# Classification of Lumbar Ultrasound Images with Machine Learning

Shuang Yu and Kok Kiong Tan

National University of Singapore, Mechatronics and Automation Lab, Singapore
{yushuang,kktan}@nus.edu.sg

**Abstract.** In this paper, we propose a feature extraction and machine learning method for the classification of ultrasound images obtained from lumbar spine of pregnant patients in the transverse plane. A set of features, including matching values and positions, appearance of black pixels within predefined windows along the midline, are extracted from the ultrasound images using template matching and midline detection. Artificial neural network is utilized to classify the bone images and interspinous images. The neural network is trained with 1000 images from 25 pregnant subjects and tested on 720 images from a separate set of 18 pregnant patients. A high success rate (96.95% on training set, 95.75% on validation set and 94.12% on test set) is achieved with the proposed method. The trained neural network further tested on 43 videos collected from 43 pregnant subjects and successfully identified the proper needle insertion site (interspinous region) in all of the cases. Therefore, the proposed method is able to identify the ultrasound images of lumbar spine in an automatic manner, so as to facilitate the anesthetists' work to identify the needle insertion point precisely and effectively.

## 1 Introduction

Epidural/spinal anesthesia (EA) is widely used in surgery and for post-surgical pain relief. A properly performed epidural procedure is the 'gold standard' of treatment to reduce pain during childbirth [1]. Around 50-90% of women in labour in developed countries choose EA for pain relief [2]. However, the failure rate of EA has been reported to be as high as 20% [3]. One of the key challenges for EA is the identification of the needle insertion site, which is traditionally identified by palpating the patients' lumbar spine [4]. This blind technique may require multiple needle insertion attempts, leading to complications in the process. The case is worse for patients with obesity problems, which is increasingly prevalent in the pregnant population.

Ultrasound imaging, as a non-radioactive, convenient and inexpensive medical imaging modality, has been introduced to EA to assist epidural needle insertion since the 1950s [5]. Previous researches have confirmed the effectiveness of ultrasound imaging compared with the traditional palpation method [6]. Despite the benefits of ultrasound, the effective interpretation of ultrasound images remains a challenge, especially for anesthetists who received limited training in reading

G. Dick et al. (Eds.): SEAL 2014, LNCS 8886, pp. 287–298, 2014.

ultrasound images [7]. The low spatial resolution and severe speckle noises of ultrasound images results in the subtle anatomical features becoming indiscernible from the surrounding background [8]. It requires professional training to fully interpret the ultrasound images and the learning curve is steep. Therefore, a large proportion of anesthetists are reluctant to adopt ultrasound imaging in the common practice.

In order to ease the ultrasound image interpretation and facilitate the applicability of ultrasound in epidural needle insertion, automatic interpretation of lumbar ultrasound images has been investigated by researchers. Train et al. utilized phase symmetry and template matching to extract the lamina and ligamentum flavum in the paramedian images [9]. Kerby et. al proposed to label the lumbar level automatically with panorama images obtained from the paramedian view [10]. Furthermore, an augmented reality system (AREA) which projected the identified lumbar vertebra levels on the patients back was developed so as to assist spinal needle insertion [11].

Although automatic interpretation of lumbar ultrasound images has been explored, it is mainly focused on the paramedian view. Ultrasound images in the transverse view, which reveal important anatomical information and frequently been used by anesthetists for precise pre-puncture localization of needle insertion site, are less researched from the automatic image interpretation perspective. In our previous research, an image processing and identification procedure was developed for the automatic interpretation of ultrasound images in the transverse view [12]. Template matching combined with position correlator (PC) was proposed to identify the interspinous images and achieved a success rate of 100% on ultrasound images obtained from lumbar spine of healthy volunteers. However, since the clarity of anatomical feature of lumbar spine might degrade during pregnancy [13], the original position correlator designed for healthy volunteer is not effectively applicable to the pregnant patient.

