceeds in parallel and repeatedly, until the following stopping criteria is met.

$$\sum_{i=1}^{n} \|X_i(t+1) - X_i(t)\| < \varepsilon \quad \text{, where } \varepsilon \text{ is a small value.} \tag{8}$$

3. Construction of Fuzzy Rules for Recognition

The fuzzy rules for organ identification is defined by incorporating the knowledge of abdominal anatomy into fuzzy sets and fuzzy rules. After the regions extracted by using Contextual Neural Network (CNN), a number of descriptions, including relative location, relative distance, tissue, area, compactness, and elongatedness, were applied to describe these regions.

(1) *Location*: The location property is described by spatial relations between two regions denoted as A R B where R denotes spatial relations containing *right_of*, *left_of*, *below*, and *above*; A is the argument region, and B is the referent region. These spatial relations are defined by fuzzy sets [7].

(2) *Distance*: The distance between two regions is computed as the Euclidean distance between the centers of the two regions divided by image width for normalization. To describe the distance in terms of linguistic, two fuzzy sets labeled as *near* and *far* are introduced.

(3) *Tissue*: According to our observations of abdominal CT images, the image pixels can be partitioned into: 1) Transparent tissues, e.g., air, fluid, and fat. 2) Soft tissues, e.g., organs and muscles. 3) Sclerous tissues, e.g., bone, calcification, and metal. The gray level distributions of these three types of tissues are distinctly different. For this reason, three fuzzy sets labeled as *transparent*, *soft*, and *sclerous* are used to state the tissue properties of regions.

(4) *Area*: This is defined as the area surrounded by the outer contour of a region normalized by the image size.

(5) *Compactness*: The measurement of compactness is defined as [8]:

$$Compactness = \frac{(region\ \ border\ \ length)^2}{number\ \ of\ \ pixels\ \ in\ \ region} \tag{9}$$

(6) *Elongatedness*: Elongatedness is defined as the ratio between the length and the width of a region's bounding rectangle [8].

To create fuzzy rules for identifying the organ of interest, we have summarized the feature characteristics associated with abdominal organs according to the knowledge of anatomy and the experts' opinions.

1) Spine: Spine is located at the bottom center of the image. Additionally, tissue of spine belongs to bone which has the highest gray level intensity.
2) Kidneys: The two kidneys are located right next to spine; one is on the left and the other is on the right. The size of kidney is not too large, and the shape is almost a circle.
3) Spleen: The location of spleen is located either right beside the left kidney, or the spine if the left kidney does not exist in this image slice. The size of spleen is not too large, its shape is compact, and its position is located somewhat far from spine.
4) Liver: Liver is located at the upper right of the right kidney, or the spine if the right kidney does not exist in this image slice. The size of liver is large, its shape is somewhat compact, and its position is located far from spine. Additionally, in some slices the liver may dissect into two regions.

These results will be translated into fuzzy rules to identify the organs.

4. PROCESS FOR ORGAN RECOGNITION

In this section, a recognition process will be proposed to recognize the organs of interest based on the fuzzy rules described above. The recognition process recognizes the organs of interest in the order of 1) kidneys, 2) spleen, and 3) liver. The process goes through every slice using spatial relationship between regions and the overlapping information between adjacent slices. In order to define the spatial relationship, a landmark needs to be first identified.

A. Landmark Definition

The landmark is used as a reference point to describe the spatial relationship between different regions in every slice. Thus, it must exist in all slices, has stable location from slice to slice, and easily be detected. According to the knowledge of anatomy, spine satisfies these conditions and, therefore, is used as the landmark to define the relative spatial relationships of other organs. Thus, the position of spine has to be first located. The locating of spine is achieved by searching for the region of maximum area satisfying the rules associated with the description of spine. Once the landmark was determined, it could help to locate other organs.

B. Region Overlapping Information

Since CT images are usually closely consecutive from the top slice to the bottom slice, the properties of regions of the same organ between adjacent slices are quite similar. Therefore, the region overlapping information between adjacent slices provides a basis for searching the possible region of the organ of interest. In order to exploit the region overlapping information, both Current Overlap Ratio (COR) and Next Overlap Ratio (NOR) are defined as follows.

$$Current\ \ Overlap\ \ Ratio = \frac{area\ \ of\ \ R3}{area\ \ of\ \ R1}, \tag{10}$$

$$Next\ \ Overlap\ \ Ratio = \frac{area\ \ of\ \ R3}{area\ \ of\ \ R2}, \tag{11}$$

where R1 is the already-recognized region in the current slice, R2 is the candidate region in the next slice, and R3 is the intersection of R1 and R2.

Both COR and NOR are considered as fuzzy variables possessing linguistic terms, low and high. From the above definitions and experimental observations, the following characteristics can be obtained.

(1) If both ratios are high, regions R1 and R2 belong to the same object.
(2) If the COR is high, but NOR is low, regions R1 and R2 belong to the same object, but there are some abnormal extrusion in R2.
(3) If the COR is low, but NOR is high, regions R1 and R2 belong to two different organs. Exceptionally, in liver recognition, the region R2 will be regarded as a dissected liver lobe.
(4) If both ratios are low, regions R1 and R2 are regarded as belonging to two different organs.

