ELSEVIER

Computers in Biology and Medicine 38 (2008) 765–784

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation

M. Alper Selver[a,*], Aykut Kocaoğlu[a], Güleser K. Demir[a], Hatice Doğan[a], Oğuz Dicle[b], Cüneyt Güzeliş[a]

[a]*Electrical and Electronics Engineering Department, Dokuz Eylül University, Buca, Izmir 35160, Turkey*
[b]*Faculty of Medicine, Department of Radiology, Dokuz Eylül University, Narlıdere, Izmir, Turkey*

Received 22 February 2007; accepted 8 April 2008

Abstract

Identifying liver region from abdominal computed tomography–angiography (CTA) data sets is one of the essential steps in evaluation of transplantation donors prior to the hepatic surgery. However, due to gray level similarity of adjacent organs, injection of contrast media and partial volume effects; robust segmentation of the liver is a very difficult task. Moreover, high variations in liver margins, different image characteristics with different CT scanners and atypical liver shapes make the segmentation process even harder. In this paper, we propose a three stage (i.e. pre-processing, classification, post-processing); automatic liver segmentation algorithm that adapts its parameters according to each patient by learning the data set characteristics in parallel to segmentation process to address all the challenging aspects mentioned above. The efficiency in terms of the time requirement and the overall segmentation performance is achieved by introducing a novel modular classification system consisting of a K-Means based simple classification system and an MLP based complex one which are combined with a data-dependent and automated switching mechanism that decides to apply one of them. Proposed approach also makes the design of the overall classification system fully unsupervised that depends on the given CTA series only without requiring any given training set of CTA series. The segmentation results are evaluated by using area error rate and volume calculations and the success rate is calculated as 94.91% over a data set of diverse CTA series of 20 patients according to the evaluation of the expert radiologist. The results show that, the proposed algorithm gives better results especially for atypical liver shapes and low contrast studies where several algorithms fail.

Keywords: Liver segmentation; Computed tomography–angiography; Multi-layer perceptron network; K-Means

1. Introduction

Living donated liver transplantation is a procedure where a healthy voluntary donor gives a part of his or her liver to another person. Measurement of the liver volume and analysis of the liver vasculature are important stages to decide whether a candidate for transplantation is suitable or not. Generally, liver volume information is used to avoid size incompatibility between donor and patient, and vasculature analysis in three dimension (3D) is used for pre-evaluation of surgery. Thus, the success of liver transplantation depends on the sufficiency of liver volume and its supplying vessels and accurate knowledge of the hepatic and portal vascular anatomy of donors for living-related transplantation would reduce the incidence of vascular complications during and after transplantation.

Routine preoperative evaluation of donors requires both computed tomography (CT) [1] and CT with contrast medium injection, namely CT–angiography (CTA), which are currently the most widely used radiographic techniques for the rendering of liver parenchyma, vessels and tumors in living liver transplantation donors. Instead of conventional angiography, CTA offers several advantages: it is minimally invasive and has diminished patient morbidity, cost, and radiation exposure to patients and staff. Moreover, CTA provides detailed information on vasculature due to the injection of contrast media. Before 3D rendering [2] of the vasculature and the measurement of liver volume, accurate segmentation of the liver from surrounding tissues and

* Corresponding author. Tel.: +90 232 4127176; fax: +90 232 4531085.
E-mail address: alper.selver@deu.edu.tr (M.A. Selver).

0010-4825/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2008.04.006

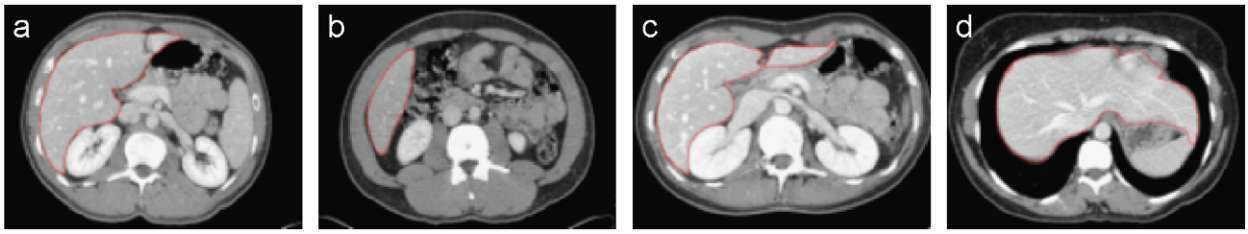

Fig. 1. Examples of abdominal CTA images: (a) very low contrast between liver and vena cava; (b) low contrast between liver and muscle tissue; (c) high contrast between liver and adjacent tissues (i.e. right kidney, muscle, vena cava); and (d) atypical liver shape that causes unclear boundary with spleen.

organs is necessary. Since the number of image slices used for 3D rendering is very large, manual segmentation of the liver on each slice is time consuming and tedious. Also the results highly depend on the skill of the operator. Therefore an automatic segmentation procedure to segment the liver in all slices is needed.

Besides its several advantages over manual segmentation, automatic segmentation of the liver is very challenging. These challenges arise from the following difficulties: First of all, the gray level values of adjacent organs of the liver are similar to each other (Fig. 1a and 1b). This similarity reduces the performance of thresholding techniques dramatically. Secondly, due to the injection of contrast media and/or different modality settings, the liver (and all other tissues) may have different gray-level values for different patient data sets, or even in different slices of the same data set (Fig. 1a–c). These effects prevent the usage of the gray level dependent segmentation techniques. Finally, the anatomical structure of the liver in different image slices is different and its shape can vary significantly from patient to patient (Fig. 1a and b). Even two or three separate regions can be seen in the same slice (Fig. 1c). Moreover, it is reported in [9] that around 15% of the patients have atypical liver shapes (i.e. unusual size or orientation of the liver, liver shape after segmentectomy) (Fig. 1d). Thus, traditional shape based segmentation techniques are not enough to segment the liver efficiently.

Our strategy for overcoming these difficulties involves a segmentation method which does not utilize a common parameter set found from all patient data sets. Instead, the method is capable of adapting the parameter set to each patient. The main reason for this approach is that the ranges of the parameter values differ significantly from patient to patient, and these wide ranges decrease the efficiency of the method when one utilizes a common parameter set for all patients.

Thus, we propose a method which examines and adapts its parameters according to each patient. We call this approach as *patient-oriented* segmentation. For qualifying 'patient oriented', the algorithm learns data set characteristics in parallel to segmentation process, and adapts its parameters to these characteristics. To our knowledge, there is no method in the literature that works in this manner and at the same time addresses all the challenging aspects mentioned above.

In the literature, different automatic and semi-automatic methods have been developed and performed for the segmentation of the liver from CTA series. These methods include but not limited to morphological techniques [3–6], deformable models [7–12], and neural networks [13–15]. However, neither in semi-automatic, nor in automatic algorithms, the problems of atypical liver shapes, different modality characteristics and data sets with low contrast adjacent tissues is handled together. Therefore, they do not deal with the all variations in CTA images at the same time.

Morphological techniques combined with gray level thresholding are used in [3] while in [4,5] these are combined with a parametrically deformable contour model which is used for boundary refinement. Although the method proposed in [4] is reported to be successful in most of the cases, a mean gray level value assumption is made for the liver at the intermediate levels of the algorithm. This assumption limits its use when the liver is more attenuating (brighter) due to the contrast media. Deformable contour models are also used by [7,8]. However the need for setting some seed points and parameters such as maximum gradient or time threshold makes it hard to use for radiologists. In [9], an automatic algorithm is proposed using deformable models; however this method does not provide correct results for atypical liver shapes. Another automatic technique is proposed in [11], in which a 3D active shape model is built from 32 samples using an optimization approach based on the minimum description length. The combination of deformable models and statistical priors [12] seems to be effective for fully automatic techniques where initial parameters for the statistical shape model (SSM) are determined with an evolutionary algorithm and a modified active shape method is used to refine the detected parameters. As in [11], the method of [12] also requires the training of the SSM with a data set to model the expected shape and appearance of the liver so resulting in a dependency on the set of CTA series used in the training.

Artificial neural networks are used for gray level classification in [13] and for feature based recognition in [14] which are discussed in detail in this study. The technique proposed in [13] is semi-automatic and require more than one manually segmented image as training data prior to the automated process. The method in [14] is not patient oriented and training is done with a limited set of images. Due to the high variation of image characteristics, a larger and more diverse database is recommended to generalize this system for reliable performance. A contextual neural network with a high segmentation performance is proposed in [15], but the results show that it fails where the gray level of the desired region is too close to the adjacent tissues. In [16], texture of the abdominal organs

is used for segmentation. Although this approach is successful in general for abdominal organs, it fails in the segmentation of liver and spleen, especially in atypical liver case, since their texture is similar in CTA data sets. Recently, Seo et al. [17] proposed a fully automatic algorithm by determining the spine first and then by using it as a reference point for segmenting the liver using morphological operators, multimodal thresholding and a decision rule. However, this approach is tested with a very limited set of CT series.

This paper proposes a robust and efficient method that can automatically segment the liver of transplantation donor candidates in any CTA series. The success rate is calculated as 94.91% over a data set of diverse CTA series of 20 patients according to the evaluation of the expert radiologist experienced on pre-evaluation of transplantation donors for more than 100 cases. The robustness of the method follows from its capability of dealing with the contrast variations and atypical liver shapes. These capabilities are provided by the patient oriented structure which learns the characteristics of a patient data set for each slice in parallel to the segmentation process and adapts its parameters according to these characteristics. Our iterative segmentation algorithm uses classification of pixels (using an unsupervised clustering method i.e. K-Means) together with adjacent slice information. A more complex classifier (multilayer perceptron network—MLP) is developed for the data sets where the K-Means clustering gives insufficient results. The efficiency in terms of the time requirement and the overall segmentation performance is achieved by introducing a novel modular classification system consisting of a simple classification system (i.e. K-Means based) and a complex one (i.e. MLP based) which are combined with a data-dependent and automated switching mechanism that decides to apply one of them. The switching is based on the detection of "low contrast" data set or atypical liver shape. If none of these is detected then K-Means based classification system is applied on a single feature (i.e. the gray level value of each pixel), otherwise MLP based classification system is utilized with three features (i.e. mean, standard deviation and distance transform). The developed method gives sufficient performance for different modalities, varying contrast, dissected liver regions and atypical liver shapes. Results indicate that we have effectively overcome the challenging difficulties explained before. This performance is achieved with the proposed modular classification system as well as introducing the distance transform as a feature for each slice and then using this information in the succeeding slice to reveal three dimensional properties of the liver which cannot be obtained by the set of slices processed individually. In other words, the approach in the paper provides the ability of dealing with the contrast variations and atypical liver shapes first by recognizing the existence of these problems, by choosing appropriate classification method, and then by solving the segmentation problem using inter-slice information provided by the distance transform. To our knowledge, there is no method in the literature which handles all these difficulties at the same time.

The rest of the paper is organized as follows. The properties of the patient data sets are presented in Section 2. The first step of the three step segmentation system, preprocessing, covers the removal of adjacent tissues to the liver (i.e. fat tissue, right kidney, spine and ribs) is explained in Section 3. The classification of the liver with modular classification system by using either K-Means or a neural network structure depending on the data set properties (i.e. contrast, atypical liver shape) and the features used for classification are established in Section 4 as the second step of the segmentation system. The last step, post-processing that consists of the removal of the misclassified objects and identification of disjointed parts of the liver is explained in Section 5. The evaluation of the system is given in Section 6. Finally, future plans for the improvement of the system are discussed in Section 7.

2. Patient data sets

Our data sets were acquired after contrast agent injection at portal phase using a Philips Secura CT with two detectors and a Philips Mx8000 CTA with 4 detectors, both equipped with the spiral CTA option and located in Dokuz Eylül University Radiology Department. Spiral CTA acquires data continuously, in a spiral path, as the patient is transported at a constant speed through the gantry. This technique scans the entire liver in 15–30 s and offers several advantages for both liver tumor detection and 3D visualization. Its speed also reduces or eliminates respiratory misregistration between slices. Now CT scanners with 64 or more detectors are available and it is possible to have higher quality images with single-breath-hold volumetric data acquisition. They also permits very thin-section coverage of large anatomic areas at speeds 3–7 times faster than the previously used helical CT scanners.

Twenty data sets (CTA series), which were obtained by these scanners, consist of 12 bit DICOM images with a resolution of 512×512. The data sets were chosen randomly from the Picture Archiving and Communication System (PACS). All of the 20 CTA series have 3–3.2 mm slice thickness and this corresponds to a slice number around 90 (minimum 77, maximum 105 slices).

Our segmentation system is designed to work with 8 bit images to support all image types. Therefore 12 bit DICOM images are reduced to 8 bit using window center and window width information, which are stored in Meta information header of the original DICOM images. Although a simple proportional scaling ($[0, 4095] \rightarrow [0, 255]$) would be the most obvious way of such a conversion, windowing is used to reduce the undersampling effect. By using windowing, the full contrast of the output display range is expanded over actually useful part of the input density range [18].

