A comparative performance evaluation of various approaches for liver segmentation from SPIR images

Evgin GÖÇERİ, Mehmet Zübeyir ÜNLÜ, Oğuz DICLE

1Electrical and Electronics Engineering Department, Izmir Institute of Technology, Urla, İzmir, Turkey
2Faculty of Medicine, Department of Radiology, Dokuz Eylül University, Narlıdere, İzmir, Turkey

Received: 04.04.2013 • Accepted/Published Online: 24.05.2013 • Printed: 30.04.2015

Abstract: Developing a robust method for liver segmentation from magnetic resonance images is a challenging task because of the similar intensity values between adjacent organs, the geometrically complex liver structure, and injection of contrast media. Most importantly, a high anatomical variability of a healthy or diseased liver is a major challenge in defining the exact boundaries of the liver. Several artifacts of pulsation, motion, and partial volume effects are also among the variety of factors that make automatic liver segmentation difficult. In this paper, we present an overview of liver segmentation methods in magnetic resonance images and show comparative results of seven different pseudo-3D liver segmentation approaches chosen from deterministic (K-means-based), probabilistic (Gaussian model-based), supervised neural network (multilayer perceptron-based), and deformable model-based (level set) segmentation methods. The results of quantitative and qualitative analyses using sensitivity, specificity, and accuracy metrics show that the multilayer perceptron-based approach and a level set-based approach, both of which use distance regularization terms and signed pressure force function, are the most successful methods for liver segmentation from spectral presaturation inversion recovery (SPIR) images. However, the multilayer perceptron-based segmentation method has a higher computational cost. The automatic method using the distance regularized level set evolution with signed pressure force function avoids the sensitivity of a user-defined initial contour for each slice, gives the most efficient results for liver segmentation after the preprocessing steps, and also requires less computational time.

Key words: Gaussian mixture model, k-means, level set, magnetic resonance image, multilayer perceptron, liver segmentation

1. Introduction

The manual liver segmentation task is not only time consuming and tedious due to the high number of slices, but also depends on skill and experience. Figures 1a, 1b, and 1c show liver edges identified differently by different radiologists, and also by the same radiologist at a different time. Therefore, despite the problematic nature of organ segmentation, automated segmentation methods are needed, especially for organs like the liver.

There are many different image segmentation techniques and procedures [1–7]. The needs and developments in this area are increasing day by day [8]. In particular, medical image segmentation is one of the most prominent topics currently studied since approaches to segmentation of medical images often have many more difficulties than other approaches. These difficulties arise from noise, low image contrast, intensity inhomogeneity, and also missing, smeared, or nonclear edges of tissues in the images, which can be due to patient movements.

*Correspondence: evgingoceri@yahoo.com
and/or a poor signal-to-noise ratio of the image acquisition devices. The blending of similar adjacent tissues enhances these challenges further. In particular, both accurate and robust liver segmentation is a challenging problem since the gray level values of the liver and all other tissues are different in different datasets, and even in different slices of a dataset because of the injection of contrast media and the existence of different image modality settings. In addition, the liver has varying anatomical shapes in different slices and has gray level values that are similar in intensity to its adjacent organs. Accordingly, the commonly used thresholding methods and shape-based automatic liver segmentation methods may not be sufficient for the purpose of automatic liver segmentation.

Figure 1. Liver edges: a) defined by radiologist R1 b) defined by radiologist R2 c) defined by radiologist R2 at a different time.

Although magnetic resonance (MR) imaging has several advantages such as efficient soft tissue contrast, free-form ionizing radiation, and multiplanar capabilities, automated segmentation of MR images is more challenging than computer tomography (CT) images. For example, MR images have small edge magnitudes compared to CT images, which causes edge-based segmentation algorithms to be more complicated. Another difficulty relates to motion artifacts and partial volume effects, which lead to more challenges in automatic liver segmentation. Consequently, in comparison with CT-based liver segmentation approaches [7,9–31], there are fewer studies for liver image segmentation from MR datasets. The present approaches for MR-based image segmentation in the literature can be listed as fuzzy c-means classification [32,33], graph-cut [34], snakes [35], the level set method [36,37], the synchronized oscillator network [38], the active shape model [39,40], watershed [41], iterative quadtree decomposition [42], the Gaussian model and Markov random field [43], modified region growing [44], and the application of free-form registration on manually segmented CT images [45]. At present, it is clear that there is no method capable of simultaneously solving all of the problems of different modality characteristics, atypical liver shapes, and similar gray values with adjacent tissues. Example SPIR images, which demonstrate why liver segmentation is a difficult task because of gray level similarities between adjacent tissues and organs, are shown in Figures 2a and 2b. Apart from all the methods described above, there are also hybrid approaches that consider the segmentation of the liver as a small part of the bigger and more difficult multiorgan segmentation problem. For instance, in [46], the authors combine the region-based and shape-based methods as a hybrid approach to segmenting the liver from CT images, along with kidneys and a spleen. The method was shown to be effective in MR organ segmentation as well. In [47], the authors show the use of the active appearance model within the scope of the graph cut and shape model. Similarly, in [48], the authors combine the graph cut algorithm and active shape modeling to segment multiple organs. Although these hybrid methods were shown to be effective in segmenting large objects, because of imaging modality differences and the
notion of individual organ segmentation in our case using a specific MR sequence for the liver, a full evaluation and comparison to the methods in this work is outside the scope of their paper.

