CHAPTER 5

FUSION BASED LIVER TUMOR DETECTION

5.1 OVERVIEW OF IMAGE FUSION

The performance of the liver tumor detection system is improved using image fusion technique. Image fusion is a technique to combine the registered images to increase the spatial resolution of acquired low detail multi-sensor images and preserving their spectral information. Some basic requirements could be imposed on the fusion scheme: (a) the fusion process should preserve all relevant information contained in the source images, (b) the fusion process should not introduce any artifacts into the image which would affect the human observer or subsequent processing stages, and (c) irrelevant features and noise should be suppressed to a maximum extent.

Image fusion aims to merge the information content from several images acquired from same or different imaging sensors of the same liver region in order to accomplish a fused image that contains the finest information coming from the original images. Hence, the fused image would provide enhanced superiority image than any of the original source images. Depending on the merging stage, fusion can be performed at three different levels—pixel level, feature level and decision level. In this work, pixel-level-based fusion is employed to represent a fusion process generating a single combined image containing an additional truthful description than individual
source image. This technique works by taking the average of the gray level source images pixel by pixel.

The proposed fusion system can be illustrated as given in Figure 5.1. Two malignant liver images taken at different orientations of the same patient are fused and then enhanced by the median filter to remove the noise present in it. The noise free fused image is then applied with Gabor transform to accurately detect the edges. Then, GLCM and laws texture features are extracted from the image and tumor regions are classified into benign or malignant by the neural network classification technique.

![Figure 5.1 Proposed image fusion based liver tumor detection methodology](image)

**Figure 5.1 Proposed image fusion based liver tumor detection methodology**

### 5.1.1 The Fusion Process

The fusion technique works by taking the average of the gray level source images pixel by pixel (Naidu & Raol 2008). Multiscale structure is simplified into two levels of scales, namely, foreground and background levels. The foreground signal includes the upper part of the original spectrum representing the small scale features like patterns, small objects and marcations which can be used in object detection and classification. The
background signals contain the neighboring components representing the large scale features like horizon position, terrain format, huge obstacles, necessary information about surrounding environment and information accountable for the natural appearance of the fused image. Important information is relatively easier to localize than in the case of the background signals and a feature selection mechanism can be implemented on pixel-level, thereby increasing the robustness of the fusion system compared to simple arithmetic fusion methods used for background signals.

The fusion of signals is performed independently at both levels. The foreground signals are obtained by finding the difference between the original and the background signals, which in turn are produced by taking product of the average filtering and combined using arithmetic fusion. The foreground signals exhibit high degrees of feature localization and are fused using simple pixel-level fusion techniques. In the fusion process, the filtering of input image signals is done using averaging of templates which produces a low-pass, background signal. The foreground signals are computed by subtracting the background from the original image. The background and foreground signals retain the same size and resolution as the original input signals since sub-sampling is not done.

The corresponding pixel for every pixel in the fused foreground image is chosen having the highest absolute value. Finally, the resultant fused signal, the background signal and foreground signal are summed up to produce the fused image. The fusion process is illustrated in Figure 5.2 clearly.
5.2  PROPOSED FUSION BASED LIVER TUMOR DETECTION

The proposed fully automatic technique to segment the liver tumor regions from liver CT is divided into two phases, namely liver region segmentation and tumor region extraction, after malignant tumor is detected. Before the Pre-processing step, pixel level fusion of liver images is done, after which Gabor filter is applied to transform the image into frequency domain. The texture features are extracted from the image and the image is classified into tumor and non-tumor regions. If malignant tumor is obtained, we apply morphological operators for exact segmentation of the tumor region from other regions. Figures 5.3 and 5.4 illustrate the process flow for benign and malignant liver image tumor detection and segmentation process, respectively.
Figure 5.3 Proposed flow of liver tumor detection system for Benign liver tumor image

Figure 5.4 Proposed flow of liver tumor detection system for malignant liver tumor image
5.2.1 Pixel Level Image Fusion

Pixel-level image fusion integrates the information from multiple images of one scene to get an informative image which is more suitable for human visual perception or further image processing. The proposed system fuses two liver CT images of the same patient taken at different angles, in order to improve the image quality and also obtain the detailed information from the fused (enhanced) liver image.