Twenty patient data sets are divided into two groups based on their volumetric histograms, namely 'high contrast' and 'low contrast'. It is observed that some data sets have three lobes in their volumetric histograms (Fig. 2a) where these lobes correspond to the fat tissue, darker soft tissues (i.e. muscles, stomach, intestines) and brighter soft tissues (i.e. heart, kidney, spleen) from left to right, respectively. The liver belongs to both second and third lobes with varying ratios due to contrast media it absorbs and modality settings. These data sets are called 'high

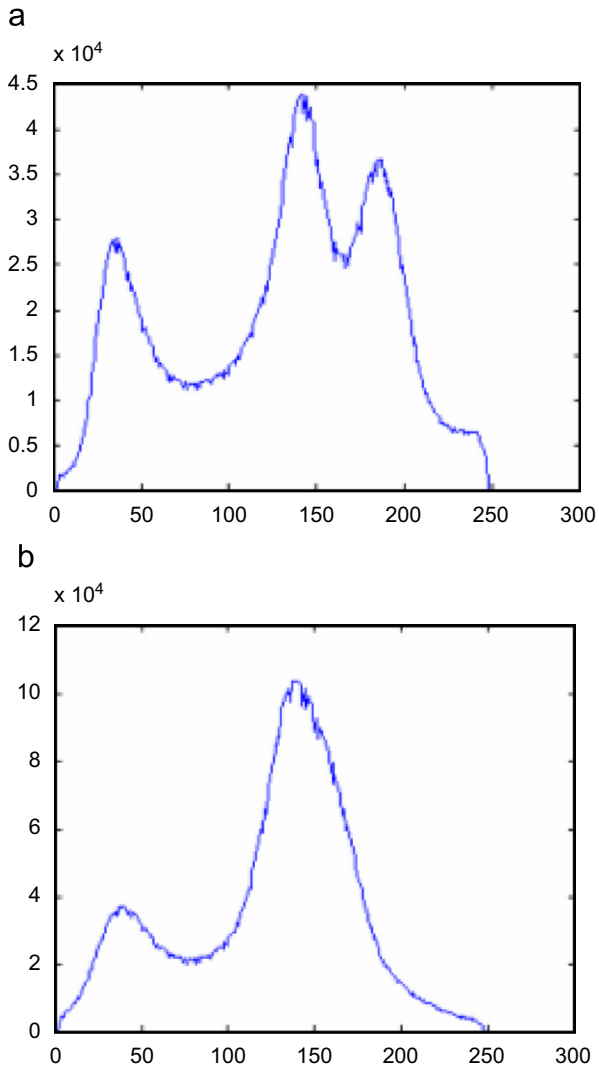

Fig. 2. Two examples of volumetric histograms: (a) volumetric histogram of a 'high contrast' CTA series one image of which is shown in Fig. 1c (three lobes); and (b) volumetric histogram of a 'low contrast' CTA series one image of which is shown in Fig. 1b (two lobes) (for illustration purposes the values of 0 and 255 are not drawn in the histograms).

contrast' because the gray level value of the liver is different than the adjacent tissues and organs (Fig. 1c). In 'low contrast' data sets, dark and bright soft tissues form only one lobe, thus the volumetric histograms have two lobes in total (Fig. 2b). In these data sets, it is harder to segment the liver because of the gray level similarity with adjacent organs (Fig. 1a). In 20 patient data sets, it is found that 15 data sets belong to 'high contrast' group while five data sets belong to 'low contrast group. The developed algorithm first determines the group of the data set and applies different classification and post-processing methods based on this decision.

3. Segmentation of the liver

The developed segmentation algorithm is designed to have three stages. The first stage is preprocessing which consists of the removal of the irrelevant tissues (the fat tissue, the spine, the right kidney and the ribs) from the original images and finding the smallest possible region of interest (ROI), where the liver tissue is known to exist. The second step of the segmentation procedure is the segmentation of the liver. This step consists of two parts: (1) Automatic selection and segmentation of an 'initial image'. (2) The segmentation of the remaining slices one by one starting from the 'initial image'. The third step, post-processing, includes necessary operations to remove small mis-segmented objects and to smooth boundaries. Moreover, identification of all components of the liver when the liver dissects into two or more regions is also done at this post-processing stage.

Before starting the process, the user selects a slice which is called 'initial kidney image'. The 'initial kidney image' is the slice where the liver and the right kidney exist together for the last time in the data set. Starting from this image, the algorithm for removing the right kidney runs through the data set until the right kidney disappears in all slices.

In the proposed algorithm, the default selection of 'initial kidney image' is the last slice of the CTA series. Although this assumption is mostly true, an interface with the user is also provided for the cases in which the right kidney does not exist at the last slice. This selection can also be done with no need to user interface in an automatically yet complicated fashion. For simplicity, we prefer to use this particular one-touch user interface, that is, selection of 'initial kidney image'. Needless to say, this is a rather simple task for the user.

3.1. Pre-processing

After the selection of the 'initial kidney image', the preprocessing starts by removing irrelevant tissues and organs including the fat tissue, the spine, the ribs and the right kidney. It is worth to point that all steps of the preprocessing stage are applied to the original images and the result of each step is removed from the rest of the images at the end of the stage.

3.1.1. Removing the fat tissue

To remove the fat tissue from a patient data set that consists of several CTA images (Fig. 3a), an adaptive thresholding method has been applied. In the volumetric histogram of a CTA series, the first lobe of the histogram corresponds to the fat tissue if there is enough fat tissue in the patient. To locate this lobe global minimum/minima of the histogram have to be found. For this reason, an averaging filter is applied to eliminate high frequency components of the histogram. Then the gradient of the smoothed histogram is calculated and global minimum/minima are found where the gradient histogram changes from negative to positive. Finally, a proper threshold value is found as the gray level value of the first global minimum. The tissues, which are removed with the application of the determined threshold value, are shown in Fig. 3b.

In young or/and fit patients, the fat tissue might be so less that it does not correspond to a lobe in the volumetric histogram. In other words, the first lobes in Fig. 2a and b might be

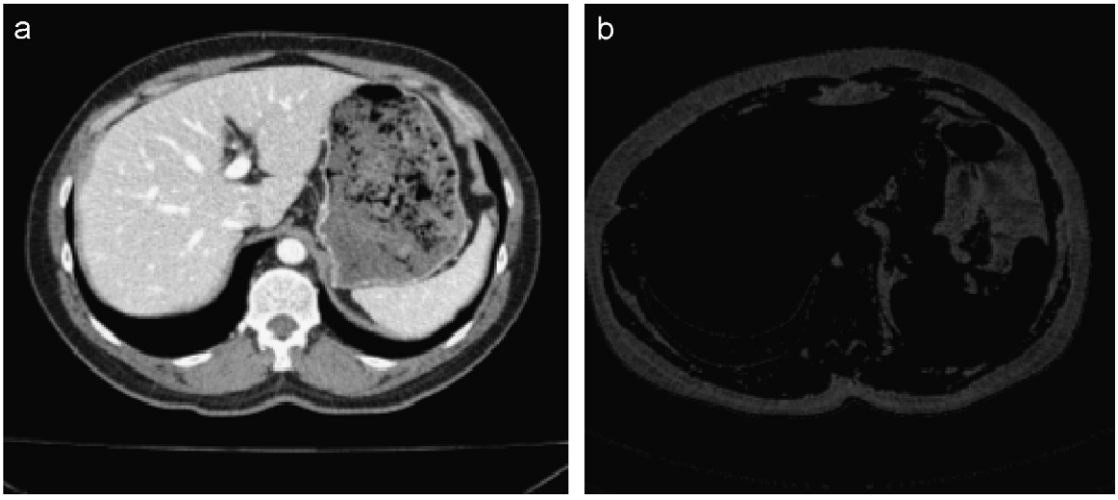

Fig. 3. Automatic determination of the threshold to remove fat tissue: (a) original abdominal CTA image; and (b) thresholded fat tissue.

suppressed due to this affect. To prevent wrong determination of the threshold value in these data sets, the maximum gray level value for the fat tissue threshold is limited to 100.

3.1.2. Removing the spine and the ribs

In the literature, the common method to remove the spine and the ribs is thresholding because of their high Hounsfield values. However, this method can also remove some vessels inside the liver or some other enhanced tissues (i.e. the kidneys) due to the brightening affect of the contrast media. These unexpected removals might cause the segmentation process to fail or the loss of important information from the liver. Therefore, thresholding is combined with anatomy knowledge to remove the spine and the ribs in the developed method.

First, an empirically determined threshold value ($T = 245$) has been applied to the image (Fig. 4a) to remove fat, skin and muscle tissues. The resulting image consists of the spine, the ribs and contrast enhanced tissues (i.e. liver vessels, kidneys) (Fig. 4b). Then, the spine and the ribs must be removed from the thresholded images without loosing any information of contrast enhanced tissues. Since it is known that the spine and the ribs surround the contrast enhanced tissues, a topology based method is used. First, the row and column projections are calculated from the thresholded image. The column projection is used to find the columns where the ribs start (from left and right) by using the first and last non-zero values on it. The row projection is used to determine the row where the spine and ribs start (from the bottom) by using the first non-zero value on it. The middle point of the corresponding row and the values determined as the starting point of the ribs are then used to make a frame through the spine direction (Fig. 4b). By dilating this frame, the ribs and the spine are excluded (Fig. 4c). To segment the spine and the ribs by combining these two images, binary morphological image reconstruction (BIMIR) [19] is used.

BIMIR is based on two images, a marker and a mask. Processing is based on the concept of the connectivity of these images. BIMIR processes the marker image, based on the characteristics of the mask image. The high gray level values in the marker image specify where the processing begins. The processing continues until the gray level values stop changing. If g is the mask and f is the marker, the reconstruction of g from f is defined by the following iterative procedure:

1. initialization of h_1 to be the marker image f,
2. creation of the structuring element, S,
3. repeat: $h_{k+1} = (h_k \Theta S) \cap g$ until $h_{k+1} = h_k$,

where the dilation operation, Θ, is defined as

$$h_k \Theta S = \{z | (S)_z \cap A \neq \emptyset\}$$

Conceptually, BIMIR can be thought as repeated dilations of the marker image until the contour of the marker image fits under the mask image. In this way, the peaks in the marker image "spread out", or dilate.

Intersecting framed image and the thresholded image by using 'AND' operation, the marker image is generated. Using the thresholded image as the mask in the BIMIR, the spine and the ribs are obtained (Fig. 4d). The advantage of using BIMIR is that it is possible to reconstruct of a rib correctly even if some parts of it remain outside the frame as in Fig. 4c.

3.1.3. Removing the right kidney

The right kidney is an adjacent organ to the liver which can have different gray level values due to the contrast media (Fig. 1a–c). Moreover, its shape and location differs from patient to patient. Also the shape of the kidney varies through the slices of a CTA series (Fig. 1a and b). Because of the partial volume affects, sometimes the border between the right kidney and the liver almost vanishes (Fig. 5a). Therefore, the elimination of the kidney should be done prior to the segmentation of the liver to increase segmentation performance.

After removing the ribs and the spine, the detection of the right kidney starts from the 'initial kidney image', which is the slice where the liver and the right kidney exist together for the last time in a data set (Fig. 5b). This user selected slice is mostly the last slice for a CTA series, however, sometimes the kidney ends before the liver. Kidney detection process starts

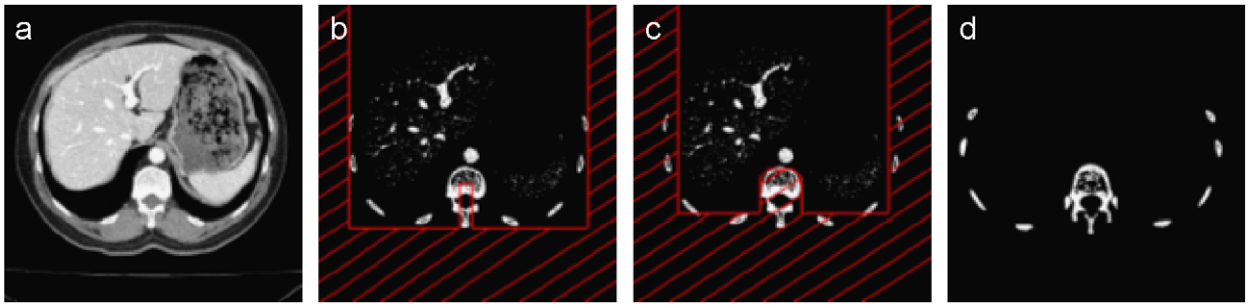

Fig. 4. Removing the spine and the ribs: (a) original image; (b) frame drawn by using the coordinates of the spine and the ribs calculated from the projections; (c) dilation of the frame and the exclusion of the ribs and the spine; and (d) combining thresholded image with the image obtained from the outer side of the frame in (c) using morphological image reconstruction.