Figure 2. Example images to show why liver segmentation from MR images is a challenging process: gray level similarities a) between liver and stomach b) between liver and heart.

In the present literature, most methods developed for automatic liver segmentation from MR images have either over- or undersegmentation or leakage problems [42–44], are tested with only a few datasets [34,44], or have complex calculations such as active contour-based approaches [45]. In [41], the watershed transformation and neural networks are used for liver detection without identifying the modality characteristics of the MR images used. Moreover, it should be noted that the T2-weighted SPIR datasets have not been used for liver segmentation until now, despite their ability to better visualize liver vessels. For all the reasons mentioned, and due to a lack of research in this area, an accurate, efficient, and robust automatic liver segmentation method from MR sequences is needed.

In this work, a probabilistic approach based on the Gaussian mixture model (GMM), a deterministic approach (K-means-based segmentation), a supervised learning method (multilayer perceptron [MLP]-based segmentation), and some recently published two level set methods and two automatic segmentation approaches based on these level set methods are applied to MR liver segmentation by explaining the properties of these techniques as well as comparing and evaluating the results. It is observed from the experimental results and quantitative performance analysis that the automatic segmentation algorithm based on the distance-regularized level set evolution (DRLSE) with the signed pressure force (SPF) function method [49] presents the best segmented liver images from SPIR image datasets.

The organization of the remaining parts of the paper is as follows. The properties of the MR image datasets used in this study are explained in Section 2. The GMM-based, K-means-based, and MLP-based liver segmentation methods are explained in Sections 3, 4, and 5, respectively. Active contour-based segmentation methods with example results for liver segmentation are presented in Section 6. The comparative results of the methods applied using our datasets are shown and conclusions are given in Section 7. Quantitative performance analyses of acceptable results using four different similarity metrics are given in Section 8. Finally, in Section 9, discussions and future works are explained.

2. Datasets
In this study, upper abdominal MR datasets are used. The datasets were obtained from ten different patients (four men and six women; age range 52–81 years) using a 1.5 Tesla MR imaging device (Gyroscan Intera, Philips, ACS-NT, Best, the Netherlands) located in Dokuz Eylül University (İzmir, Turkey) Radiology Department. The
examined 16-bit DICOM images are fat suppressed T2-weighted (TR/TE, 1600/70 ms; flip angle, 90°; slice thickness, 8 mm) SPIR images (0.63 pixels per mm and the pixel size is 1.58 × 1.58 mm) in the axial plane with a resolution of 256 × 256.

The reason for choosing T2-weighted SPIR images is that they are better able to visualize liver vessels. As an extension of this work, the segmented liver images will be used to label hepatic and portal veins, and then to obtain Couinaud segments. Therefore, the first aim is to obtain segmented liver images from SPIR datasets that show a healthy liver.

3. Liver segmentation with GMM

The GMM [50–52] is a probabilistic and unsupervised classification. Probability density functions of images are estimated during the training phase of the classifier. The assumption of this classifier for pixel-labeling is that the gray level value for each pixel of the observed image is a sample from a finite mixture distribution.

The applied iterative liver segmentation algorithm using GMM from MR images in this section is similar to the kidney segmentation using K-means clustering from CT angiography images given in [25]. The first step is to select a slice in which the border of the liver is very clear in the dataset. It is usually the middle slice of the datasets, which comes just after the first disappearance of the right kidney. In general, the liver boundaries do not overlap in this slice with other organ boundaries such as the heart or the right kidney. The selected slice is used as the initial slice to be segmented (Figure 3).

The second step is to use spine location as a landmark to find the liver location in the selected initial slice since the liver is inside the ribs and located on the left-hand side of the selected slice. The spine will be in the bright cluster after the initial slice is clustered into three clusters, which are background, dark gray tissues (skin), and the bright gray tissues (organs) (Figure 4a) by using GMM with expectation maximization [53]. The location of the spine is inside the ribs and in the middle of the bottom side of the abdominal image. To find the spine position, anatomical information is used. Therefore, by searching from left to right and from top to bottom in the clustered image, the first nonzero values are found. Afterwards, a rectangular frame is obtained using these values (Figure 4b). The point of having the first nonzero value from the bottom is that it helps to find the row of the spine. The middle point of this row is used to draw another frame through the spine (Figure 4c). At the right side of the spine, a seed region that includes some part of the liver is selected after detecting the spine location. Finally, image reconstruction [54] is performed to obtain the liver image using the selected seed image on the liver as a marker, and the clustered image as a mask (Figure 4a). The segmented initial liver
image is obtained after morphological filling and closing operations (Figure 4d).