In order to improve the identification accuracy for pregnant patients and make the classification algorithm more generally applicable, a feature extraction and classification procedure is developed. Three contributions are achieved with this paper. First, a set of features, which are composed of important parameters, are extracted from the lumbar ultrasound images with template matching and midline detection methods. Secondly, a multi-layer neural network is utilized to classify the interspinous images and bone images with the extracted feature vector. A high success rate is achieved with the proposed feature extraction and neural network structure on images collected from the pregnant patients. Last but not least, the trained neural network model is also tested on 43 videos and it successfully identify the interspinous region and bone region on all of the cases collected, with a computational speed fast enough for real-time processing.

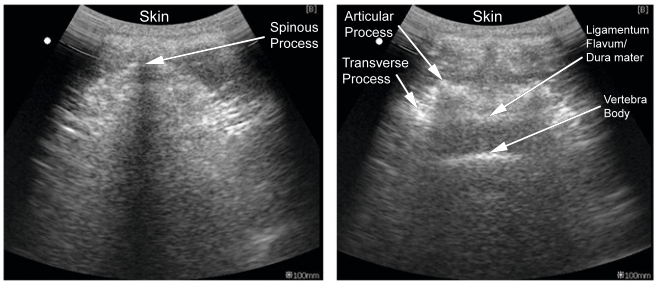


**Fig. 1.** Ultrasound Image of Lumbar Spine. (a) typical ultrasound image when the probe is placed above spinous process, featured by the triangular anechoic window; (b) ultrasound image when probe is placed on interspinous space, where the articular processes, epidural space and vertebra body are visible.

## 2 Materials and Methods

### 2.1 Ultrasound Image Feature of Lumbar Spine

The ultrasound images taken at different region of the lumbar spine have different features, determined by the region where the probe is placed. When the probe is placed directly on the spinous process (not proper for needle insertion), the ultrasound wave will be impeded by bones, creating a long triangular hypo-echoic acoustic shadow Fig 1(a)). The ultrasound image will be dark with a triangular dark window along the midline, which is the main feature of bony images. When the probe is moved to the interspinous region (proper for needle insertion), more details beneath the skin can be noted, as shown in Fig 1(b). The 'flying bat' alike shape on the ultrasound image indicates that the location of the probe is a suitable site for needle insertion [14].

### 2.2 Feature Extraction

Before feature extraction, raw ultrasound images are pre-processed with difference of Gaussian enhanced local normalization, so as to remove the speckle noises and extract the anatomical structure [12]. After pre-processing, local intensity variance induced by uneven ultrasound wave reflection rate are also eliminated. Therefore, a potential element which might deteriorate the image classification is removed.

Feature extraction procedure is extraordinarily important for image classification. Medical images generally suffer from limited training samples, thus the feature vector length shall be limited. Otherwise, the learning models will have high variance and cannot be optimally trained. In this paper, image features are extracted with two approaches, the template matching method to detect the key anatomical features and midline detection approached to obtain image features along midline.

**Template Matching.** The visibility of 'flying bat' shape is the criterion adopted by anesthetists to recognize interspinous images [14]. However, in computer vision, due to the variation and distribution extent of the 'flying bat' shape in the image, the recognition of the entire 'flying bat' shape is not a easy task. In our previous research, we proposed to decompose the 'flying bat' shape into three sub-features: the 'bat ear' (articular process), epidural space and vertebra body. The decomposed sub-features recognized the articular process and vertebra body with high accuracy on images obtained from volunteers [12].