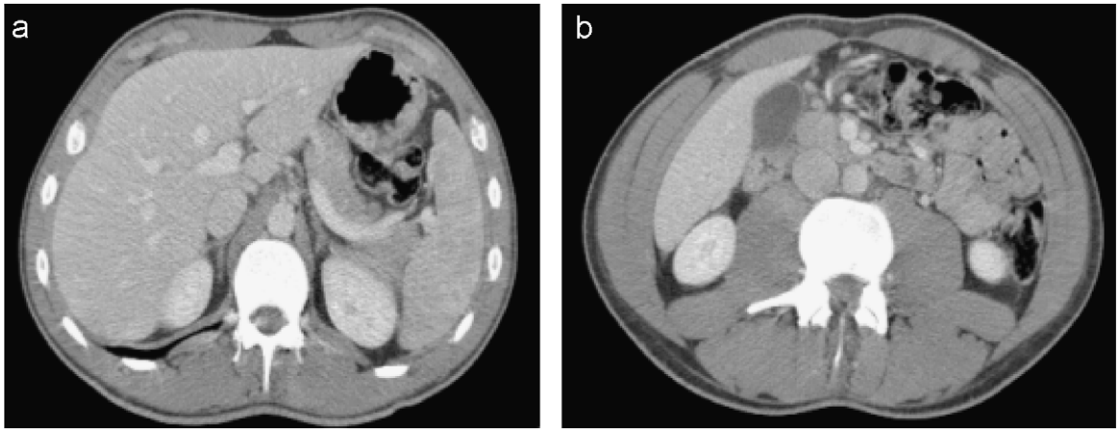

Fig. 5. The right kidney and the liver: (a) at the preceding slices where they cannot be separated easily, and (b) at the succeeding slices where the border between them is more clear.

from 'initial kidney image' and continues through the beginning of the series until the right kidney is removed from the data set. The reason for the direction of the process is the gray level value similarity and unclear boundary between the liver and the right kidney at the slices, where the right kidney begins to appear (Fig. 5a). In such images, the liver and the right kidney could not be separated even by complicated methods. The only way to segment the kidney in those conditions is to use the information obtained from a slice where the liver and the right kidney can be separated easier (Fig. 5b).

The removal of the right kidney starts with the classification of the 'initial kidney image' into five clusters by using the K-Means algorithm. K-Means algorithm [20] partitions the pixels in the image into n clusters by using an iterative procedure. The aim is to minimize the sum, over all clusters, of the within-cluster sums of gray level value-to-cluster centers:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \|x_i^{(j)} - c_j\|^2$$

where $\|x_i^{(j)} - c_j\|^2$ is a chosen distance measure between a gray level value $x_i^{(j)}$ and the cluster center c_j is an indicator of the distance of the n data points from their respective cluster centers. In our system, Euclidean distance metric and batch update method [21] are used where every iteration consists of reassigning gray level values to their nearest cluster centers, all at once, followed by recalculation of cluster centers.

The initial centers of the clusters are chosen to be 30 (for the background), 255 (for the spine and the ribs), threshold found for the fat tissue (generally around 80) and two equidistant gray level values between 255 and the threshold found for the fat tissue. After the application of the K-Means method, it is observed that the kidneys are always assigned to the brightest cluster together with the spine and other very bright tissues (Fig. 6 (top right)).

Then, a seed region is generated at the right of the spine, which is found during the removal of the spine and the ribs (Fig. 6 (top left)). The seed region is determined at a location, where at least some part of the kidney is most likely to exist, by using anatomical information.

To segment the kidney from this result, BIMIR is used. In the 'initial kidney image' the marker is the seed region and the mask is the image that consists of pixels which belong to the brightest cluster The largest connected component after the BIMIR operation forms the kidney (Fig. 6—output image).

The kidneys at the other slices are then detected iteratively. For each preceding slice, the skeleton of the previously segmented kidney is used as the new marker image. And the image

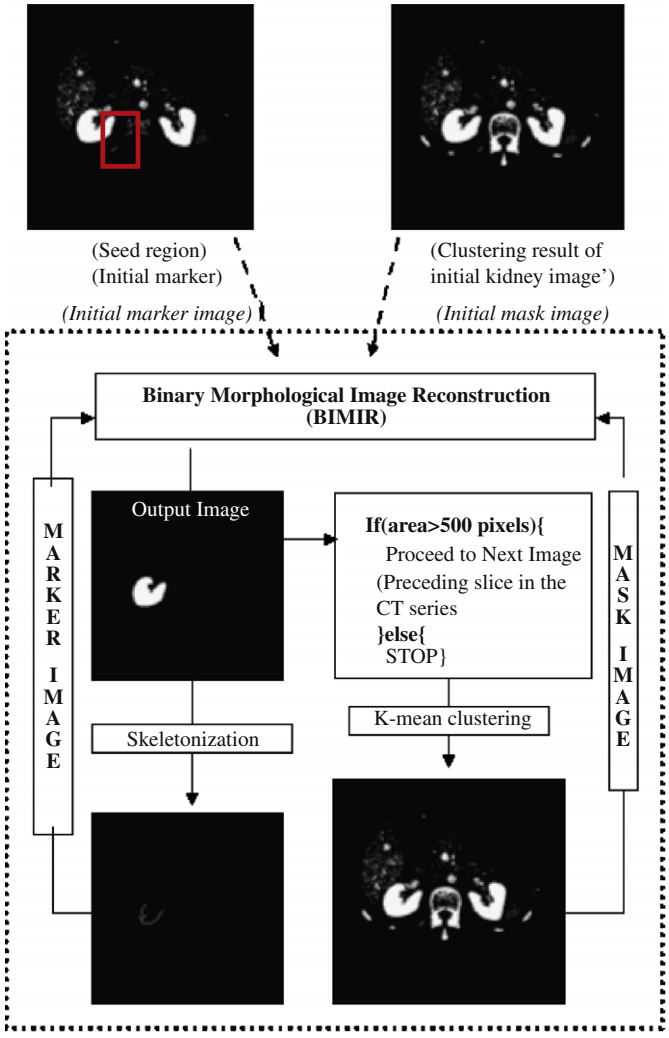

Fig. 6. Determination of the kidney: initially, seed region inside the red frame is used as the marker image and the image after K-Means clustering (brightest cluster) is used as the mask image. Application of BIMIR using these images gives the right kidney. Then, skeleton of the detected kidney (marker image) and clustering result of the next image (mask image) is used to detect the kidney in the next image.

of the pixels that belong to the brightest cluster of the current slice is used as the mask image. Using BIMIR algorithm gives the kidney structure of the current slice. This procedure is illustrated in Fig. 6.

The skeleton of a kidney is calculated using iterative skeletonization (thinning), which is a method to reduce all objects in an image to lines, without changing the essential structure of the image [22]. It removes the pixels on the boundaries of the objects but does not allow objects to break apart. The pixels remaining make up the image skeleton.

Since the skeleton of the previously segmented kidney can be thought as an iteratively eroded image, BIMIR restore exactly the shape of the kidney in the current slice without depending on the similarity between the shapes and the structuring element. By using this method kidneys are detected automatically until the kidney area drops to a user defined value (default = 500 pixels).

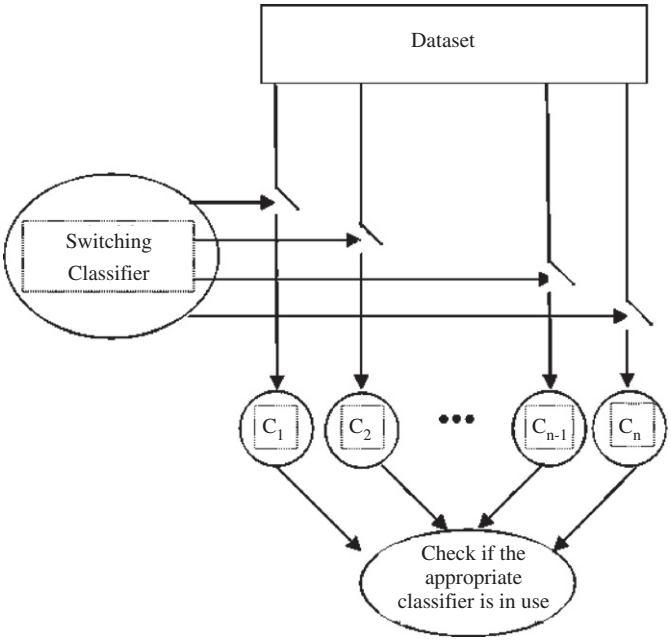

Fig. 7. Proposed modular classification system which constitutes a new kind of system of classifiers. In our case, *N* is equal to two since we have K-Means and MLP classifiers.

This method works efficiently for varying gray level values, shapes and positions of the kidney. It also gives sufficient results invariant of contrast and even when the kidney has lesions. This method; skeletonization combined with BIMIR, is also used in the iterative segmentation of the liver.

3.1.4. ROI selection

For reducing the computation complexity and for increasing the performance of the segmentation algorithm, as much irrelevant information as possible should be removed from the image at the preprocessing. From the anatomy knowledge, we know that the liver is surrounded by the ribs from the left, right and bottom, thus the pixels at the outer side of the ribs can be removed. Also the unnecessary parts from the top (starting from the first non-zero pixel) can be removed. The remaining ROI decreases the image size by 40% in average and reduces the computational complexity significantly.

These four steps of the preprocessing are done for the complete series of CTA images. At the end of the preprocessing, the fat tissue, the bones, the ribs, and the right kidney are removed from the original images and these images are resized using the ROI mentioned above. An example of a preprocessed image is shown in Fig. 7a. In what follows, we applied the segmentation algorithm to these preprocessed images.

3.2. Classification of the liver

The second step of the segmentation procedure is the segmentation of the liver. This step consists of two parts: (1) Automatic selection and segmentation of an 'initial image'. (2) The segmentation of the remaining slices one by one starting from the 'initial image'.

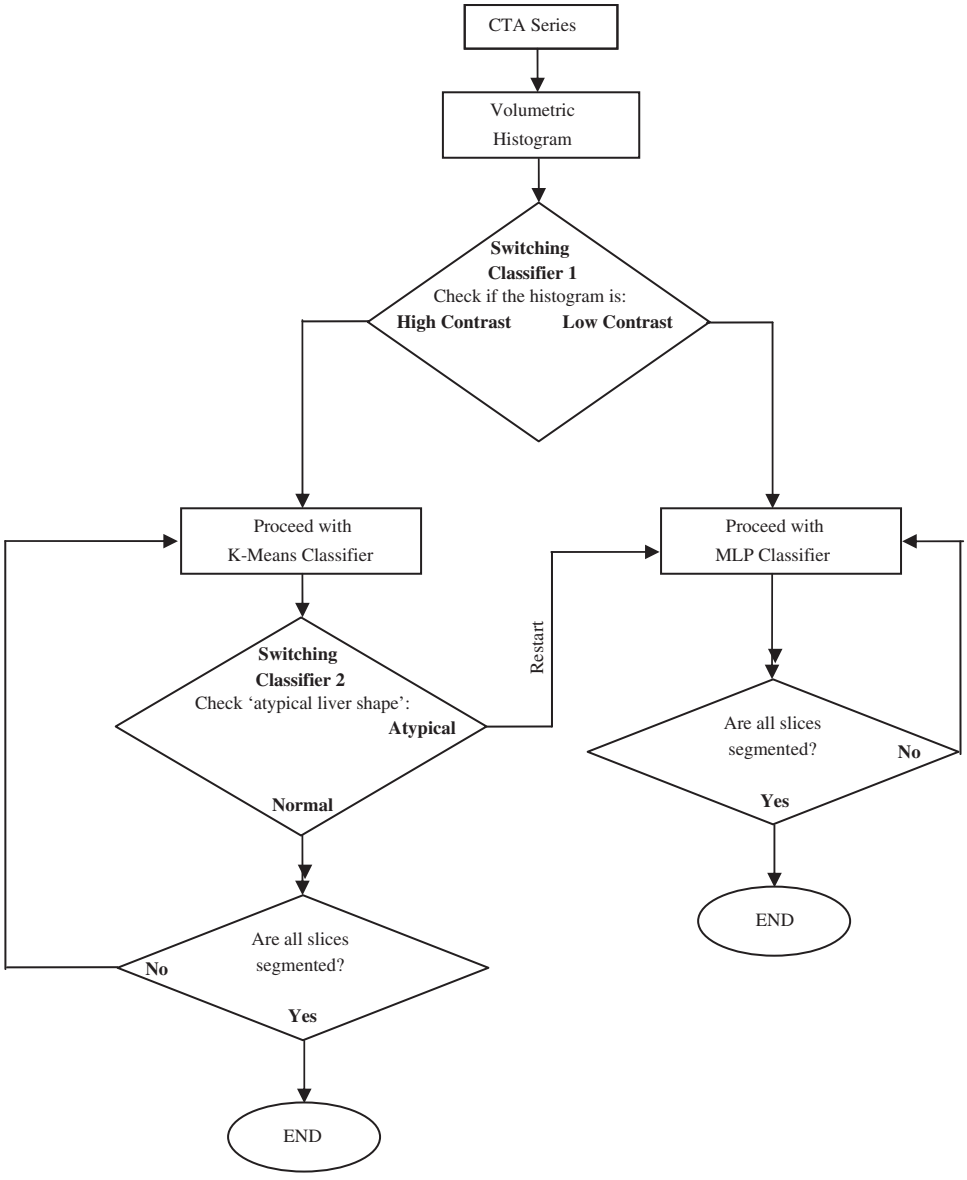

Fig. 8. Switching mechanism and selection of the appropriate classifier for automatic detection of an atypical liver shape and a low contrast image characteristic.