![Figure 4. a) Initial slice clustered image b) rectangular frame around the image c) spine in the blue frame and the selected seed region in the red frame d) reconstructed liver region after filling and closing operations.](image)

Figure 5a shows the segmented grayscale initial liver image to use as a reference image for the first succeeding (Figure 5b) and preceding slices (Figure 5c).

![Figure 5. a) Initial liver image b) the next slice c) the previous slice to be processed.](image)

The succeeding slices of the initial liver slice also are clustered into three clusters. The spleen disappears and the heart seems darker, while the kidneys appear in each preceding slice of the initial liver slice. This leads to changes in the distribution of gray level and to the addition of a new cluster, which shows the brightest regions. Therefore, the preceding slices are clustered into the four clusters. The tissues of the liver are in the second and third clusters in the preceding and succeeding slices.

The skeleton (Figure 6a) of the segmented initial liver image (Figure 5a) is used as a marker for both the next (Figure 5b) and previous slices (Figure 5c). The required masks for each slice to be segmented are the clusters that include liver tissues in the GMM (Figure 6b). The image reconstruction process is applied with the binary mask and marker images to obtain the reconstructed liver image. From this, the segmented grayscale liver image (Figure 6c) is obtained. The marker image for each previous slice is the skeleton image (Figure 6d) of the segmented previous liver. The necessary marker image for each succeeding slice, which is the skeleton of the next liver image, is used with the new mask image (Figure 6e) for image reconstruction.

All remaining slices in the dataset are segmented iteratively by starting from the initial liver image to the end of the dataset and from the initial liver image to the beginning of the dataset. When the borders are unclear between the spleen and liver, or when there is an atypically shaped liver, in order to decide where the liver border is, the radiologist examines the other slices that come just after or just before where the borders are more visible. The result of the GMM-based method for the first succeeding slice of the initial liver image in Figure 6c shows that the probabilistic approach is not successful at segmenting the liver from SPIR datasets.
Figure 6. a) Skeleton of the initial liver image (Figure 5a) to use as initial marker b) the result of GMM-based clustering of the next slice (Figure 5b) to use as a mask c) reconstructed grayscale liver image using the mask and marker images d) the skeleton image as a new marker for the next slice e) the next clustered image.

4. Liver segmentation with K-means

The K-means method [55] is commonly known as a deterministic approach. It is an unsupervised clustering method used in many applications [56–59]. K-means minimizes the sum of distances from each data item and measures the distance from each data item to its cluster center. The computation of cluster centers is performed for each cluster after assigning all data items to their closest centers repeatedly. The main known distances are squared Euclidean distance, Hamming distance, and cosine dissimilarity. Generally, the squared Euclidean distance metric is used as a distance measure for K-means. The sum of squared error function is used as the criterion function to be optimized in a Euclidean K-means method [60].

We have applied the same liver segmentation algorithm explained in Section 3 using the original K-means method [61] instead of GMM-based segmentation. In our application, Euclidean distance with batch update [62] is implemented. The gray values are assigned to their closest cluster centers at the same time, and then
recomputation is performed to reassign the gray values to each cluster center at each iteration. Figure 7 shows an example result obtained by this method for the same slice shown in Figure 5b. Figures 7a and 7b are the reconstructed binary liver image, and the grayscale liver image corresponds to Figures 6c and 6d, respectively. It was observed that the K-means-based iterative algorithm is not successful in segmenting liver organs from all SPIR slices.

![Figure 7. a) Reconstructed liver image b) grayscale segmented image.](image)

5. Liver segmentation with MLP

Liver segmentation from SPIR datasets is performed using neural networks, i.e. MLP, in this section. Similar to the MLP-based liver segmentation algorithm used for CT images in [25], we applied preprocessing steps to remove irrelevant organs (e.g., kidneys and spine) from our abdominal MR image datasets and used three features of the preprocessed images. However, we have fitted the probability distribution of the GMM to image intensity histograms instead of using the K-means method for the clustering process when we apply the proposed algorithm. Moreover unlike CT images, SPIR images show gallbladders as bright white organs (Figure 8a), which affects the MLP-based segmentation results. Therefore, we have segmented gallbladders in the preprocessing stage.

In the first step, the spine position is detected to find the kidney locations in the preprocessing stage. Kidney segmentation is performed using GMM following the same steps shown in Figure 6 and presented in [63]. In this case, an initial kidney slice is selected instead of an initial liver slice, and kidney segmentation is applied iteratively. The next step is spleen segmentation. It is performed similarly to the kidney segmentation algorithm by selecting a seed region on the spleen instead of a kidney. From this, we obtain the preprocessed images without the spine, spleen, or kidneys (Figure 8b).