A typical pixel level fusion system consists of six sub-systems: imaging, registration, preprocessing, fusion, post-processing and displaying. Various pixel level fusion algorithms have been proposed. The simplest pixel level fusion method, namely the weighted averaging (WA) fusion is employed. The simplest image fusion based on weighted averaging is by taking the average of the source image pixel by pixel, such as:

\[ C(m, n) = \alpha A(x, y) + \beta B(x, y) \]  \hspace{1cm} (5.1)

where, \( \alpha \) and \( \beta \) are the scalar weights. The WA method is simple and fast to implement. This method also reduces the noise present in the source image.

5.2.2 Median Filter

The median filter (Zhao et al 2009) is normally used to reduce noise in an image, rather like the mean filter. However, it is better than the mean filter as it preserves the useful details in the image. This class of filter belongs to the class of edge preserving smoothing filters which are non-linear filters. This means that for two images \( A(x) \) and \( B(x) \):

\[ \text{median}[A(x) + B(x)] \neq \text{median}[A(x)] + \text{median}[B(x)] \]  \hspace{1cm} (5.2)
These filters smooth the data while keeping the small and sharp details in the image. The median is just the middle value of all the values of the pixels in the neighborhood. The median is not the same as the average or mean, instead, the median has half the values in the neighborhood larger and half smaller. The median is a stronger “central indicator” than the average.

The median filter considers each pixel in the image and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. The median is calculated by first arranging all the pixel values in ascending order and then the median (middle) pixel replaces the central pixel value in the 3×3 sub-image. Figure 5.5 illustrates an example of median filtering process considering a 3×3 sub-image.

![Figure 5.5 Computation of median value in a median filter](image)

Figure 5.5 Computation of median value in a median filter
5.2.3 Gabor Filter

In digital image processing, a Gabor filter (Ji et al 2004) can be defined as a linear filter employed in detecting the edges of an image. The Gabor filter represents the orientation and frequency of the image very similar to the human visual system. The 2D Gabor filter acts as a Gaussian kernel function in the spatial domain modulated by a sinusoidal plane wave, and is particularly appropriate for texture representation and differentiation. In Gabor filters, all filters can be generated from one mother wavelet by dilation and rotation and hence, they are self-similar.

Gabor filters can be designed for a number of dilations and rotations, thus directly related to Gabor wavelets. Gabor filters exhibit desirable characteristics of spatial locality and orientation selectively and is optimally localized in the space and frequency domains have been extensively and successfully used in face recognition. The Gabor kernels used are defined as follows:

\[
\psi_{\mu,\nu} = \frac{k_{\mu,\nu}^2}{\sigma^2} \exp \left( - \frac{k_{\mu,\nu}^2 z^2}{2\sigma^2} \right) \times \left[ \exp(i k_{\mu,\nu} z) - \exp \left( - \frac{\sigma^2}{2} \right) \right]
\]  

(5.3)

where, \(\mu\) & \(\nu\) are the orientation and scale of the Gabor kernels, respectively, \(z=(x,y)\), and \(k_{\mu,\nu}\) is the wave vector. Finally, the noise removed liver CT image is transformed to frequency domain image after application of the Gabor filter.

5.2.4 GLCM Features

Gray Level Co-occurrence Matrix (GLCM) is a second order statistic measurement containing information about the positions of pixels with similar gray level values. Numerous statistical features named Haralick texture features (Haralick et al 1973), were extracted using the GLCMs.
Second-order statistics are similar to that of observing a pair of gray values occurring at the endpoints of a dipole (or needle) of random length placed at a random location and random orientation over the image. GLCM directions of analysis includes, Horizontal (0° or 180°), Vertical (90° or 270°), Right Diagonal (45° or 225°), and Left diagonal (135° or 315°). The GLCM contains the second-order statistical information of spatial relationship the pixels of an image.