The 'initial image' is a slice where the liver boundary does not overlap with any adjacent organ boundaries, especially the heart and the right kidney. It is selected automatically by choosing the slice that comes just before the first appearance of the right kidney which is determined during the kidney removal stage. After the preprocessing stage, the segmentation algorithm starts from this 'initial image' and then runs through the end of the data set. Then starting from the 'initial image' again, it runs through the beginning of the data set to complete the segmentation process. After the segmentation of the initial image, liver structures at other slices are segmented iteratively.

For this purpose, we introduce a novel modular classification system consisting of a simple classification system (i.e. K-Means based) and a complex one (i.e. MLP based) which are combined with a data-dependent and automated switching mechanism that decides to apply one of them. The introduced modular classification system with data-dependent and automated switching mechanism constitutes a new kind of system of classifiers, some of which are simple and therefore efficient in time-memory requirements with good generalization ability and the others are complex providing a high classification performance, such that depending on the data set, herein the CTA series, one of the classifiers become active.

The switching is based on the detection of "low contrast" data set or atypical liver shape. The switching mechanism does indeed perform a classification task that assigns the CTA series, based on the histogram evaluation and in some cases also according to intermediate results of the K-Means based classifier, into one of the following three categories: (a) low contrast (MLP is employed for this category), (b) high contrast (K-Means is employed for this category), and (c) high contrast—atypical liver shape (MLP is employed for this category). In other words,

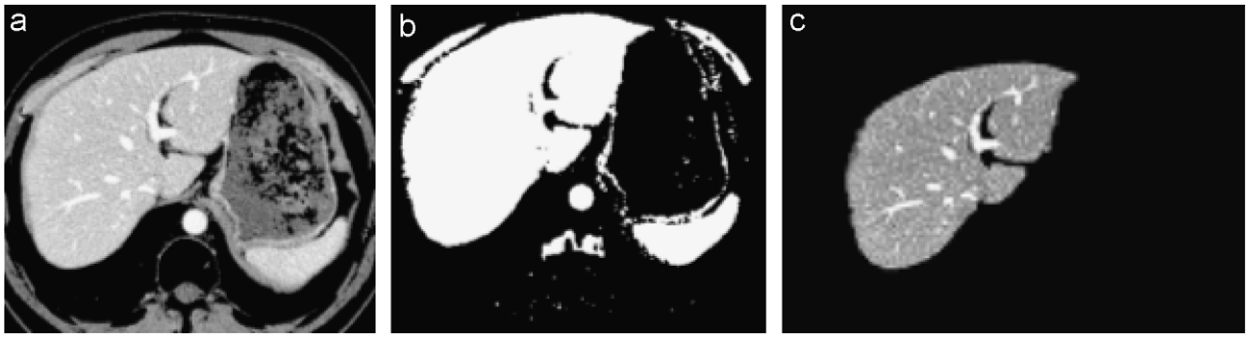

Fig. 9. Initial image segmentation: (a) the preprocessed 'initial image'; (b) clustering result; and (c) the biggest component in the slice.

the switching mechanism selects MLP classifier instead of the K-Means classifier where the K-Means clustering method does not give sufficient results (i.e. if the liver has atypical shape (Fig. 1d) or the gray level difference of adjacent tissues is very low (Fig. 1b). For this reason we develop a process for automatic detection of an atypical liver shape and a low contrast image characteristic (Fig. 8).

In this manner low contrast image characteristics are detected by determining the number of lobes in the volume histogram (discussed in Sections 2 and 3.1). If the histogram has three lobes, the algorithm proceeds with K-Means. Otherwise (if it has two lobes) the algorithm switches to MLP.

The other criterion in the selection of the classifier to use is the atypical liver case. Existence of atypical liver shape means that the liver elongates (stretches) to the left side of the abdomen and becomes adjacent to the spleen which has mostly the same gray value range as the liver (Fig. 1d). In those cases, the boundary between the spleen and the liver is uncertain even for human eye. Therefore, the algorithm may not detect the boundary between the spleen and the liver and segments them together. This causes a dramatic increase in the area of the segmented object. To detect atypical liver shape, we find the area of the segmented object for each slice. Since the change in the area should be smooth for successive slices the algorithm switches to MLP if there is a sudden increase in the area.

It is worth to point that, the result of both classifiers (i.e. K-Means and MLP) gives rough results and these results are refined at the post-processing stage (i.e. removal of the small mis-segmented objects, identification of all components of the liver when the liver dissects into two or more regions). However, when MLP is used as the classifier, the necessary post-processing operations is significantly less than the operations used after K-Means classifier.

3.2.1. Initial image segmentation

As previously mentioned, 'initial image' is a slice that should satisfy three constraints. First of all, the liver should have a relatively big area in the slice but it does not have to be the biggest organ. Second, it should consist of one connected component and third, the liver boundary should not overlap with any adjacent organs, (i.e. the heart and the right kidney). An example of an initial image is given in Fig. 4a. After preprocessing (Fig. 9a), the image can be classified into three clusters: background, bright organs and dark organs. Excluding the background pixels from the process, the two clusters are found using the Otsu's method [23] by finding the optimal threshold to separate dark organs (stomach and muscle) and brighter organs (liver, spleen, and heart).

Otsu's method chooses the threshold value k that maximizes the between-class variance σ_B^2 which is defined as

$$\sigma_B^2 = w_0.(\mu_0 - \mu_T)^2 - w_1.(\mu_1 - \mu_T)^2$$

where

$$w_0 = \sum_{q=0}^{k-1} p_q(r_q), \quad w_1 = \sum_{q=k}^{255} p_q(r_q), \quad \mu_T = \sum_{q=0}^{255} q \cdot p_q(r_q),$$

$$\mu_0 = \sum_{q=0}^{k-1} q \cdot p_q(r_q)/w_0, \quad \text{and} \quad \mu_1 = \sum_{q=k}^{255} q \cdot p_q(r_q)/w_1$$

Here $p_q(r_q)$ is the discrete probability density function, as in $p_q(r_q) = n_q/n$ where n is the total number of pixels in the image, and n_q is the number of pixels that have intensity level $r_q (q = 0, 1, \ldots, 255)$.

Selecting the cluster which has the foreground organs (i.e. the right side of the threshold) (Fig. 9b) and then taking the biggest connected component in the slice results with the segmented liver structure (Fig. 9c).

It is important to point that this method gives sufficient results only under the constraints given for the 'initial image'. For our data sets, the slices, which are located around the one third of the series (i.e. 30th slice of 90 slice CTA series), are determined to be suitable for being 'initial image'. However, it is also possible to select any slice in the series, which satisfies the necessary conditions, as 'initial image'.

3.2.2. Segmentation with K-Means

After segmenting the initial image, the algorithm first runs downwards to the last slice and then upwards to the first slice starting from the initial image.

It is clear that the preprocessed images contain the liver, the tissues and the organs that have similar or darker gray level values than those of the liver because the brighter tissues and

organs are removed at the preprocessing phase (except the bright tissues inside the liver).

By using the K-Means method an unsupervised clustering is applied to these preprocessed images for classifying the organs into two clusters. At this step, the initial cluster centers are determined as follows: For the 'initial image', the first cluster center is given as the minimum gray level value of that image (excluding the background). The second cluster center is determined as the mean value of the segmented initial image. Then for the other slices, the centers found in the preceding slice are used as the initial centers.

The brighter cluster at the clustering result is preserved since it always consists of the liver. The identification of the liver including its dissected parts and removal of the incorrectly segmented objects is done at the post processing stage.

3.2.3. Segmentation with MLP

Although segmentation of the liver using the K-Means clustering method generally gives sufficient results, it fails when the liver has atypical shape (Fig. 1d) or the gray level value difference between the liver and the adjacent tissues are very low (Fig. 1b). To obtain acceptable results also for these cases, a feature based segmentation process is developed. In feature based segmentation, first, K-Means is tested as the classifier. However, the segmentation results were not adequate hence a more complex classifier, MLP is used instead. The overall structure of the segmentation process using MLP is shown in Fig. 10.

In segmentation with the MLP, the preprocessed 'initial image' and the segmented 'initial image' are used for initial training of the network. From the preprocessed 'initial image' two features (Mean and standard deviation) are calculated. Then, the distance transform is applied to the segmented 'initial image' and the pixel values after the transform are taken as an additional (third) feature.

The features are calculated in a window of size 9×9 centered for a given pixel. Although some sudden changes in the image (i.e. edges) can be identified more accurately with a smaller window size, the optimal size for identifying the liver region is decided to be 9×9 after extensive experimentation by concerning to prevent finding details inside the liver area and to get better information about the orientation of the liver. Also in [13,14], 5×5 window size is presented to be optimal for 256×256 and 240×320 images, therefore using 9×9 for 512×512 images is appropriate.

To represent the homogeneous regions (i.e. Liver parenchyma) in the current slice, we use the mean feature (Fig. 11a). For a pixel the mean gray level value is calculated by

$$\overline{x}_{ij} = \frac{1}{N} \times \left(\sum_{i-4}^{i+4} \sum_{j-4}^{j+4} x_{ij} \right)$$

To represent the edges (i.e. liver boundary) we use the standard deviation feature (Fig. 11b) which is calculated as

$$\sigma_{ij} = \frac{1}{N} \times \left(\sum_{i-4}^{i+4} \sum_{j-4}^{j+4} (x_{ij} - \overline{x}) \right)$$

where $\overline{x}_{ij}$ is the mean value of the pixel $\times$ located at the position i, j and σ_{ij} is the standard deviation of the pixel located at the same position. N is the total number of pixels in the window which is equal to 81 in our approach.

Finally, the distance transform feature is used to represent the previously segmented image. The distance transform provides a metric that measures the separation of the pixels in the image. The metric is calculated to measure the total Euclidean distance along the horizontal, vertical, and diagonal directions (Fig. 12a).

In our algorithm, the distance transform gives information about the liver location at the adjacent (preceding/succeeding) slice. Since the liver size and location does not change dramatically between adjacent slices, the distance transform of a segmented liver (Fig. 12b) gives quite important information about the liver location at the adjacent (preceding/succeeding) slice.

The distribution of the features in the feature space is shown in Fig. 13. The lighter data points (x : [0 0.1], y : [0 1]) correspond to the pixels that belong to the liver in the figure. From Fig. 13a, it is observed that the most discriminative feature is the distance transform. However the distance transform is not sufficient to discriminate the data inside the circle in Fig. 13b. It is observed that the data inside the circle becomes separable by using the mean and standard deviation features.

After the extraction of features, first, these three features, distance transform, mean and standard deviation, are used as the inputs of the K-Means clustering algorithm. Although, there is a slight increase in the performance of classification, it is observed that atypical liver cases and 'low contrast' CTA series cannot be segmented properly. Therefore, a more complex classifier is needed and MLP is chosen as the classifier.

The network is trained initially by using the segmented initial image as the desired output and the calculated three features as the training data. At the network output, each input pixel is classified as belonging to the liver region or lying outside the liver region on the basis of these features.

After this initial training, weights are updated and iteration proceeds to the next slice. Mean and standard deviation features are calculated for the preprocessed next slice (which is the current slice to be segmented) and the distance transform is calculated from the previously segmented image. By using these features and the weights (obtained from the previous slice), the current slice is segmented. After the segmentation of the current slice, the network is trained again by using the features (mean, standard deviation and the distance transform calculated for the segmented images of current slice) as the input and the new segmented image as the desired output. After the training and calculation of the weights, the algorithm proceeds with the next slice and this iterative procedure continues until all images are processed. Using the previously adjusted weights as the initial weights of the next training phase, the training time is reduced significantly.