Gallbladders may have a different shape, size, and position in a human body. They may not appear in each dataset. However, they are adjacent to the liver and affect the MLP-based liver segmentation results if they exist in an abdominal SPIR dataset. Therefore, gallbladders are extracted in the preprocessing stage using the image that has no kidney, spine, or spleen (Figure 8b). The slices that have gallbladders (Figure 8a) are among the slices that come before the selected initial slice. A preprocessed image with a gallbladder (Figure 8b) is clustered into seven clusters, which are background, dark gray tissue (skin), four different gray levels, and the brightest white tissues using GMM-based segmentation. The gallbladder is in the last cluster and is extracted easily since the gray value of the liver, which is the remaining organ after the preprocessing steps, is very different from that of the gallbladder. Afterwards, the obtained preprocessed image (Figure 8c) is used for MLP-based liver segmentation.
An initial slice is selected after the preprocessing steps to obtain an initial liver image and to use as the first reference image for the segmentation of the next and previous slice. The selected preprocessed initial slice (Figure 9a), where the liver is larger than the liver areas in other slices, is clustered by applying an optimal threshold with Otsu's method [64], which gives a successful result only for the selected initial slice. The cluster that contains the liver (Figure 9b) is selected from the three clusters, which are background, dark regions (skin), and bright regions (liver) after classification. Then binary morphological operations are applied. The first operation is median filtering, which removes the small white spots from the background and the black spots from the liver region (Figure 9c). The next operation is erosion to break off the small connections (Figure 9d), and a connected component analysis to extract the biggest part of the image. Finally, a dilation process is applied to restore the completely undeleted objects by erosion (Figure 9e), giving the segmented initial liver image (Figure 9f).

The three features that are used are the mean to represent the homogeneous regions, standard deviation to represent the edges, and the distance transform process to measure the separation of pixels in the image in order to provide information about liver position. The mean and standard deviation features are computed using the preprocessed initial image in the initial training step. The Euclidean distance transform feature is
obtained using the same initial image after its clustering. We chose to use probabilistic GMM-based clustering instead of K-means clustering in this step. These three features are used as training data and the clustered initial liver image is used as the target output in the initial training step to obtain the initial weight values.

All slices are segmented iteratively after the initial training step by updating weights. The new current image, which is the next preprocessed image to be segmented, is used to calculate the new values of the mean and standard deviation. The third feature (distance transform) is computed using the previous segmented slice. These three features and the weights that are obtained from the previous image are used for image segmentation by applying MLP. The new weight values to use in the next segmentation are computed by training the network. The three features are computed again for this training step with the resulting image and used as input. Furthermore, the resulting image is used as the desired target image for this training. The segmentation procedure proceeds using the next image after the computation of the new weights. Using the previous weights for the next training step requires significantly less training time. Therefore, this iterative process increases segmentation performance.

6. Liver segmentation with level set methods

Active contours have great importance and have become a well-established technique in the area of image segmentation [1,2,4,5,65–72]. A review can be found in [73]. The reason for the extensive research on active contours is that they can generate closed curves using images and achieve subpixel accuracy. They also provide robustness for spurious edges by incorporating smoothness constraints.

A geometric active contour approach is proposed in two independent studies [2,68]. The active contour methods in this category are based on the curve evolution theory [74] and the level set technique [65]. According to the curve evolution theory, the deformation of the curve is written by a partial differential equation (PDE). The importance of PDEs in image segmentation has been increased since the first application of level set methods as image segmentation at the beginning of the 1990s. The main reason for this is that PDEs can transform a problem of segmentation modeling into a PDE framework. Moreover, PDEs can use regularizers with these segmentation models. Another reason is that finite difference methods can be used for PDE solutions. It is also possible to extend the PDEs from two dimensions to higher dimensions. In addition to these advantages, the solutions from PDEs are fast, and they are able to perform an image segmentation operation interactively.

Although the original level set method does not contain any energy terms, level set equations have been applied using an energy minimization term of the contour to handle topological changes and obtain accurate results in some improved level set methods [1,5,67,70–72,75–81]. The energy minimization in level set methods is applied like an edge detection method. In these alternative approaches, an energy functional equation that will be minimized is defined on the level set function and an evolution PDE is obtained using this energy functional instead of a Eulerian equation, as in classical level set methods [1,2,68,82]. For example, for the variational level set method in [80], minimization of the contour energy is performed by optimizing a constraint that indicates the interior homogeneity degree of a contour. A constraint optimization method in [77] uses two selected points on the object boundary and computes the minimum global energy for a contour between these points. The supervised segmentation method in [76] needs prior information for the object pattern to do energy minimization. Shape-based prior information has been integrated into level set equations for accurate segmentation in [75,78,79]. Only region-based features are used in [78] by assuming the existence of only the object and the background of the image to construct the energy in order to minimize them. However, the global information used in the proposed formulation is not accurate when there is intensity inhomogeneity inside or
outside of the contour. Similarly, a piecewise constant model-based level set method in [6] cannot handle the intensity inhomogeneity problem. Local intensity information is incorporated into the level set method as a solution for intensity inhomogeneity [70–72]. However, only mean values of local intensities are not enough for an efficient segmentation when there is severe noise and intensity inhomogeneity. Therefore, both mean and variance values of local intensities, which are obtained by Gaussian distributions, are used to deal with intensity inhomogeneity [5]. Although this approach has not been tried for liver images, it gives accurate results for brain images, but creates a high computational cost.