GLCM is formed from a gray-scale image and it contains information about how often a pixel with gray-level (gray scale intensity or Tone) value $i$ occurs either horizontally, vertically, or diagonally to adjacent pixels with the value $j$, where $i$ & $j$ are the gray level values (tone) in an image. GLCM is a $N \times N$ matrix, in which $N$ is the number of gray levels in the input image. Number of pixel pair repetitions are counted and updated in the GLCM matrix. GLCM can be created for four directions namely 0°, 45°, 90°, 135°.

For the purpose of feature extraction, GLCM should be a symmetric and normalized matrix. To make a matrix symmetric, transpose of GLCM is taken and added with the original GLCM. To get a normalized matrix, sum of all elements in a GLCM is calculated and each element of the matrix is divided with the obtained sum. From the normalized symmetrical GLCM, the texture features are extracted. Five properties of GLCM that are used for our evaluation are energy, correlation, contrast and homogeneity. Consider a matrix element $P(i,j \mid \Delta x, \Delta y)$ which is the relative frequency of two pixels separated by pixel distance ($\Delta x$, $\Delta y$). The GLCM properties within a given neighborhood with gray level intensity $i$ and intensity $j$ are computed as,

$$\text{Energy} = \sum p(i,j)^2 \quad (5.4)$$
Correlation = \sum (i - \mu_i)(j - \mu_j) \frac{p(i,j)}{[\sigma_i \sigma_j]} \quad (5.5)

Contrast = \sum (|i - j|^2 \times p(i,j)) \quad (5.6)

Homogeneity = \frac{\sum p(i,j)}{1+|i-j|} \quad (5.7)

The number of gray levels in an image determines the size of GLCM. A GLCM contains information about the positions of pixels having similar gray level values. A GLCM \( P[i, j] \) is defined by first specifying a displacement vector \( d = (dx, dy) \) and counting all pairs of pixels separated by a distance ‘d’ having gray levels \( i \) and \( j \).

### 5.2.5 Law's Texture Feature Extraction

Feature extraction is the process of obtaining higher-level information of an image such as color, shape and texture. Texture is a key component of human visual perception. The Law’s method uses filter masks to extract secondary features from natural micro-structure characteristics of the image (level, edge, spot and ripple) which can then be used for segmentation or classification. Laws developed five labeled vectors which could be combined to form two dimensional convolution kernels. When convolved with a textured image these masks extract individual structural components of the image (Laws 1980).

After a series of particular convolution with selected Laws’ masks, the outputs are passed to texture energy measurement (TEM) filters for the analysis of the texture property of each pixel. These consists of a moving non-linear window operation, every pixel of the image is replaced by comparing the pixel with its local neighborhood based on three statistical
descriptors (mean, absolute mean and standard deviation). These descriptors are computed as follows:

\[
mean = \frac{\sum_{W \text{neighboring pixels}}}{w} \tag{5.8}
\]

\[
\text{absolute mean} = \frac{\sum_{W \text{abs(neighboring pixels)}}}{w} \tag{5.9}
\]

\[
\text{standard deviation} = \sqrt{\frac{\sum_{W (\text{neighboring pixels - mean})^2}}{w}} \tag{5.10}
\]

where W is the window size. The operation will lead to the creation of three TEM images corresponding to each statistical descriptor. After the windowing operation, all the obtained images are normalized in order to be presented well as images. Min-max normalization method is utilized in this work.

5.2.6 Gray Level Based Features

These features are based on the differences between the gray-level intensity values in the candidate pixel and its surrounding pixels. Since liver tumor lesions are always darker than their surroundings, features based on the gray-level variations in the surroundings of the tumor region are an excellent option to segment the tumor region. A set of gray-level-based descriptors are derived from homogenized image H considering only a small pixel region centered on the described pixel \((s, t)\). \(\mathcal{S}_{m,n}^k\) stands for the set of coordinates in a \(L \times L\) sized square window centered at the point \((m, n)\). The Gray level based feature descriptors \((D)\) each represent a feature image and are represented as follows:
\[ D_1(m, n) = H(m, n) - \min \{H(s, t)\} \]
\[ D_2(m, n) = \max \{H(s, t)\} - H(m, n) \]
\[ D_3(m, n) = H(m, n) - \text{mean} \{H(s, t)\} \]
\[ D_4(m, n) = \text{std}\{H(s, t)\} \]
\[ D_5(m, n) = H(m, n) \]