The MLP structure used for the segmentation consists of three neurons at the input layer, which corresponds to the number of input features. There are eight neurons at the hidden layer, each of which has a bias input that ranges between ±1. The biases are updated along with weights during error backpropagation [21]. The output layer consists of one neuron. The

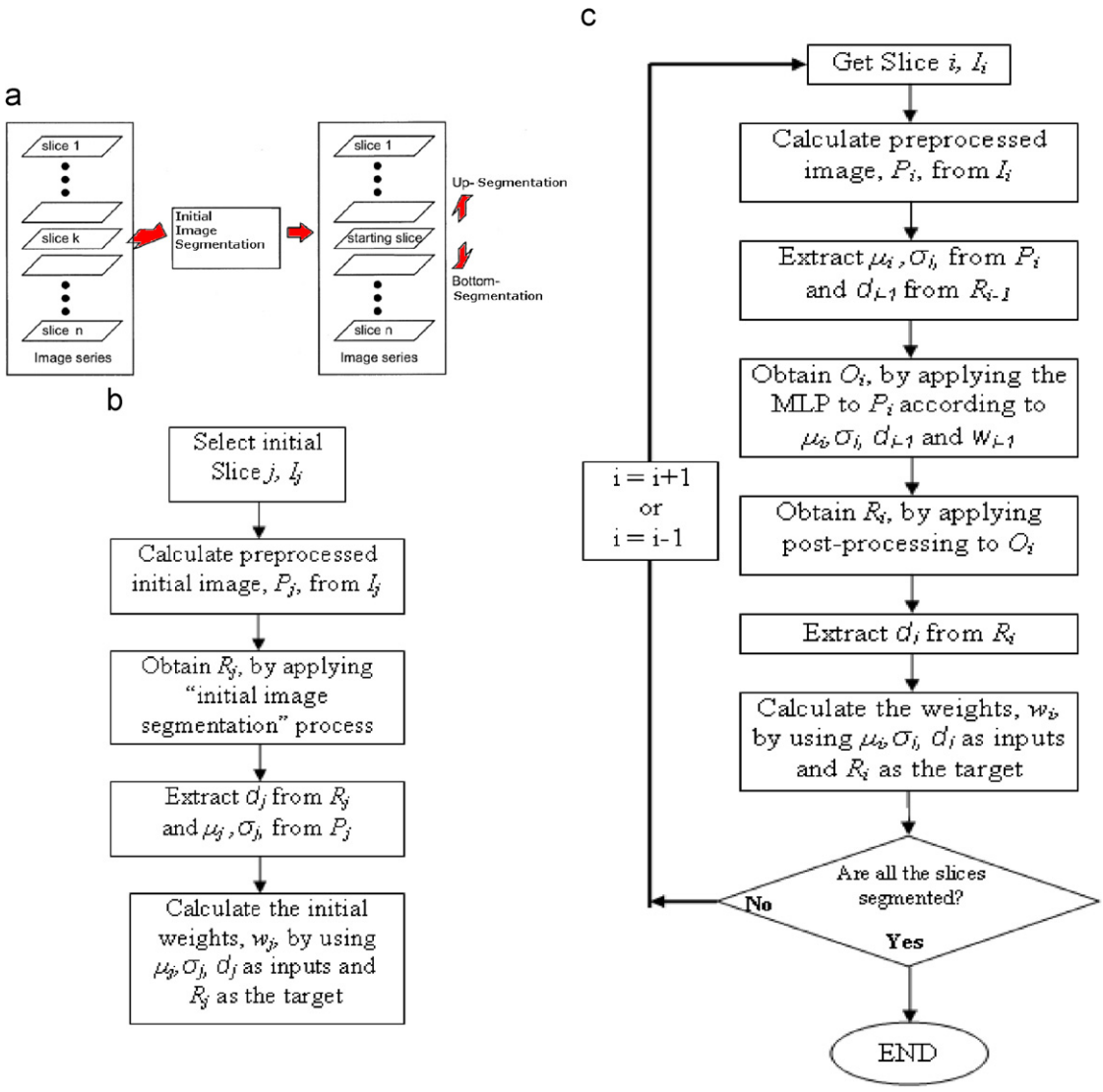

Fig. 10. (a) Segmentation process by using MLP; and (b) preprocessing of all images is followed by the selection and segmentation of the initial image. The initial training is done by using the segmented initial image as the desired output and the feature vectors obtained from initial and segmented initial images as training data, (c) then, the algorithms proceeds to next slices and at each slice previously found weights are used for classification with MLP. The weights are updated using the current segmented image as the desired output and the feature vectors obtained from the current slice as training data. In the figure μ_i, σ_i, and d_i corresponds to mean, standard deviation and the distance transform features obtained from slice I, respectively where $i = j-1, j-2, \ldots, 2, 1$ for up-segmentation and $i = j+1, j+2, \ldots, N-1, N$ for bottom segmentation with N is the number of slices.

output of the network lies between 0 and 1 for each pixel and it is thresholded by 0.5. Then, for an input region belonging to the liver class, the output is designed to be unity whereas for all other the output is designed to be zero. This network structure is determined to be the optimum after extensive experimentation and due to a compromise between efficiency and reliability.

3.3. Analysis and comparison of features and classifiers

In [14], five optimum features are reported as follows: mean gray level, standard deviation, skewness, entropy, homogeneity. However, these statistical descriptors are not sufficient to differentiate two organs/tissues that have similar texture and/or statistical properties. This drawback limits the usage of these features in atypical liver shapes where the border between the spleen and the liver vanishes. Since the spleen has almost the same texture and statistical properties as liver, it becomes very difficult to segment the liver without the spleen. Moreover, in 'low contrast' CTA series, the statistical properties of muscle tissues and vasculature (i.e. aorta, inferior vena cava) get closer to the liver which also hardens the correct segmentation of the liver by only using these features without any spatial information.

To overcome these problems, the distance transform is selected as an additional feature. Our approach in the selection of the distance transform is to simulate the decision process of a radiologist. For instance, in the case of atypical liver shapes, the unclear border between the liver and the spleen is

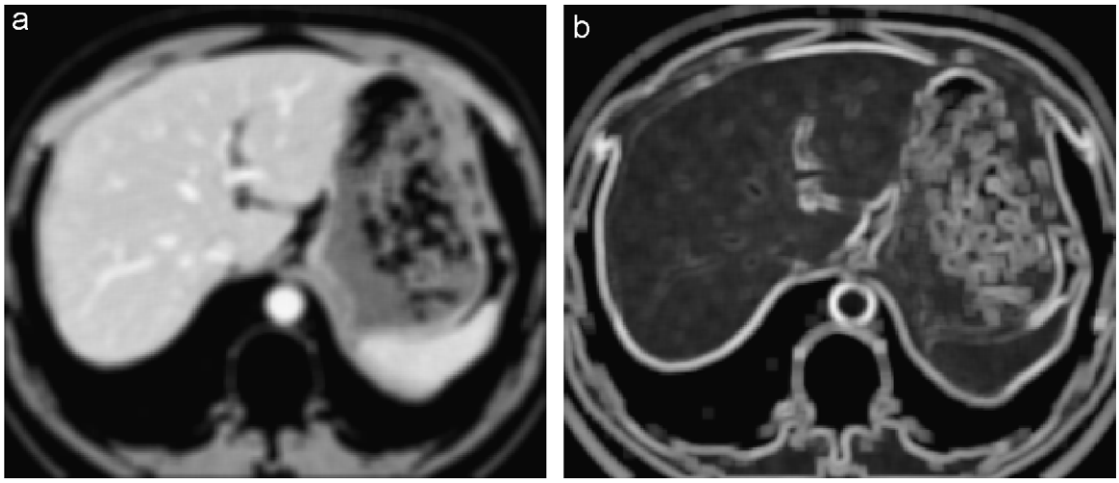

Fig. 11. Feature images: (a) mean; and (b) standard deviation.

a

2.8	2.2	2	2.2	2.8
2.2	1.4	1	1.4	2.2
2	1	0	1	2
2.2	1.4	1	1.4	2.2
2.8	2.2	2	2.2	2.8

b

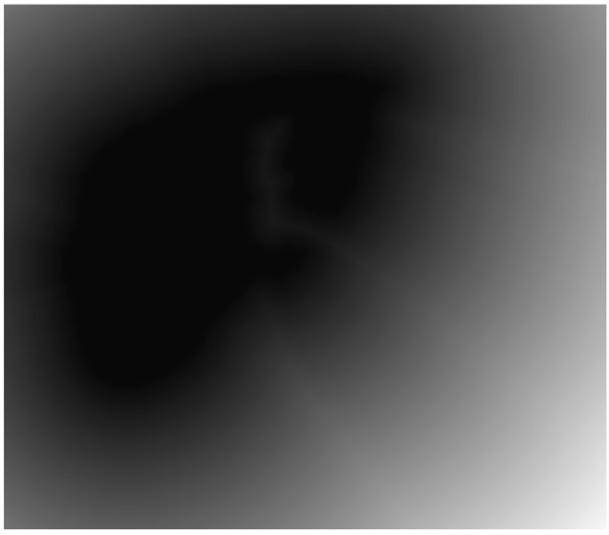

Fig. 12. Distance transform process: (a) distance transform of a binary image in which only the pixel at the center has value 1; and (b) distance transform of a segmented slice.

determined by the radiologist by following the slices (especially the ones just before and after) where the border is more visible. Similarly, the distance transform feature provides a metric to represent the liver at the previously segmented slice which gives information about the liver location at the adjacent (preceding/succeeding) slice. Although this feature is affected by the slice spacing of CTA data, the slice thickness (ST) of the data sets used in this study is 3.2 mm, which is a rather big value considering the emerging technology of CT modalities. Even with this thick ST, the information provided by the distance transform is enough to handle atypical liver shapes and 'low contrast' data sets.

In the following figure, an example case is presented to show the advantages of using distance transform for detection of the liver. Fig. 14 shows a set of images that are selected from a CTA series where the liver has atypical shape (Fig. 1d). Fig. 14a–d shows the outputs of the K-Means classifier while Fig. 14e–h represents the outputs of the MLP classifier. The slice numbers are 33, 31, 28, and 24 from left to right where the slice 33 is the initial image to be segmented. As shown in Figs. 14a and e; both MLP and K-Means are successful in classifying the liver without any connectivity with other objects. However, as the segmentation process continues through the beginning of the CTA series, the liver gets closer to the spleen (Fig. 14b and c) and they finally they merge (Fig. 14d). Since the texture and gray level variation of the liver and the spleen are very similar, it is necessary to use the information obtained from the previously segmented slices. The distance transform allows the usage of this information as it limits the search area to a region that is slightly bigger than the previously segmented liver region. As shown in Fig. 14f–h, the distance transform prevents the spleen to appear at the output of the classifier.

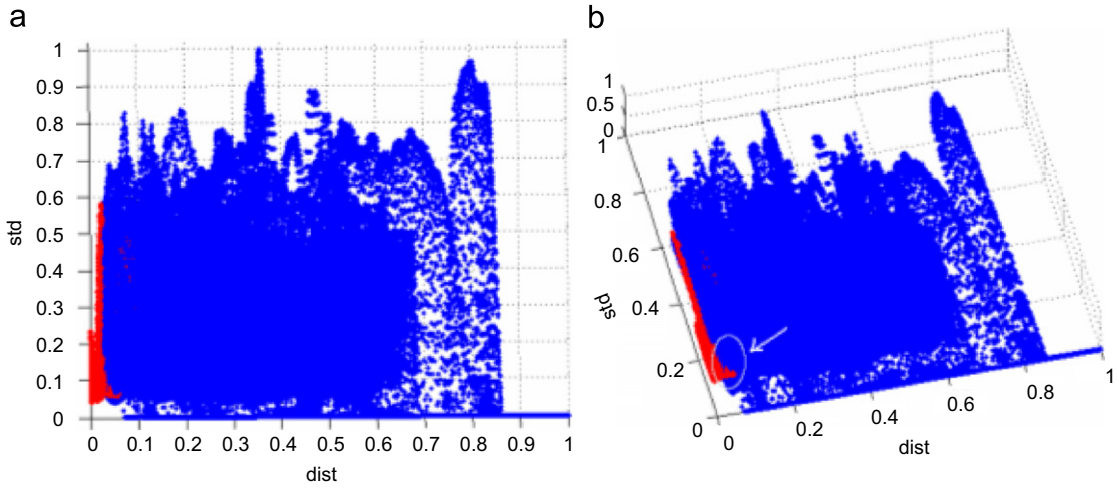

Fig. 13. Feature space from two different views: (a) the most discriminative feature is the distance transform since most of the data is separable along the distance transform space; and (b) however, it is not sufficient to discriminate the data inside the circle and additional features (i.e. mean, standard deviation) are required.

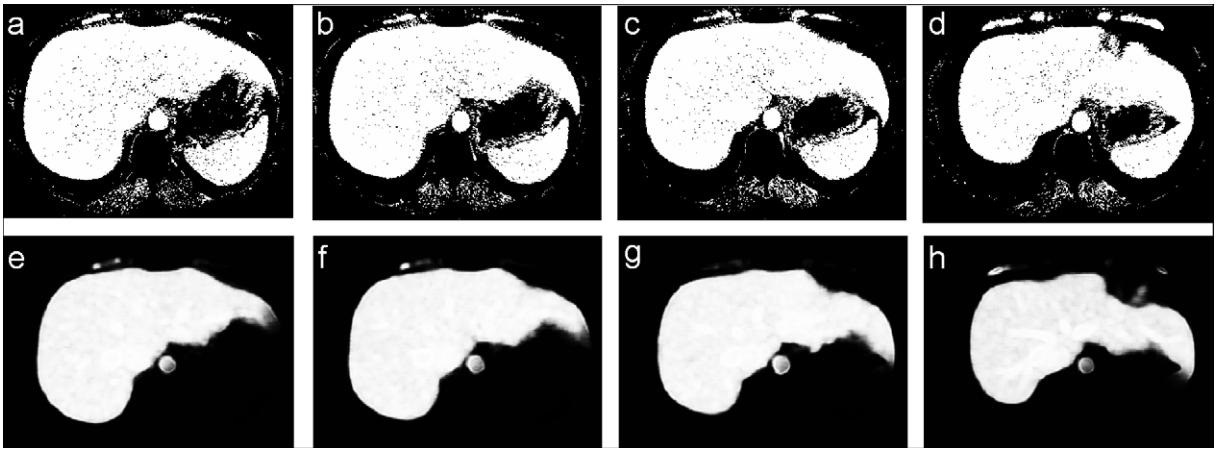

Fig. 14. The effect of the distance transform in segmentation of atypical liver. Each figure shows the output of the classification results with K-Means or MLP without any post-processing: (a) the 'initial image' can be segmented by K-Means (Slice 33); (b) as the slice by slice segmentation proceeds, the spleen, which has almost the same texture and statistical properties as liver, becomes closer to the liver (Slice 31); (c) as spleen and liver gets closer, their border cannot be identified clearly since they belong to the same cluster due to the similar characteristics (Slice 28); (d) when the borders of the liver and the spleen intersects, the clustering is not enough to classify these two organs (Slice 24); (e) the 'initial image' can also be segmented by MLP (Slice 33); (f, g) as the slice by slice segmentation proceeds, the distance transform limits the classifier output to a region close to the location of previously segmented liver (Slices 31–28); and (h) the effect of distance transform prevents the misclassification of the spleen even when the border between the liver and the spleen vanishes (Slice 24).