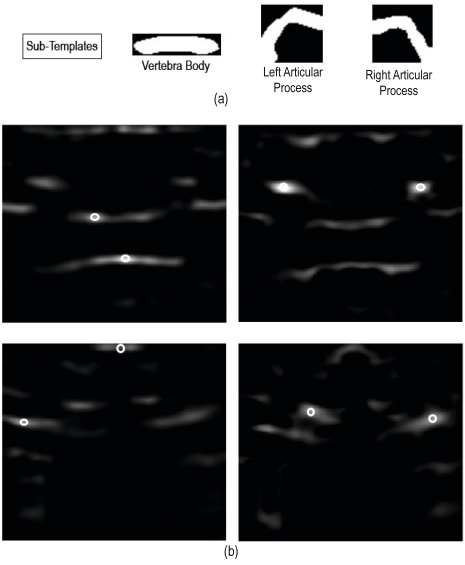


**Fig. 2.** Feature Extraction with Template Matching. (a) Sub-templates for anatomical features, from left to right: Vertebra body and epidural space, left articular process and right articular process; (b) Matching result of key anatomical features: The left column: matching result of vertebra body sub-template; the right column: matching result for articular process sub-templates; the upper row: interspinous image; the lower row: bone image. The optimal matching position is marked by a circle.

In this paper, similar decomposition is employed. Template matching is used to obtain the matching position and matching value between the sub-features and the images. Among the three sub-features, the appearance of the epidural space and the vertebra body both resemble a line. Thus, the same linear sub-template (as shown on Fig 2(a)) is employed for the recognition of both vertebra body and epidural space. Of the two maximum matching blobs, the one that locates lower in the image is vertebra body and the superior one is epidural space, which follows the anatomical structure of the lumbar spine. In the interspinous images, the visibility of vertebra body and epidural space is clear and both

of them can be correctly recognized. While in the bone images, the maximum matching of the sub-template will occur at different regions in the image; and the matching values for both epidural space and vertebra body are low, as indicated by Fig 2(b). The situation is the same for the matching of articular processes, except that the maximum matching of articular processes will appear on the left and right side of the midline. Therefore, based on the matching position and matching value, it is possible to partially discriminate the interspinous images and bone images.

Fig 2(b) shows the template matching results of the epidural space, vertebra body and left & right articular process, with the optimal matching position been marked. The parameters obtained with template matching can be utilized to constitute part of the feature vector for the purpose of image classification, including the retrospective depth measurement of epidural space ($\mathcal{D}_1$) and vertebra body ($\mathcal{D}_2$), their matching values ($\mathcal{V}_1$ and $\mathcal{V}_2$), matching position of two articular processes ($\mathcal{P}_3$, $\mathcal{D}_3$ for left articular process and $\mathcal{P}_4$, $\mathcal{D}_4$ for right articular process) and their matching values ($\mathcal{V}_3$ and $\mathcal{V}_4$).

**Midline Detection.** The image features along the midline of the ultrasound image is different for interspinous images and bone images. For the bone images, ultrasound wave is impeded by the spinous process, resulting in an anechoic region along the midline (Fig 1(a)); while for interspinous images, the epidural space and vertebra body along the midline will be visible (Fig 1(b)). Therefore, the appearance of black pixels along the midline serve as an important feature for the classification of interspinous / bone images.

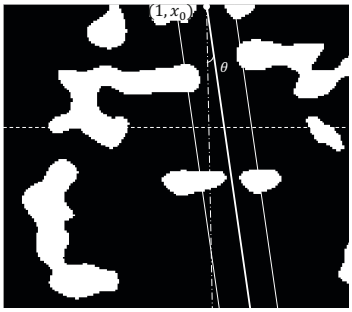


**Fig. 3.** Feature Extraction with Midline Detection. (Note: the background image is the pre-processed binary image; The horizontal dashed line is the depth threshold used to calculate the white pixels rate$\mathcal{R}_w$. $x_0$ denotes the center of scanning window along the horizontal axis and $\vartheta$ denotes the angle of scanning window against the vertical axis.).