5.2.7 Neural Network Classification

The neural network (NN) classifier (Rehna & Jeyakumar 2011) is defined as an information-processing system inspired by the structure of the human brain. Inspired by the biological neuron in the brain, NNs consist of a number of interconnected neurons. A neuron is an information-processing unit that receives several signals from its input links, each of which has a weight assigned to it. These weights correspond to synaptic efficiency in biological neurons. Weights are the basic means of the long term memory in NNs. Neural networks (NNs) are adaptive non-linear statistical data modeling or decision making tools.

NNs can be used to model complex relationships between inputs and outputs or to find patterns in data, thereby making it suitable for brain tumor segmentation. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. Feed-forward ANNs allow signals to travel one way only, i.e. from input to output. Feed-forward ANNs tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition. The behavior of an ANN (Artificial Neural Network) depends on both the weights and the input-output function (transfer function) that is specified for the units. This function can be grouped into one of the three categories:
The activation function controls the amplitude of the output of the neuron. An acceptable range of output is usually between 0 and 1, or -1 and 1. The functional model illustrating this process is shown in Figure 5.6. The output of the neuron, $O_k$, is produced due to the outcome of an activation function $A_k$, such that,

$$A_k = \sum_{h=1}^{q} W_{kh} P_h$$ \hspace{1cm} (5.12)

The diagram shows the practical model of the neural network.

**Figure 5.6 Practical model of the neural network**

Neural networks are employed in this method to classify the tumor tissues and non-tumor regions of the liver. The NN is trained with the features extracted from the liver CT images and it maintains a database with it containing the feature sets/values of the image. The image is divided into several small regions for the purpose of feature extraction. The NN works by...
comparing these feature values extracted from each small region of the liver with the features extracted from the test image. The liver region is selected and the segmentation parameter is adjusted, such that the segmentation is repeated until the difference between the feature values of the liver region and NN output did not decrease. Finally, the non-tumor tissues and the tumor tissues are classified as benign or malignant.

A single layer feed-forward network consists of one or more output neurons, each of which is connected with a weighting factor to all of the inputs. A simple feed-forward network consists of only two inputs and a single output, as depicted in Figure 5.7.

![Figure 5.7 A single layer feed-forward neural network](source: Rehna & Jeyakumar 2011)

The input of the neuron is the weighted sum of the inputs and the bias term. The output of the network is formed by the activation of the output neuron, which is some function of the input:

\[ y = A_F \left( \sum_{i=1}^{2} w_i x_i + \theta \right) \]  

(5.13)
The activation function $F$ can be linear thus making the network to be linear. In this work, we consider the threshold function, given by,

$$F(s) = \begin{cases} +1, & \text{if } s > 0 \\ -1, & \text{otherwise} \end{cases}$$

(5.14)

The output of the network thus is either +1 or -1 depending on the input. The network can now be used for a classification task, to decide whether an input pattern belongs to one of the two classes.

5.2.8 Morphological based Operations for Tumor Segmentation

The tumor regions are accurately segmented using morphological operations (Moltz et al 2008). The morphological operations are applied on the gray scale liver image to segment the abnormal regions. Erosion and dilation are the two elementary operations in mathematical morphology. Both the operations are combined to characterize the remaining operations. The symbols $\Theta$, $\Theta$, $\circ$, and $\bullet$, denote the dilation, erosion, opening and closing operations, respectively. A function $f(x, y)$ denotes the image, and the function $h(x, y)$, or $h$ denotes the structuring element. The four operations are defined as follows:

**Dilation:**

$$dilation: (f \oplus h)(x, y) = \sup_{(r,a) \in H} \{x - r, y - s\} + h(r,s)$$

(5.15)