Here, it is worth to point that the distance transform is combined with the features proposed in [14] as another experiment. However, after extensive simulation studies no noticeable increase is observed in the performance of segmentation when skewness, entropy, and homogeneity are included in the feature set. But these three features increase the computational burden and slow down the process dramatically. Therefore, skewness, entropy, and homogeneity features are not used during classification.

Another important issue is the training of the neural network since it has significant effect on the performance and computation time. The technique proposed in [13] requires more than one manually segmented image as training data which is undesired for automated processes. In [14], training is done with a limited set of images and due to the high variation of image characteristics, a larger and more diverse database is recommended to generalize this system for reliable performance. After the adjustment of the network weights using the training set, these fixed weights are used for segmentation of other data sets. However, this approach is error prone and needs new training sets for the data sets with new image characteristics such as different modalities and modality settings. Moreover, our simulations show that using a fixed hyper-plane (network weights) to segment all images in a CTA series, decreases the segmentation performance. This means that a new set of weights for each image increases the segmentation performance and the weights should be adjusted during the segmentation of different slices of the same CTA series. To provide this adjustment,

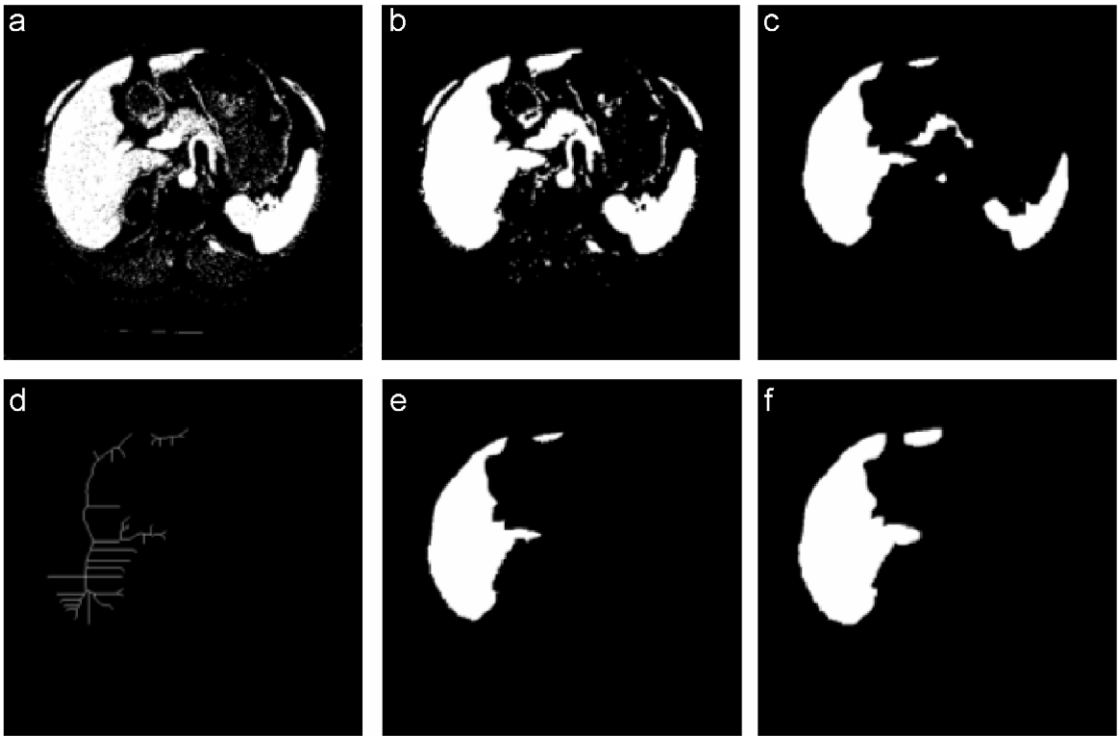

Fig. 15. Post-processing stage for a slice: results of: (a) K-Means clustering; (b) median filtering; (c) erosion; (d) skeletonization of the previously segmented liver; (e) BIMIR using median filtering result as the marker and skeleton as the mask; and (f) dilation to restore the contours that have not been completely removed by the erosion.

the weights of the proposed network are updated at each slice. This is done by using the original form of a slice as the input and the segmentation result as the desired output of the training. Initialized by the "initial image", the previously adjusted weights are used in the segmentation of the next slice. The same weights are also used as the initial weights of the next training phase which reduce the training time significantly. A similar training methodology is proposed in [15], where a contextual neural network with a high segmentation performance is proposed. But the results show that it fails where the gray level of the desired region is too close to the adjacent tissues since the proposed method is designed for the segmentation of all abdominal organs.

3.4. Post-processing

The results obtained from the segmentation algorithms are roughly segmented liver structures (Fig. 15a). To remove small mis-segmented objects and for boundary smoothing, a post-processing is needed. Moreover, identification of all components of the liver when the liver dissects into two or more regions is also done at the post-processing stage.

Post processing is handled differently for the slices before and after the "initial image" slice because of the different image characteristics.

In the post processing phase of the slices after the 'initial image', a series of nonlinear filtering and morphological operations is applied to separate weakly connected components in the clustered binary image. First, a median filter is applied to remove the white spots appear at the background and black spots appear inside the liver (Fig. 15b). Then, erosion operation is applied to eliminate the small unconnected objects that do not sit within the structuring element (Fig. 15c). Next the skeleton of the previously segmented liver is obtained using skeletonization (Fig. 15d). By using BIMIR, in which the skeleton is used as the marker image and K-Means or MLP result as the mask image, the liver is identified (Fig. 15e). Usage of skeletonization also provides the important information about the separated parts of the liver when it dissects into two or more regions (Fig. 15d). The algorithm searches for the second and the third components automatically for the slices after the 'initial image'. After that, dilation operation is applied to restore the contours that have not been completely removed by the erosion (Fig. 15f). Finally a Gaussian filter is applied to the edges to smooth the contours.

The difference between post processing of the slices before and after the initial image is due to the existence of heart tissue in the slices before the initial image. For the slices before the initial image, the liver area from the previously segmented slice is not dilated because the heart can also be included due to gray level similarity of it with the liver. Another difference is that the algorithm looks for only one connected component for the first slices.

Post processing is handled differently also for the low and high contrast data sets. In low contrast data sets; nonlinear filtering is applied in a different way such that the value after

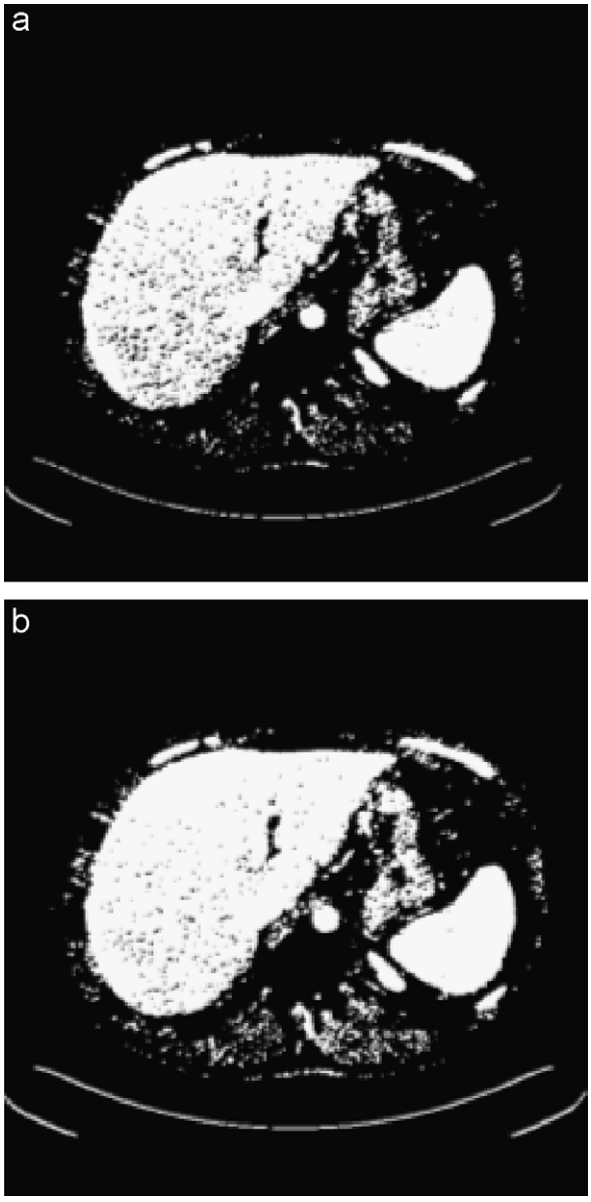

Fig. 16. (a) Affect of median filtering and (b) affect of modified nonlinear filtering.

the median is taken instead of the median value. This kind of filtering increases the connectivity of the liver by eliminating more dark spots than median filtering (Fig. 16a and b).

4. Evaluation

The segmentation results are evaluated by using the area error rate (AER) [17]. AER is defined as the area difference between the region segmented by the algorithm (RA) and the region segmented manually (RM). Defining a union region RU as RA ∪ RM and an intersection region RI as RA ∩ RM, AER is equal to:

$$AER = \frac{RU - RI}{RM} \times 100\%$$

AER is similar to the criteria volumetric overlap error (VOE) which is used in [25]. VOE is defined as

$$VOE = \frac{RU - RI}{RU} \times 100\%$$

VOE is 0 for a perfect segmentation and has 100 as the lowest possible value, when there is no overlap at all between segmentation and reference. AER also takes zero values for a perfect segmentation however the lowest possible value is not limited to 100. In our evaluation, AER is calculated directly (without any boundary modification) between the manually and automatically segmented images. The manually segmented images are segmented by an expert radiologist from Dokuz Eylül University (DEU) Radiology Department.

The slice by slice average AER for 20 data sets (‘5’ low contrast and 15 ‘high contrast’, three of which have atypical liver shape) with the K-Means algorithm is shown in Fig. 17a. It is observed that the algorithm shows better performance for the slices at the middle and at the end of the data sets. The high AER values for the slices at the beginning of the data sets are due to unclear boundary between the heart and the liver. The average AER for the complete data set is calculated as 12.15% by using the K-Means algorithm.

As explained in Section 3.2 the overall algorithm switches automatically from K-Means to MLP when it is needed. To illustrate the effect of this switch on the segmentation performance we performed several experiments. Fig. 17b shows the average AER for each patient data set when the system just uses the K-Means. The high AER values in data sets 8, 9 and 15 are due to atypical liver shapes.

As mentioned above, the main points where the K-Means algorithm fail are the first slices where the heart and the liver can hardly be segmented even with the human eye, the patient data sets with atypical liver shapes (Fig. 18a) and the separation of the tissues in ‘low contrast’ data set where the gray level value of the adjacent tissue (organ or vessel) is very close to the liver (Fig. 18c and e). The algorithm with the neural network, MLP, classifier solves these problems and performs better segmentation in those cases (Fig. 18b, d, and f). These high AER values are reduced from 41.20% to 12.73% in data set 8, 18.20% to 9.95% in data set 9, and 29.8% to 11.30% in data set 15 by using MLP. The high AER values in 2, 12 and 13 are due to low contrast between the liver and its adjacent tissues, organs. These high AER values are reduced from 14.7% to 10.16% in data set 2, 17.40% to 11.6% in data set 12, and 17.1% to 10.24% in data set 13 by using MLP.

However, the time required for the algorithm with K-Means classifier is less than MLP. Therefore, it is necessary to use the overall system, which takes the advantage of both classifiers, to obtain the optimum results.