For the detection of midline, a cost function $J(\vartheta, x_0)$ based on the summation of white pixels within a predefined scanning window is formulated. The window scanned though the entire image within [-45, +45] degrees. The position and degree that gives the minimum cost function value will locate the midline, as

demonstrated by Fig 3. In order to increase the accuracy of midline detection for interspinous images, a penalty which decreases its weight as a function of depth is imposed on the cost function, so as to allow the appearance of epidural space and vertebra body to be less penalized in the cost function. The cost function is formulated in Equation 1.

$$J(\vartheta, x_0) = \sum_{i=1}^{n} \sum_{j=-C}^{C} [0.5 + exp(-0.05i)] \times f(i, itan\vartheta + x_0 + j) \times \sqrt{(1 + 0.3|\vartheta||x_0 - n/2|)} \tag{1}$$

The first part of the Equation 1 is the penalty term for the appearance of white pixels at different depths. The third part is the penalty term if the detected midline is not near the middle of the ultrasound image or that it is not vertical. In equation 1, $f(i, j)$ denotes the binary image of the pre-processed ultrasound image with a dimension of $n \times m$; C represents half size of the predefined window, which can be optimally set between 5 - 10.

After optimal $\vartheta'$ and $x'_0$ is obtained and midline is located, the rate of black pixels within the predefined scanning window can be calculated using the following equation:

$$\mathcal{R}_b = 1 - \frac{\sum_{i=1}^{n} \sum_{j=-C}^{C} f(i, itan\vartheta' + x'_0 + j)}{2Cn} \tag{2}$$

The depth of epidural space is reported to range from 3-8 cm, indicating that the epidural space and vertebra body appear deeper than 3cm in the image. Thus, the rate of potential epidural space and vertebra body within the scanning window can be calculated with:

$$\mathcal{R}_w = \frac{\sum_{i>=3cm}^{n} \sum_{j=-C}^{C} f(i, itan\vartheta' + x'_0 + j)}{2Cn} \tag{3}$$

Because of the presence of epidural space and vertebra body at the lower part of the interspinous image, the parameter $\mathcal{R}_w$ is bigger than 0. On the contrary, for the images obtained from bone regions, the lower part of the image is black. Thus, $\mathcal{R}_w$ approximates 0 for bone images.

After the midline is located via the cost function approach, symmetry measurement is utilized to double confirm the accuracy of midline detection. The introduction of symmetrical parameter is based on the fact that the anatomical structure of lumbar spine exhibits mirror symmetry with respect to the midline. The symmetrical parameter $\mathcal{S}$ is simply calculated with Equation 4

$$\mathcal{S} = \frac{\sum |f(x, y) - f(x', y')|}{nx'_0} \tag{4}$$

where $(x, y) and (x', y')$ represent the coordinates of one pair of pixels which are symmetrical to each other against the detected midline.

$\mathcal{R}_b$, $\mathcal{R}_w$ and $\mathcal{S}$ add another three parameters for the feature vector. Therefore, combining the 10 parameters obtained from template matching and 3 parameters from midline detection, a feature vector of length 13 is formulated. A detailed description of template matching and midline detection on lumbar ultrasound image processing can be further approached at [15].

### 2.3 Multi-Layer Neural Network

After the feature set has been extracted and normalized, the multi-layer neural network (MLP) is employed for the classification of the interspinous images and bone images. The MLP contains three or more layers, with an input layer, an output layer and one or more hidden layers of nonlinearly-activating nodes (usually sigmoid or tanh function). Theoretically, MLP can approximate any bounded continuous function [16]. In order to simplify the network structure, only one hidden layer is utilized in this paper, as shown in Fig 4.

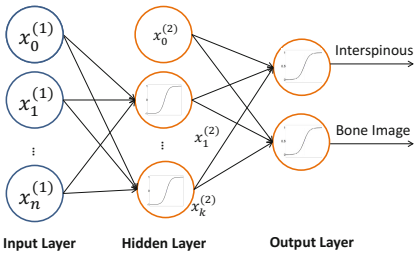


**Fig. 4.** Multi-layer Neural Network Structure: on input layer with 13 neurons, one hidden layer with k neurons and one output layer with 2 neurons

The learning procedure of MLP is fulfilled by changing synaptic weights of neurons after training data is processed. The most popular method for training of MLP is the back-propagation (BP) algorithm, witch provides a computationally efficient approach for MLP training [17]. The training processing using BP algorithm can be divided into two phases: the forward phase and the backward phase. In the forward phase, the synaptic weights are fixed and the training samples are propagated through the network. In the backward phase, the resulting error produced by comparing the output with desired output is propagated through the network backwards, during which successive adjustments are made to the synaptic weights [18].