**Erosion:**

$$erosion: (f \ominus h)(x, y) = \inf_{(r,a) \in H} \{x + r, y + s\} - h(r,s)$$

(5.16)

**Opening:**

$$opening: f \circ h = (f \ominus h) \oplus h$$

(5.17)

**Closing:**

$$closing: f \bullet h = (f \oplus h) \ominus h$$

(5.18)
where, $\inf{\{\}}$ and $\sup{\{\}}$ denote the infimum and supremum operations, respectively. Erosion and Dilation are fused to form the Opening operation, by which objects that are adjacent are spaced and objects that are adjoined are detached and the holes within the objects are enlarged.

### 5.3 RESULTS AND DISCUSSION

The proposed liver tumor detection system is evaluated for its performance appraisal. To keep the objectiveness of the evaluation, the results obtained are compared with the ground truth images for evaluation of the segmentation results to obtain relevant measurements and scores. The segmentation results are compared with the ground truth image obtained by expert physician. The following parameter (Ballin et al 2009) help in determining the classification performance, and are given by,

\[
\text{Accuracy} = \frac{(TP+TN)}{(TP+FN+TN+FP)} \tag{5.19}
\]

where, TP denotes true positive, FP denotes false positive, FN is false negative and TN is true negative. TP refers to the correctly identified tumor pixels, TN refers to the wrongly identified tumor pixels, FP refers to the correctly identified non- tumor pixels and FN refers to the wrongly identified non- tumor pixels. The results of the proposed method are shown in Figure 5.8. The results include the simulation results after the proposed method is applied on the malignant liver CT image.

The image in Figure 5.8b represents the image obtained after fusion, Figure 5.8c, 5.8d, 5.8e represents the median filtered image, thresholded image and Gabor filter applied image, respectively. Figure 5.8f denotes the tumor regions after segmentation. The same tumor regions are again represented within the liver image in Figure 5.8g.
Figure 5.8 Results of proposed methodology: (a) Malignant liver Image, (b) Fused image, (c) Median filtered image, (d) Thresholded image, (e) Gabor transformed image, (f) Detected tumor regions, (g) Result of tumor segmentation
The proposed algorithm was able to detect tumors from the liver images based on the extracted features with the help of the neural network classifier. The best results were obtained for large tumors with strong edge gradients and for tumors surrounded by homogeneous liver tissue. Even smaller tumors near the liver boundary with weak gradients were detected accurately.

**Table 5.1 Performance comparison of proposed method**

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>98.5</td>
</tr>
<tr>
<td>Particle Swarm Optimization (Sharma &amp; Kaur 2013)</td>
<td>93</td>
</tr>
<tr>
<td>Seeker Optimization algorithm (Sajith &amp; Hariharan 2013)</td>
<td>60</td>
</tr>
<tr>
<td>Non-density based method (Zhang et al 2007)</td>
<td>95.7</td>
</tr>
</tbody>
</table>

Table 5.1 depicts the accuracy levels of classification of other well known methods, proposed by several researchers. The accuracy in classifying the tumor regions from the liver images was done using various different algorithms as proposed by Freiman et al (2010) and others. The proposed method is being compared in terms of classification accuracy, which clearly shows that the proposed method classifies the tumor more accurately than most other methods (Freiman et al 2010; Sharma & Kaur 2013; Sajith & Hariharan 2013; Zhang et al 2007). Since the proposed technique uses texture based features and a good classifier (neural network), the tumor regions are more precisely segmented from the normal tissues, with an accuracy of 98.5% which is higher than conventional methods as stated in Figure 5.1.
5.4 SUMMARY

In this chapter, an automatic method for the extraction of the liver region from CT images of liver is proposed using neural network classifier. The experimental results demonstrate that the proposed system is effective in the segmentation of liver tumors, and it is expected to be useful for radiologists in the clinical examination of liver images. The tumor segmentation method proposed in this chapter includes a novel method for tumor classification which helps the medical experts to detect the tumor regions more accurately. The main advantage of this method is that it yields accurate results about 98.5% of classification accuracy with respect to ground truth images.