In sections 6.1 and 6.3, we review two different level set-based image segmentation methods that have been proposed recently, and present their implementation results for our abdominal MR image datasets. We explain automatic segmentation methods using these two methods (which do not require user-defined initial contours for each slice) in sections 6.2 and 6.4, respectively.

6.1. Level set method without using PDE

Narrow-banded level set methods [83,84] have been proposed to increase segmentation performance with a reduced computation cost. In these methods, PDEs are solved only around the neighborhood of the zero level set instead of in the whole image to increase the evolution speed of the curve. A narrow-banded level set method with a bandwidth of two pixels is proposed in [85]. The authors propose the curve evolution with the fast two cycles (FTC) algorithm, which approximates the level set technique using integer operations without PDEs to reduce the computation cost. An example result of the FTC method using an original image (Figure 10a) and an initial contour (Figure 10b) is shown in Figure 10c, which has a misclassified region (Figure 10d).

![Figure 10. a) Original image b) initial contour c) result of the FTC method d) misclassified region.](image)

6.2. Automatic liver segmentation with the FTC algorithm

Initial contours can be propagated to outside or inside. The position of initial contours is a key challenge for level set-based segmentation methods [86]. Each user can draw different initial contours at different locations and sizes. Finding the correct number of iterations for each initial contour is not easy. The segmentation results are not robust even when the iteration numbers are given by users, or a constant number of iterations is used due to different initial contours defined by users. Therefore, we need to apply level set approaches for liver segmentation automatically to overcome these drawbacks and obtain acceptable results from all slices in a dataset.

The liver shape in a slice is close to the liver shape of the preceding or succeeding slices with respect to other slices in the dataset. Therefore, the shape of an initial segmented liver image can be used as the initial contour for the next and previous slices. In our previous work [87], we obtained the initial contour automatically...
for each slice using this relation between sequential slices in a dataset and used the FTC algorithm without any user interaction for liver segmentation. However, we observed that the segmentation results are not robust. Therefore, in this section, we have applied the same algorithm after the preprocessing steps as explained in Section 5 to increase segmentation performance.

The first step of this automatic segmentation algorithm is to obtain an initial liver image (Figure 9f). Then the shape of the initial liver image is used as the initial contour for liver segmentation from the first preceding and succeeding slices. Finally, binary morphological postprocessing steps, which are erosion to remove weakly connected objects, connected component analysis to get the biggest part, and dilation to restore the completely undeleted objects by erosion, are applied. All remaining slices are segmented iteratively with the same number of iterations, which is a very small number. We have used only ten iterations for each slice since the initial contour is very close to the desired liver edges. Therefore, the required time for the segmentation of each slice is reduced and better segmentation results are obtained automatically from all slices in the dataset. Example experimental images are shown in Figure 11 with two consecutive slices. The preprocessed image given in Figure 11a is used for the automatic level set-based segmentation with the initial contours (shapes of the previously segmented liver image) shown in Figure 11b, and the image given in Figure 11c is obtained. The segmented liver image after postprocessing is presented in Figure 11d. Similarly, Figure 11e shows the next preprocessed image to be segmented with the initial contour given in Figure 11f. The applied automatic level set-based segmentation result is shown in Figure 11g and the image after postprocessing is given in Figure 11h.

Figure 11. a) and e) Preprocessed images b) and f) shapes of the previous liver images to use as initial contours c) and g) results of the automatic segmentation with the FTC algorithm d) and h) segmented liver images after postprocessing.

6.3. An edge-based level set method with a SPF function

A recently published edge-based level set method proposes to use both local and global image information for liver segmentation [49]. Therefore, the authors use a PDE that includes a distance regularization term to solve the reinitialization problem of the level set function, and a SPF function to control the level set evolution using the average values of the intensities both inside and outside of the active contour.

Figure 12a shows an initial contour on the original image. Figure 12b shows that although the abdominal
SPIR images have inhomogeneous intensities and gray level similarities with the liver and heart, DRLSE with the proposed SPF function gives a successful result.

![Figure 12. a) Initial contour on the original image b) result of the DRLSE with SPF function method.](image)

Our aim is to successfully obtain segmented liver images not only from a given slice, but also from all slices in the dataset. However, a small number of iterations defined by users for each slice affects the accuracy of the results, while a high number of iterations increases the computation time. Therefore, we apply an automatic segmentation algorithm using the DRLSE with a SPF function method and examine the results from different images in the next section.

### 6.4. Automatic liver segmentation using the DRLSE with SPF function

Similar to the automatic liver segmentation algorithm explained in Section 6.2, the initial liver image shown in Figure 9f is used to obtain the initial contour from its shape for the previous and next slices to be processed. Liver segmentation from all remaining slices in the dataset is performed iteratively using the initial contour obtained from succeeding slices. Example results of this iterative application using the original images shown in Figures 13a–13c with initial contours are given in Figures 13d–13f respectively. As can be seen, we can obtain successfully segmented liver images (Figures 13d and 13e) without applying any postprocessing steps. However, if there is a gallbladder in the original image, then the iterative application does not give an acceptable result (Figure 13f).