Moreover, it is clear that, AER is very sensitive to the pixel differences between automatically and manually segmented images. Therefore even 1 or 2 pixel difference between these images increases the error rate significantly especially at the boundaries even when no modification is needed. Therefore, a qualitative evaluation is also made by an expert radiologist. The evaluation of the expert is based on his idea if a slice needs

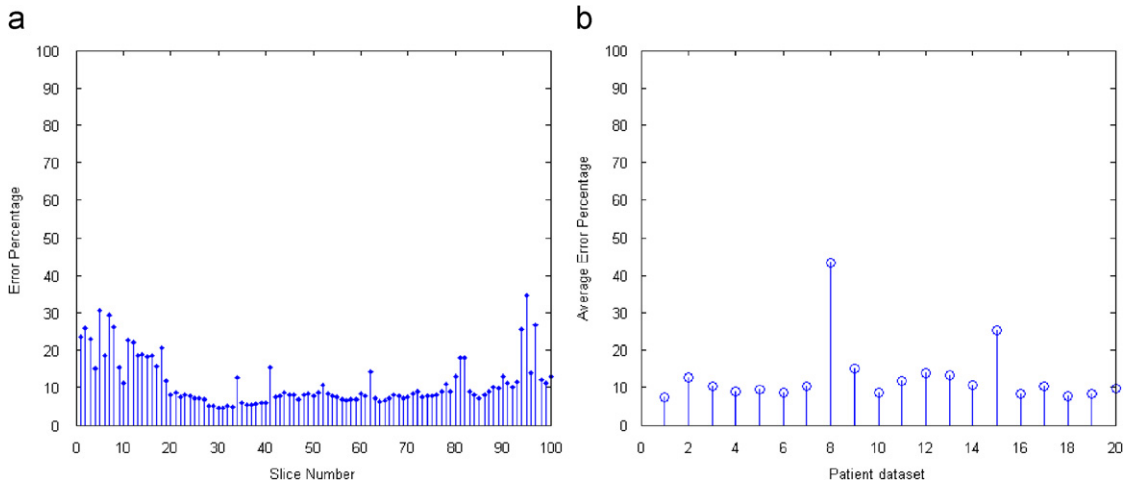

Fig. 17. (a) Average AER calculated for the slices, which have identical number, in all data sets. (i.e. the error percentage for the 40th slice represents the average AER of 40th slices in all data sets); and (b) average AER calculated for each data set if just K-Means algorithm is used.

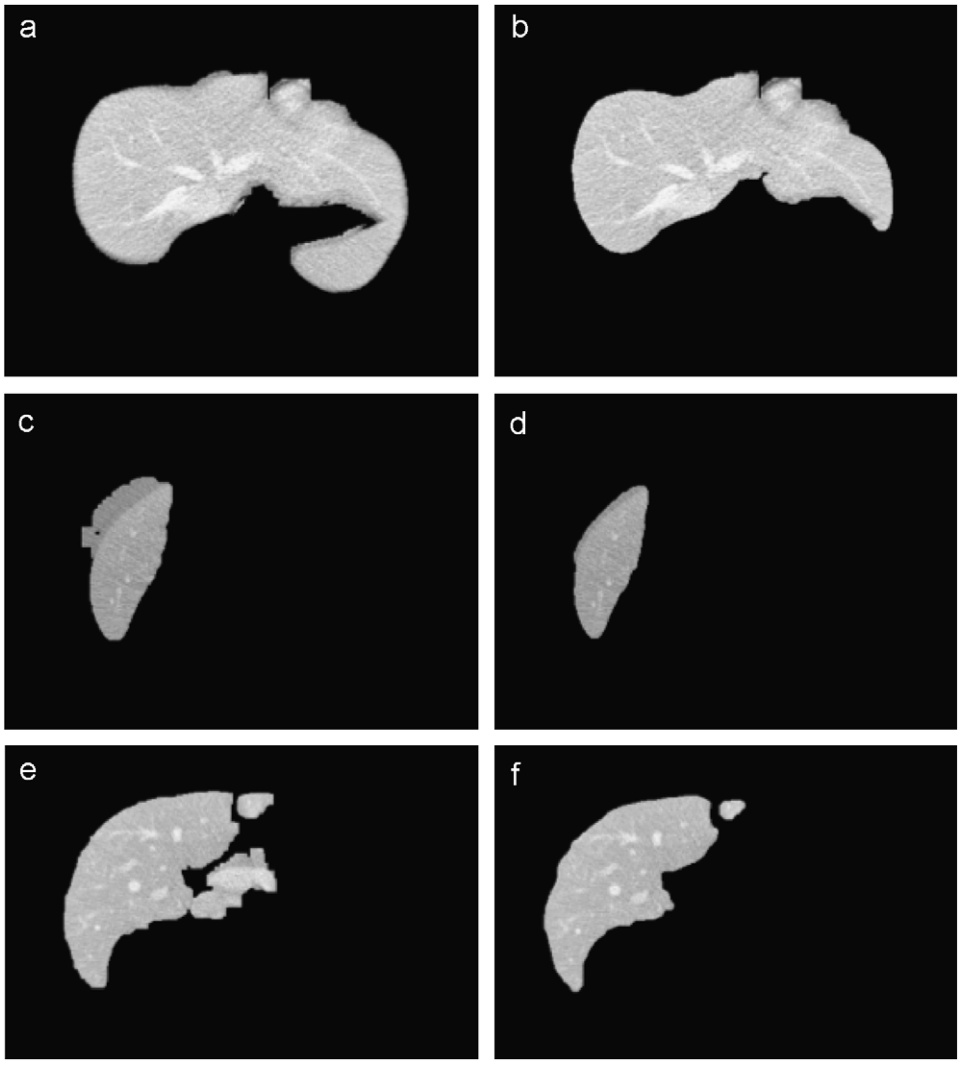

Fig. 18. (a) Segmentation results for Fig. 1d: algorithm result with K-Means; (b) algorithm result with MLP; (c) segmentation results for Fig. 1b: algorithm result with K-Means; (d) algorithm result with MLP; (e) segmentation results for Fig. 1a: algorithm result with K-Means; and (f) algorithm result with MLP.

modification or not. Then the AER is calculated for the slices that need modification. Fig. 19 shows the average AER calculated for the slices that need modification for each patient data set by using the overall system. The results show that the average AER is reduced significantly to 5.09% with a minimum of 2.4% and a maximum of 7.63.

As a last analysis we calculated volume measurement of the liver for comparing the segmentation performances of the algorithms (K-Means, MLP and overall) and the success rate of our approach is (Fig. 20a). To measure the volume of the segmented parts, pixel spacing and slice thickness values are used from DICOM Meta information. The error range of the volumes obtained with the K-Means algorithm is between 20 and 210 mm^3. The percentage errors of these measurements are found between 0.7% and 16.26% with a mean error of 4.7%. The same analysis for MLP show that the error range of the volumes obtained with the K-Means algorithm is from 40 to 160 mm^3 and from 2% to 12.8% with a mean of 7.5%. The results again show that the MLP shows better performance at the data sets with a typical liver shape (8, 9, and 15) and low contrast adjacent tissues (2, 12, and 13) although the K-Means is better in overall average error.

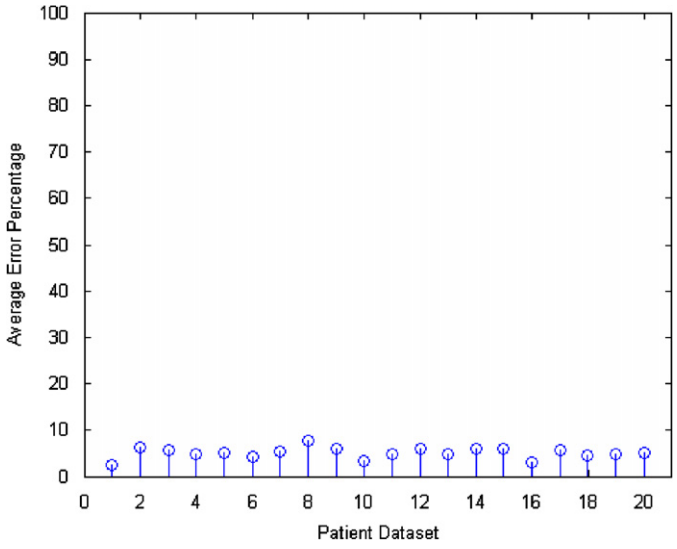

Fig. 19. AER calculated for all data sets but using only the slices that are determined by the expert as further modification is needed

The overall system performance is shown in (Fig. 20b). The error range of the volumes obtained with the algorithm is from 20 to 140 mm^3. The percentage errors of these measurements are found between 0.7% and 12.8% with a mean error of 6.4%.

The Java version of the program with the K-Means algorithm runs for 12–17 min in a standard PC with 2 GB Ram and 3 GHz processor and requires 750 MB of memory [24]. The Matlab® version runs more slowly and takes around half an hour with K-Means classifier. The algorithm with the MLP classifier ends approximately in 45 min both in Matlab® and in Java. On the other hand it takes around 60–90 min for an experienced user to segment liver from 100 slices manually and it requires user experience both on the liver and the software which should consist of the necessary tools for manual segmentation of the liver. In comparison with manual segmentation tool that is currently in use, our algorithm is clinically feasible and much more efficient in terms of time and efficiency.

5. Discussion and future work

The paper proposes a robust and efficient method that can automatically segment the liver of transplantation donor candidates in any CTA series. The success rate is calculated as 94.91% over a data set of diverse CTA series of 20 patients

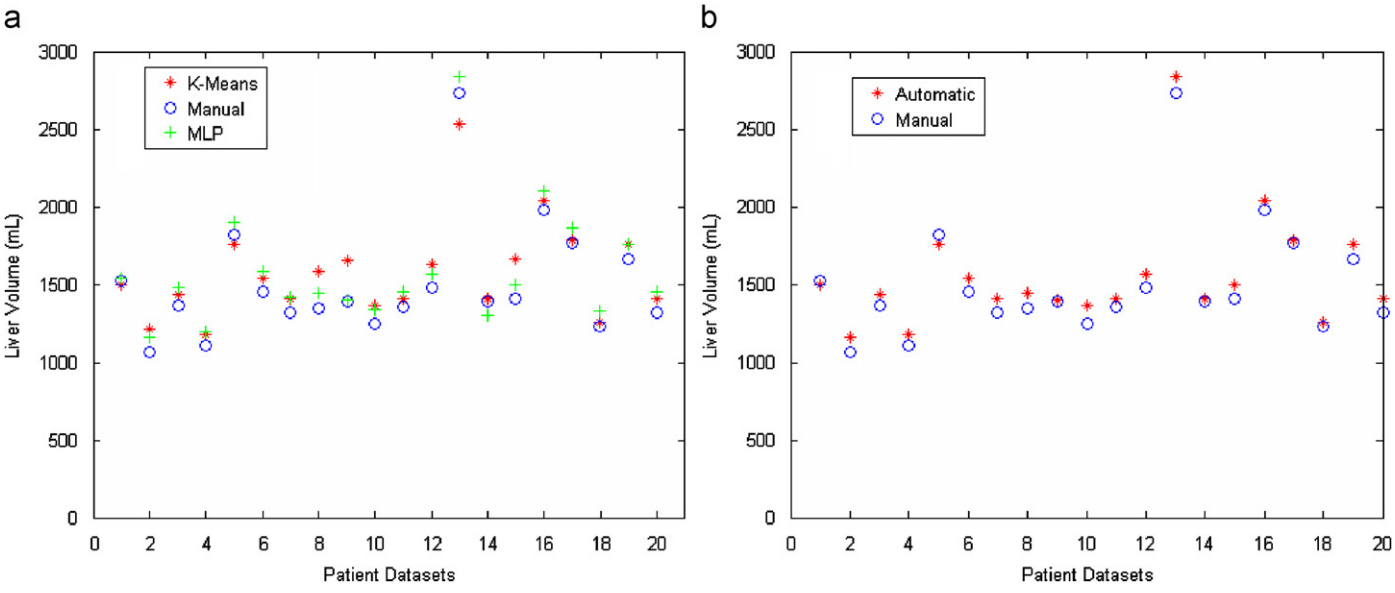

Fig. 20. Volume measurement results of segmented parts: (a) comparison of volumes obtained by MLP and K-Means with manually segmented; and (b) overall system performance.

according to the evaluation of the expert radiologist experienced on pre-evaluation of transplantation donors for more than 100 cases.

In this study, a robust algorithm, that can automatically segment the liver in any CTA series, is established. The robustness of the method follows from its capability of dealing with the contrast variations and atypical liver shapes and this capability is provided by the patient oriented structure of the algorithm. For qualifying 'patient oriented', the algorithm learns the data set characteristics in parallel to segmentation process, and adapts its parameters to these characteristics. This strategy involves a segmentation method which does not utilize a common parameter set found from all patient data sets. Instead, the method is capable of adapting the parameter set to each patient. So the wide ranges of the parameter values are covered and the developed system is sensitive to all variations in a data set by adopting its parameters due to the data set characteristics.