C. Recognition Process

The overall recognition process is conducted as follows. First, the landmark (spine) is identified for each slice. Then the spatial relationship of every region versus spine is captured. Based on the spatial relationship, the slice containing the region with maximum likelihood of organ of interest is identified and the position of organ of interest in this slice is located. After that, from the starting slice, both up-searching and bottom-searching are performed to search for the region of the same organ in other slices, based on the region overlapping information, as shown in Fig. 2. The recognition process recognizes the organs of interest in the order of 1) kidneys, 2) spleen, and 3) liver. Furthermore, it is worth to mention that in searching the latter organs, the spatial relationship with reference to the previous already-identified organs can be suitably incorporated as part of the spatial constraints in the recognition.

D. Contour Modification

Due to very similar gray levels of adjacent organs, it is possible to see the abnormal extrusions in some of obtained organ's contour. The occurrence of abnormal extrusions can be easily detected, as mentioned previously, based on the overlapping information between adjacent slices. For obtaining more accurate organ contours, the abnormal extrusions will be removed by referring to the contour of the already-recognized region. Let R1 be the already-recognized region in the current slice, and R2 be the region which is to-be-processed in the next slice. The abnormal extrusions A_i, i=1..N, in R2 are the connected components belonging to R2-(R1 $\cap$R2). These extrusions will be removed in the order of size, from large to small, until the percentage of extrusions is within a range. Assume that $A_1 > A_2 > \ldots > A_N$. The step for removing the abnormal extrusions is performed as follows.

do

$$R2 = R2 - A_i$$

$$i = i + 1$$

while $[R2-(R1 \cap R2)]/R1 > \varepsilon$

The ε here represents the difference tolerance between two adjacent slices. After the modification, a contour smoothing algorithm is applied to smooth the obtained region boundary.

6. EXPERIMENTAL RESULTS

The system was tested by several cases obtained form hospital of National Cheng Kung University. Fig. 3. shows the result of one typical case, from which we can see that all organs of interest such as spine, kidneys, spleen, and liver are correctly recognized. However, because of the effect of partial volume and patient motion causing the edges of organ regions blurred, the obtained organ regions are slightly smaller than they look like. Even so, these results are still satisfactory for clinician, especially if we consider the near boundary area is actually formed from the spatial average of the organ and other abdominal areas, such as air or other organs.

Fig. 4. shows the results of another case, in which the liver dissected into two disjointed regions. By applying region overlapping information between adjacent slices, these two regions are both identified as the same organ, i.e., liver.

7. CONCLUSIONS

In this paper a system for the recognition of abdominal organs is established based on a contextual neural network, organ shape descriptors, anatomic knowledge, and fuzzy rules. The contextual neural network is used to label the image pixels into disjointed regions, followed by a fuzzy-rule-based recognition system to identify organs from these regions. In the neural network, contextual information of each pixel is embedded inside its state evolution so as to achieve a consistency of labeling. Therefore, the difficulties resulted from the partial volume effect and similar gray values of different organs are resolved. The fuzzy rules are constructed according to shape descriptions and anatomic knowledge of abdominal organs. In our process, a starting slice containing the region with maximum likelihood of organ of interest is first located by the fuzzy rules. After that, from the starting slice, both up-searching and bottom-searching are performed to search for the region of the same organ in other slices. In the searching stage, the overlapping information between adjacent slices is also exploited to help the identification of organ locations. Furthermore, because of the proper usage of the overlapping information, separated regions of a same organ, such as the left and right lobes of the liver, can be successfully recognized.

REFERENCES

[1] H. Li *et al.*, "Object recognition in brain CT-scans: Knowledge-based fusion of data from multiple feature extractors," *IEEE Trans. Med. Imag.*, vol. 14, no. 2, pp. 212-229, 1995.

[2] C. W. Chang *et al.*, "A fuzzy rule-based system for labeling the structures in 3D human brain magnetic resonance images," in *Proc. of IEEE Int. Fuzzy Systems.*, pp. 1978-1982, 1996.

[3] S. P. Raya, "Low-level segmentation of 3D magnetic resonance brain images: A rule-based system," *IEEE Trans. Med. Imag.*, vol. 9, no. 3, pp. 327-337, 1990.

[4] M. Kobashi and L. G. Shapiro, "Knowledge-based organ identification from CT images," *Pattern Recognition*, vol. 28, No.4, pp. 475-491, 1995.

[5] A. Rosenfeld, R.A. Hummel and S.W. Zucker, "Scene Labeling by Relaxation Operations," *IEEE Trans. Syst. Man Cybern.*, vol. 6, pp.420-433, 1976.

[6] Wei-Chung Lin, Eric Chen-Kuo Tsao and Chin-Tu Chen, "Constraint Satisfaction Neural Networks for Image Segmentation," *Pattern Recognition*, Vol. 25, No. 7, pp.679-693, 1992.

[7] Raghu Krishnapuram, James M. Keller and Yibing Ma, "Quantitative Analysis of Properties and Spatial Relations of Fuzzy Image Regions," *IEEE Trans. Fuzzy Systems*. Vol. 1, No. 3, pp.222-233, 1993.