The results obtained from the preprocessed images and initial contours shown in Figures 14a, 14b, and 14c are given in Figures 14d, 14e, and 14f, respectively, after implementation of the automatic segmentation algorithm. As observed from these results, liver segmentation from preprocessed images gives correctly segmented livers even if the original slices have gallbladders.

### 7. Experimental results

Experiments on original and preprocessed images showed that fitting the probability distribution of GMM to intensity histograms gives better results than the deterministic K-means method, which is affected by noise more than GMM. However, because of the similar gray level texture of the liver and adjacent organs such as the spleen and heart, even though they used skeleton information from the previous slice, the results obtained from the GMM and K-means-based segmentation methods are not correctly segmented liver images. It is not possible to obtain the desired liver shapes from these results by applying the same postprocessing operations for all slices.
The results of the supervised MLP-based segmentation method without any preprocessing and postprocessing step are not acceptable.

The integer operations-based level set method (FTC method) is sensitive to user-defined parameters, which are iteration numbers and initial contours on each slice.

The automatically applied method based on DRLSE with SPF function presents liver images segmented...
more successfully without applying any postprocessing steps. However, the results of the automatic liver segmentation algorithm using this method as explained in Section 6.4 are not acceptable when we use original slices that have kidneys and gallbladders (Figure 13f). Therefore, we applied the automatic segmentation algorithm using this method on preprocessed images.

Figure 15 shows the results obtained by using the original slices given in Figures 15a–15c in which kidneys do not appear with the automatic GMM (Figures 15d–15f), automatic K-means (Figures 15g–15i), MLP (Figures 15j–15l), the FTC method (Figures 15m–15o), the FTC algorithm-based automatic method (Figures 15p–15s), DRLSE with SPF function (Figures 15t–15v), and also the automatic DRLSE with SPF function-based liver segmentation method (Figure 15x–15z).

Figure 15. a) b) c) Original slices d) e) f) results of the automatic GMM g) h) i) automatic K-means
Figures 16a–16c show the preprocessed images of the original slices shown in Figures 15a–15c. The liver segmentation results from these preprocessed images using the automatic GMM (Figures 16d–16f), automatic K-means (Figures 16g–16i), MLP (Figures 16j–16l), the FTC method (Figures 16m–16o), the automatically applied FTC algorithm-based method (Figures 16p–16s), the DRLSE with SPF function (Figures 16t–16v), and also the automatic DRLSE with SPF function-based liver segmentation method (Figures 16x–16z) are shown in Figure 16. It is observed that the results of these automatic methods are more efficient when we use preprocessed images instead of the original slices for liver segmentation.
Figure 16. a) b) c) Preprocessed slices d) e) f) results of the automatic GMM g) h) i) automatic K-means j) k) l) MLP m) n) o) the FTC method.
Figures 17a–17c show original slices that have kidneys and gallbladders. The results obtained from these slices using automatic GMM (Figures 17d–f), automatic K-means (Figures 17g–i), MLP (Figures 17j–l), the FTC method (Figures 17m–o), the automatically-applied FTC-based method (Figure 17p–s), the DRLSE with SPF function method (Figures 17t–v) and also the automatic DRLSE with SPF function-based liver segmentation method (Figures 17x–z) are shown in Figure 17.

Figures 18a–18c show preprocessed images of the slices given in Figures 17a–17c. The liver segmentation results from these preprocessed images using the automatic GMM (Figures 18d–18f), automatic K-means (Figures 18g–18i), MLP (Figures 18j–18l), the FTC method (Figures 18m–18o), the automatic FTC algorithm-based method (Figures 18p–18s), the DRLSE with SPF function method (Figures 18t–18v), and also the automatic DRLSE with SPF function-based liver segmentation method (Figures 18x–18z) are presented in Figure 18.

It is observed from the results (Figures 15 and 17) that the methods presented in this paper may not work efficiently without applying preprocessing steps. This is because the gray level values of the liver, which are very similar to the gray level values of other abdominal organs in SPIR images, are inhomogeneous. Moreover, the existence of the gallbladder, which appears white on SPIR images and has different shape, size, and position, affects the intensity-based liver segmentation methods presented in this paper.
Figure 17. a) b) c) original slices d) e) f) results of the automatic GMM g) h) i) K-means j) k) l) MLP m) n) o) the FTC method.
We have shown in the results (Figures 16 and 18) that some of the methods presented in this paper might be unsuccessful for liver segmentation in some cases, even if we apply preprocessing steps, because the liver has a vascular structure and the vessels on the liver have a bright white color. Therefore, the thickness of the vessels and inhomogeneous intensity values of the liver affect segmentation performances. Low image contrast of abdominal SPIR images is another important factor that affects the efficiency of liver segmentation methods. In addition, if user-defined initial contours and iteration numbers for the FTC algorithm and the DRLSE with SPF function method are not chosen appropriately, segmented results may not be efficient.