The ability of dealing with the contrast variations and atypical liver shapes first by recognizing the existence of these problems and then by solving the segmentation problem using inter-slice information provided by the distance transform. To our knowledge, there is no method in the literature which handles all these difficulties at the same time. This ability is gained:

(i) by introducing the distance transform as a feature for each slice and then using this information in the succeeding slice to reveal three dimensional properties of the liver which cannot be obtained by the set of slices processed individually,
(ii) by devoting different (MLP) classifiers for different slices each of which is fed by three features such as mean, standard deviation and distance transform as opposed to the automatic organ segmentation methods available in the literature which use a single classifier for the whole set of slices, some uses statistical features extracted from three dimensional data [16] and some uses five features for each slice [14],
(iii) by reducing the number of features and by initializing each (MLP) classifier's weights with the weights of the previous one, so getting a good efficiency in terms of time and memory requirements,
(iv) firstly by segmenting the initial slice in an unsupervised way, secondly by using the segmented image as the target in the (supervised) training of the classifier devoted to the initial slice, and finally by segmenting each slice in a supervised way with its associated classifier whose weights are obtained via training the classifier of the preceding slice which uses the segmented image of that slice as the target. (Such an approach makes the design of the overall classification system fully unsupervised that depends on the given CTA series only without requiring any given training set of CTA series. This is a very interesting feature of the overall classification system preventing the generalization errors originated from the dependence of the classifiers' performance on the used training set of CTA series.)

Results show that several problems in liver segmentation are addressed including gray level value similarity of adjacent organs, partial volume effects, atypical liver shapes and different modality settings. The method's ability of adaptation to data set characteristics increases the tolerance capability of the system and makes it feasible for clinical usage.

Although some of them produce very effective results in CT series, the deformable model based and gray level value based techniques generally produce segmentation results with holes inside the liver volume in CTA series even when the outer border is found correctly, because of the fact that the internal structure of the liver is acquired heterogeneous due to contrast media injection in CTA series. This results with incorrect measurement of the liver volume which is also handled properly with the proposed method.

The proposed algorithm has also been applied to the data sets provided in [25]. Although these data sets are acquired with CT, the series obtained for patient data sets with no tumors are segmented with high performance. This result also shows the benefit of patient oriented approach which is affected minimally from modality settings and does not need a training set prior to the application. In the data sets with tumors, it is observed that the segmented area generally includes the liver without including the tumor area. This is an expected result since the proposed algorithm is designed to segment healthy liver parenchyma for the evaluation of transplantation donors who should not have any tumors in their liver

Together with the data sets provided in [25,26], the proposed method has been tested with the CT and CTA series acquired from four different modalities. The successful results obtained by all these modalities also show that the proposed method does not have dependence on the modality.

The disadvantage of the proposed method is its dependency to the correct segmentation of the 'initial image'. If the automatically selected initial image does not satisfy the necessary requirements, the user should select an appropriate slice or might need to manually segment the initial image.

Since the algorithm is developed for the pre-evaluation of the transplantation donor candidates, the series that are acquired with rotated patient position (i.e. data sets 1 and 2 in [25]) cannot be segmented.

Areas that may further be examined include speeding up the process by improving the programming structure, more robust detection of the 'initial image' and the 'initial kidney image'.

6. Summary

Our regular article submission consists of the authors' original work about the implementation of an algorithm for automatic segmentation of liver in contrast enhanced CT images. Due to gray level similarity of adjacent organs, injection of contrast media and partial volume effects; robust segmentation of the liver is a very difficult task. Moreover, high variations in liver position, different image characteristics of different CT modalities and atypical liver shapes make the segmentation process even harder.

Our strategy for overcoming these difficulties involves a segmentation method which does not utilize a common parameter set found from all patient data sets. Instead, the method is capable of adapting the parameter set to each patient. The main reason for this approach is that the ranges of the parameter values differ significantly from patient to patient, and these wide ranges decrease the efficiency of the method when one utilizes a common parameter set for all patients.

Thus, we propose a method which examines and adapts its parameters according to each patient. We call this approach as *patient-oriented* segmentation. For qualifying 'patient oriented', the algorithm learns dataset characteristics in parallel to segmentation process, and adapts its parameters to these characteristics. To our knowledge, there is no method in the literature that works in this manner and at the same time addresses all the challenging aspects mentioned above.

Our iterative segmentation algorithm combines classification of pixels (using an unsupervised clustering method i.e. K-Means) with adjacent slice information (obtained by skeletonization) via morphological reconstruction. A more complex classifier (multi-layer perceptron network—MLP) is developed for the data sets where the K-Means clustering gives insufficient results. Here, the neural network is designed to use features extracted from the current and adjacent (previously segmented) slices and therefore intrinsically robust to gray level and shape variations. The decision between using either K-Means or MLP is also done automatically by the algorithm. The developed algorithm gives sufficient performance for different modalities, varying contrast, dissected liver regions and atypical liver shapes. Results indicate that we have effectively overcome the challenging difficulties explained before and our algorithm is clinically feasible.

Acknowledgments

This work is supported by TUBITAK with Grant number 104E178. The authors would like to thank Ömer KARAL and Selen ŞAYLISOY for their support on this study. The authors would like to thank the anonymous reviewer for valuable critics.

Conflict of interest statement

None declared.

References

[1] G. Flohr, S. Schaller, K. Stierstorfer, H. Bruder, B.M. Ohnesorge, U.J. Schoepf, Multi detector row CT systems and image reconstruction techniques, Radiology 235 (2005) 756–773.

[2] D.R. Ney, E.K. Fishman, D. Magid, R.A. Drebin, Volumetric rendering of computed tomography data principles and techniques, IEEE Comput. Graphics Appl. 10 (2) (1990) 24–32.

[3] K.T. Bae, M.L. Giger, C.T. Chen, C.E. Kahn Jr., Automatic segmentation of liver structure in CT images, Med. Phys. 20 (1993) 71–78.

[4] L. Gao, D.G. Heath, B.S. Kuszyk, E.K. Fishman, Automatic liver segmentation technique for three-dimensional visualization of CT data, Radiology 201 (1996) 359–364.

[5] S.J. Lim, Y.Y. Jeong, Y.S. Ho, Automatic liver segmentation for volume measurement in CT images, J. Vis. Commun. Image R. 17 (2006) 860–875.

[6] J. Masumoto, M. Hori, Y. Sato, T. Murakami, T. Johkoh, H. Nakamura, S. Tamura, Automated liver segmentation using multislice CT images, Systems Comput. Jpn. 34 (9) (2003) 2150–2161.

[7] J. Montagnat, H. Delingette, Volumetric medical image segmentation using shape constrained deformable models, CVRMed-MRCA, Lecture Noes in Computer Science, vol. 1205, Springer Verlag Publisher, Berlin, 1996, pp. 13–22.

[8] J.S. Chou, S.Y. Chen, G.S. Sudakoff, K.R. Hoffmann, C.T. Chen, A.H. Dachman, Image fusion for visualization of hepatic vasculature and tumors, Medical imaging 1995: image processing, SPIE Proc. 2434 (1995) 157–163.

[9] L. Soler, et al., Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery, Acad. Radiol. 12 (2005) 1178–1189.

[10] J. Gao, A. Kosaka, A. Kak, A deformable model for automatic CT liver extraction, Medical imaging 2000: image processing, SPIE Proc. 2434 (2000) 157–163.

[11] T. Heimann, I. Wolf, H.P. Meinzer, Active shape models for a fully automated 3D segmentation of the liver—an evaluation on clinical data, in: Proceedings of MICCAI 2006, Lecture Notes in Computer Science, vol. 4191, 2006, pp. 41–48.

[12] T. Heimann, H.P. Meinzer, I. Wolf, A statistical deformable model for the segmentation of liver CT volumes, in: T. Heimann, M. Styner, B. van Ginneken (Eds.), MICCAI 2007 Workshop Proceedings: 3D Segmentation in the Clinic—A Grand Challenge, 2007, pp. 161–166.

[13] D. Tsai, Automatic segmentation of liver structure in CT images using a neural network, IEICE Trans. Fundamentals E77-A (11) (1994) 1892–1895.

[14] S.A. Husain, E. Shigeru, Use of neural networks for feature based recognition of liver region on CT images, in: Neural Networks for Signal Processing—Proceedings of the IEEE Workshop, vol. 2, 2000, pp. 831–840.

[15] C.C. Lee, P.C. Chung, H.M. Tsai, Identifying multiple abdominal organs from CT image series using a multimodule contextual neural network and spatial fuzzy rules, IEEE Trans. Inf. Technol. Biomed. 7 (3) (2003) 208–217.

[16] J.E. Koss, F.D. Newman, T.K. Johnson, D.L. Kirch, Abdominal organ segmentation using texture transforms and a hopfield neural network, IEEE Trans. Med. Imag. 18 (7) (1999) 640–648.

[17] K.S. Seo, L.C. Ludeman, S.J. Park, J.A. Park, Efficient liver segmentation based on the spine, ADVIS 2004, Lecture Notes in Computer Science, vol. 3261, Springer Verlag Publisher, Berlin, 2004, pp. 400–409.

[18] L. Bidaut, Data and Image Processing for abdominal imaging, Abdominal Imaging, vol. 25, Springer Verlag, Berlin, 2000 pp. 341–360.

[19] L. Vincent, Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, IEEE Trans. Image Process. 2 (2) (1993) 176–201.

[20] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematics Statistics and Probability, vol. 281, 1967.

[21] S. Haykin, Neural Networks: A Comprehensive Foundation, second ed., Prentice-Hall, Prentice, 1998 pp. 176-201, ISBN: 0132733501.

[22] H.A. Blum, Transformation for extracting new descriptors of shapes, in: W. Wathen-Dunn (Ed.), Models for the Perception of Speech and Visual Form, MIT Press, Cambridge, MA, 1967, pp. 362–380.

[23] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Systems Man Cybernet. SMC-9 (1) (1979) 62–66.

[24] F. Fischer, M.A. Selver, W. Hillen, A platform independent visualization software for DICOM series, in: Proceedings: IEEE EMBS Symposium: Innovative Medical Imaging Modalities Proceedings, 2005, p. 330.

[25] B. Van Ginneken, T. Heimann, M. Styner, 3D segmentation in the clinic: a grand challenge, in: T. Heimann, M. Styner, B. van Ginneken (Eds.), MICCAI 2007 Workshop Proceedings: 3D Segmentation in the Clinic—A Grand Challenge, 2007, pp. 7–15.

[26] A. Rosset, L. Spadola, O. Ratib, OsiriX: an open-source software for navigating in multidimensional DICOM images, J. Digit. Imaging 17 (3) (2004) 205–216.

M. Alper Selver (1980) graduated in Electrical and Electronics Engineering at Gazi University in 2002. He has been working as a Research Assistant at the Department of Electrical and Electronics Engineering in Dokuz Eylul University since 2002. He got his M.Sc. degree in 2005 and he is currently a Ph.D. student from the same department. His main research interest is in the field of radiological image processing, software development, artificial neural networks, and fuzzy computing. He is involved in several national projects on computer science in medicine.

Aykut Kocaoğlu was born in Usak, Turkey, in 1982. He received B.Sc. degree from Electrical and Electronics Engineering Department of Dokuz Eylul University, Turkey, in 2005. Since 2005, has been with Dokuz Eylul University, Turkey, where he is a Researcher and Graduate Student at the Department of Electrical and Electronics Engineering. His main research interests are in the field of image processing, software development, artificial neural networks, and vehicle tracking.

Güleser K. Demir received her M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Dokuz Eylul University (DEU), Izmir, Turkey, in 1996 and 2001, respectively. She worked in the Department of Computer Science and Engineering at University of Minnesota as a Visiting Researcher from January 2002 to August 2003. She is currently an Assistant Professor in the Department of Electrical and Electronics Engineering at the Dokuz Eylul University. Her research interests include classification algorithms, neural Networks and image processing.

Hatice Doğan received the B.Sc., M.Sc., and Ph.D. degrees in Electrical and Electronics Engineering from Dokuz Eylul University, Izmir Turkey, in 1996, 1999 and 2005, respectively. She has been a research and teaching assistant at Engineering Faculty of Dokuz Eylul University since 1996. Her research interests include learning theory, neural networks and biomedical signal processing.

Oğuz Dicle was born in 1959. He received the M.D. degree, in 1982, from Dokuz Eylul University and became a Professor of Radiology in 2000. Since 2001 he is the Head of Radiology Department of Dokuz Eylul University School of Medicine. His main interest in radiology is interventional and abdominal radiology. He principally involved in many pioneering projects such as the first MR imaging center and the first Digital Radiology Department in Turkey. He is also the Chairman of Medical Informatics Department in the same University. He was the Vice Dean of the Medical School before he started his current role as the Coordinator of the health campus. Dr. Dicle has more than 35 international scientific papers and has two books dedicated to radiology physics and Turkish Radiology History.

Cüneyt Güzeliş received the B.Sc., M.Sc., and Ph.D. degrees in Electrical Engineering from Istanbul Technical University, Istanbul, Turkey, in 1981, 1984, and 1988, respectively. He had worked in Istanbul Technical University from 1982 to 2000. He worked between 1989 and 1991 also in the Department of Electrical and Computer Engineering at University of California, Berkeley, CA, as visiting researcher and lecturer. He is now a professor in the electrical and electronics engineering department, Dokuz Eylül University, Izmir, Turkey. His interest areas include artificial neural networks, biomedical signal processing, nonlinear circuits, systems and control, and educational systems.