[8] Milan Sonka, Vaclav Hlavac and Roger Boyle, *Image Processing, Analysis and Machine Vision*, Chapman & Hall Publishing Company, 1993.

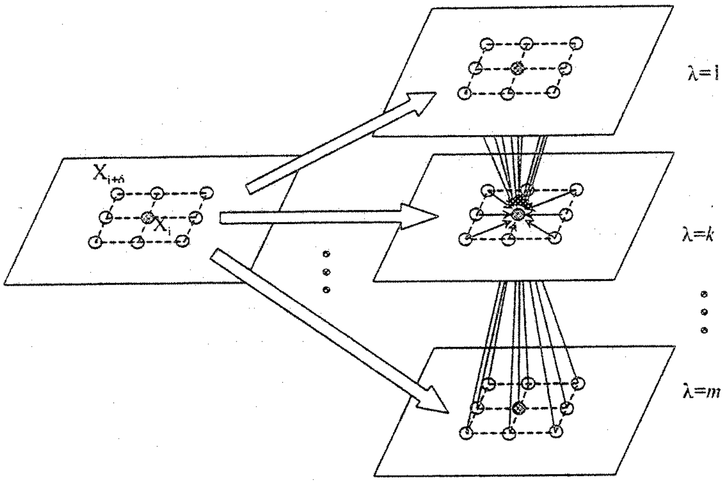

Fig. 1. The architecture of the contextual neural network.

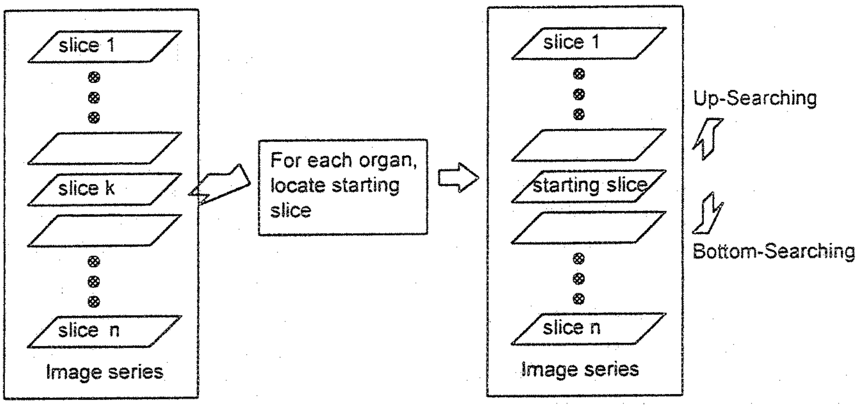

Fig. 2. The organ recognition process diagram.

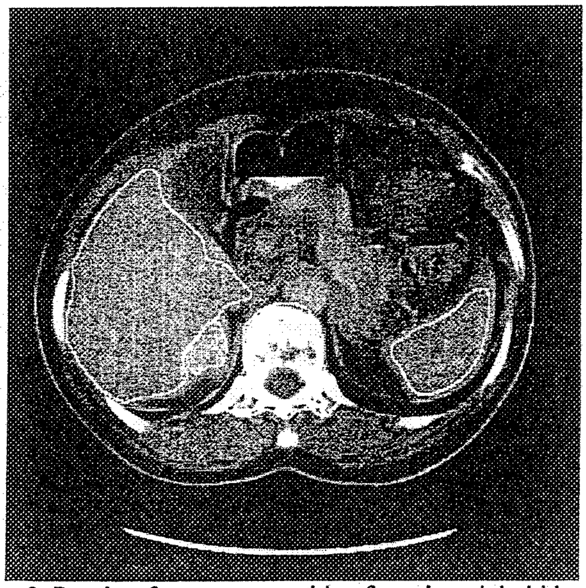

Fig. 3. Results of organs recognition for spine, right kidney, spleen, and liver.

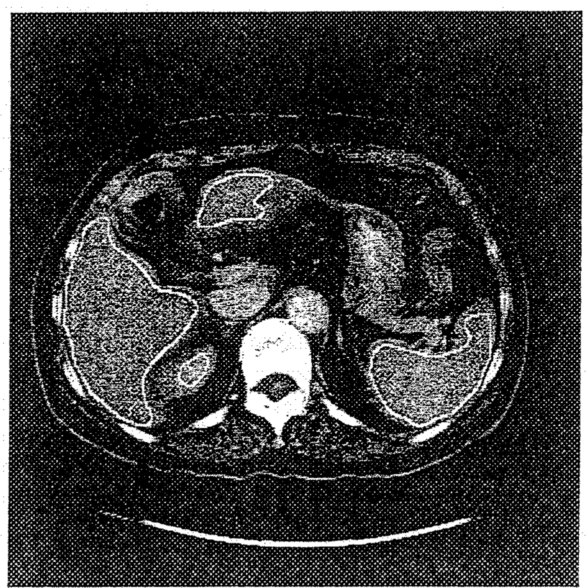

Fig. 4. Results of organs recognition for spine, right kidney, spleen, and liver which appears in two regions.