We have presented four conclusions in this section. First, it is observed that K-means and GMM-based segmentation approaches do not give acceptable results for liver segmentation from SPIR image datasets, even if preprocessed images are used. Second, the results of the MLP-based method are not successful when the preprocessing steps are not applied. Third, the automatic segmentation technique using the FTC algorithm represents desired liver images from the preprocessed slices. However, this method is not successful for the slices that have kidneys or gallbladders in their originals. Finally, only the automatically applied DRLSE with SPF function-based method does not need any preprocessing and gives correctly segmented liver images, except in the slices which have a gallbladder. Quantitative performance analyses have been performed to compare the results obtained from these approaches according to their segmentation efficiency and computation time.
Figure 18. a) b) c) Preprocessed slices d) e) f) results of the automatic GMM g) h) i) K-means j) k) l) MLP m) n) o) FTC method.
Performance evaluations of these methods with sensitivity, specificity, and accuracy measures are presented in the next section.

8. Performance analysis

There are several performance measures to evaluate methods in the literature. Generally sensitivity, specificity, and accuracy are used. They are defined as follows:

\[
\text{Sensitivity} = \frac{TP}{TP + FN}, \quad \text{Specificity} = \frac{TN}{TN + FP}, \quad \text{Accuracy} = \frac{TN + TP}{TP + TN + FP + FN}
\]  

(1)

where TP (True Positive) is the number of pixels in the foreground that are correctly classified, TN (True Negative) is the number of pixels in the background that are correctly classified, FP (False Positive) is the number of pixels in the background that are classified as foreground, and FN (False Negative) is the number of pixels in the foreground that are classified as background.

We have evaluated all segmentation methods presented in this paper by using these three metrics. Table 1 shows measured average values of sensitivity, specificity, and accuracy for the results obtained using 10 datasets.

In addition, Table 1 lists the total computational costs, which are obtained using a 2 GB RAM and 2.40 GHz Intel Pentium CPU for all numbers of slices in a dataset.
Table 1. Quantitative analysis of all methods presented in this paper.

<table>
<thead>
<tr>
<th>Method</th>
<th>In percentage</th>
<th>Required time for a dataset (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>GMM</td>
<td>87.355</td>
<td>84.585</td>
</tr>
<tr>
<td>K-means</td>
<td>83.105</td>
<td>78.555</td>
</tr>
<tr>
<td>MLP</td>
<td>75.165</td>
<td>94.5</td>
</tr>
<tr>
<td>FTC</td>
<td>84.625</td>
<td>87.885</td>
</tr>
<tr>
<td>Automatic FTC</td>
<td>84.795</td>
<td>94.115</td>
</tr>
<tr>
<td>DRLSE with SPF</td>
<td>66.95</td>
<td>95.37</td>
</tr>
<tr>
<td>Automatic DRLSE with SPF</td>
<td>87.255</td>
<td>99.615</td>
</tr>
</tbody>
</table>


Figure 19 shows the manually segmented reference images (Figures 19a–19f) that we have used for the quantitative analysis of the implementation results.

Sensitivity measures the success of a method for identification of the TP and FP cases. Specificity indicates how well the method can identify the TN and FN cases. Therefore, accuracy is low (high) when both specificity and sensitivity are low (high). However, accuracy alone is not a good measure for a performance evaluation, because the value of accuracy is biased towards either specificity or sensitivity when either specificity or sensitivity has a high value and the other has a low value. Therefore, we have presented images whose sensitivity and specificity values are mostly close to each other. The comparison in Table 2 summarizes the advantages and disadvantages of the applied methods.

Figure 19. a) b) c) d) e) f) Manually segmented reference images.

9. Discussion and conclusions
The GMM-based probabilistic technique is more flexible than K-means, a heuristic method that uses distance values of gray levels and lacks a statistical foundation. The K-means algorithm cannot adapt to any cluster shape. Its cluster-modeling capability with a similar number of data points is limited to spherical clusters. In
Table 2. Advantages and disadvantages of the presented methods for liver segmentation from SPIR datasets.

<table>
<thead>
<tr>
<th>Method</th>
<th>Requirement of manual initialization</th>
<th>Required number of parameters</th>
<th>Requirement of postprocessing</th>
<th>Accuracy</th>
<th>Fast/Slow</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMM</td>
<td>No</td>
<td>3 (number of classes, variance of the Gaussian function, selected initial slice)</td>
<td>Yes</td>
<td>-Higher than K-means method</td>
<td>Slower than all presented methods in this paper except MLP method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Lower than all presented methods in this paper except K-means method</td>
<td></td>
</tr>
<tr>
<td>K-means</td>
<td>No</td>
<td>2 (number of classes, selected initial slice)</td>
<td>Yes</td>
<td>-Lower than all methods presented in this paper</td>
<td>Faster than GMM and MLP method</td>
</tr>
<tr>
<td>MLP</td>
<td>No</td>
<td>4 (number of iterations, learning rate, error tolerance, selected initial slice)</td>
<td>Yes</td>
<td>-Higher than GMM, K-means and FTC method</td>
<td>Slower than all methods presented in this paper</td>
</tr>
<tr>
<td>FTC</td>
<td>Yes</td>
<td>5 (number of iterations for regularization and data dependent cycles, initial curve, size of Gaussian filter, variance of Gaussian filter)</td>
<td>Yes</td>
<td>-Higher than GMM and K-means method</td>
<td>Slower than automatic FTC method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Lower than DRLSE with SPF, automatic FTC and automatic DRLSE with SPF method</td>
<td></td>
</tr>
<tr>
<td>Automatic FTC</td>
<td>No</td>
<td>5 (number of iterations for regularization and data dependent cycles, initial curve, size of Gaussian filter, variance of Gaussian filter)</td>
<td>Yes</td>
<td>-Higher than GMM, K-means and FTC method</td>
<td>Faster than GMM, MLP and automatic DRLSE with SPF method</td>
</tr>
<tr>
<td>DRLSE with SPF</td>
<td>Yes</td>
<td>6 (3 coefficients for the area, length and regularization terms, number of iterations, variance of the Gaussian kernel, initial curve)</td>
<td>No</td>
<td>-Lower than automatic DRLSE with SPF method</td>
<td>Faster than GMM and MLP method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Higher than GMM, K-means and FTC method</td>
<td></td>
</tr>
<tr>
<td>Automatic DRLSE with SPF</td>
<td>No</td>
<td>6 (3 coefficients for the area, length and regularization terms, number of iterations, variance of the Gaussian kernel, initial curve)</td>
<td>No</td>
<td>-Higher than all presented methods in this paper</td>
<td>Slower than the automatic FTC, K-means and DRLSE with SPF method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(GMM: Gaussian Mixture Model, MLP: Multi-Layer Perceptron, FTC: Fast Two Cycles, DRLSE: Distance Regularized Level Set Evolution, SPF: Signed Pressure Force Function)

the GMM, each mixture component is considered as a cluster of an arbitrary ellipsoidal shape. Therefore, the decision in the K-means clustering algorithm is difficult, due to assigning each data item to a single cluster. However, the drawback of the GMM-based segmentation with the EM technique is sensitivity to chosen initial values. The result of the GMM method depends on the number of chosen components. We have chosen a constant kernel number in our studies. The parameter values of the GMM are recalculated iteratively starting from the initial parameters until convergence. We have shown that neither the proposed GMM-based nor the K-means-based iterative segmentation method has the ability to accurately classify the liver in SPIR datasets because of the gray level similarities, intensity inhomogeneity, and partial volume effects. Using only unsupervised approaches such as K-means or model-based techniques to segment the liver from SPIR images does not give sufficient results. Both supervised and unsupervised methods should be used. Therefore, we
have applied the MLP-based liver segmentation method after several preprocessing steps and obtained more successful results. However, the drawback of this method is a very high computational cost.

We also have explained the properties of two different recently published level set-based segmentation techniques and presented their implementation results for liver segmentation. The FTC algorithm, which uses a switching mechanism, seems successful and can give acceptable results when preprocessed abdominal images are used. However, the drawback of the FTC algorithm is its sensitivity for initial contours that are defined in each slice. This is because the accuracy of the results depends not only on the size of the initial contours drawn by users, but also on the number of initial contours and their positions. In addition, the user-defined iteration numbers for each slice affect the segmentation results. Therefore, this approach is not robust, and it generates oversegmented or undersegmented images on some slices. In order to overcome these drawbacks, we have applied an automatic method iteratively using the FTC algorithm without any user interaction and with a small fixed number of iterations for liver segmentation from SPIR datasets. However, we have observed that automatic liver segmentation using the FTC method is not successful for SPIR datasets, even if preprocessed images are used. Therefore, we have applied the DRLSE with SPF function-based automatic segmentation method, which gives fast and acceptable results without any postprocessing operation. Moreover, this method has the ability to segment liver images without extracting adjacent organs except the gallbladder. In addition, it is more efficient than the MLP-based segmentation method in terms of the required segmentation time.

Both qualitative and quantitative comparison results of eight different active contour methods (except the application specific methods) are presented in [73] for brain MR images, ultrasound pig heart images, kidney CT images, knee MR images, and microscopy blood cell images. However, there is no comparative study for liver segmentation on SPIR images, which show the vascular structure of the liver very clearly and are very useful for vessel segmentation. None of the proposed approaches, namely, the deterministic iterative method (K-means based segmentation), the probabilistic model-based iterative method (GMM-based segmentation with EM), the supervised learning method (MLP-based segmentation), or the four different level set-based methods have been applied to liver segmentation from SPIR images. The contribution of this study is to make a comparison of state-of-the-art methods and present them for liver segmentation on abdominal MR images. Seven different algorithms have been implemented and the results obtained from SPIR image datasets are presented in this section. In addition, we propose an automatic liver segmentation approach using preprocessed images without any user interaction or postprocessing operations in Section 6.4.

The quantitative comparison results given in Table 1 show that the automatic DRLSE with SPF function-based segmentation approach using preprocessed images presents the most effective segmented liver images with the least computational cost among all the applied methods. Efficiency of the regularization of the level set function can be increased to get more successful results from this method as a future study. In this way, the computational cost of this method can be reduced. Vessel segmentation from the segmented SPIR images will be a future study.

References