### 2.4 Evolutionary Neural Network

Back Propagation is computationally efficient for the training of feed forward neural network. However, it is prone to be trapped in local minimum [18].

For the complex problems, evolutionary computation, e.g. genetic algorithm, can be involved in the training of neural network to obtain the global optimum by performing searches over a complex and multi-mode space [19]. For the classification problem, firstly, calculate the fitness function, usually the misclassification rate with the generated neural network; select the individuals with largest fitness and reserve them to the next generation; then perform the crossover and mutation with the current population to generate the new generation. The above procedure is repeated to evolve the inital weights until the training goal is achieved. Another commonly used evolutionary neural network training is hybrid algorithm, which combine genetic algorithm and back propagation together. The hybrid training fist use genetic algorithm to optimize the initial weight distribution and locate certain search spaces in the solution space. Then use back propagation to search the optimal solution in the small solution spaces [20].

In this paper, the problem is not very complex, since only two classes are involved and the dimension of the input vector is not high. Back propagation is good enough to train the neural network. However, in order to avoid the local minimum problem, the neural network is trained 5 times consequently.

### 2.5 Materials and Image Acquisition

The ultrasound video streams utilized in this research were collected from KK Women's and Children's Hospital (Singapore), with institutional review board (IRB) approval and patients' consent obtained. Pregnant women scheduled for a caesarean procedure were recruited before they were sent to the operation theater (OT). During the study, 43 ultrasound video streams are collected from 43 different subjects. After video streams are collected, the image database is obtained by extracting still images from the video streams. 40 images are randomly extracted from each of the video streams, constituting 1720 ultrasound images in the training and test database in total. The extracted images are then labelled by an experienced sonographer: '1' for interspinous images, '0' for bone images and other images not proper for needle insertion.

**Table 1.** Statistics of Training Set, Validation Set and Test Set

| | Training Set | Validation Set | Test Set |
|---|---|---|---|
| Subject Number | 21 | 10 | 12 |
| Image Number | 840 | 400 | 480 |
| Interspinous | 404 | 182 | 204 |
| Bone Images | 436 | 218 | 276 |

## 3 Results and Discussions

Of the 43 video streams collected, 21 (49%) of them are randomly selected as training set, another randomly selected 10 subjects (23%) are used for validation

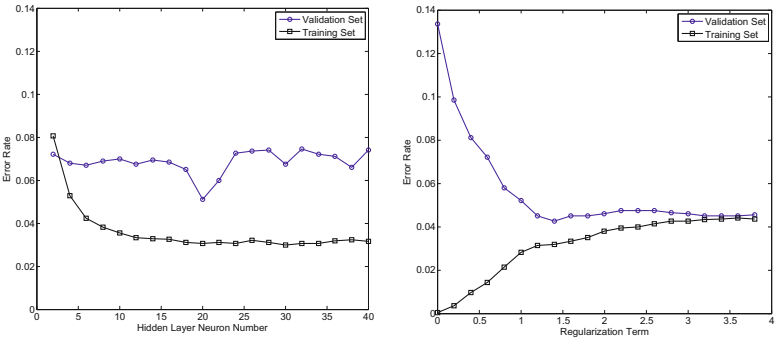


**Fig. 5.** Parameter Tuning of the Multi-layer Neural Network. (a). Fix the regularization term as 1 and change the hidden neuron numbers. (b). Fix the hidden neuron number as 20 and change the regularization term.

set and the reset 12 (28%) are used as test set. The training, validation and test sets are divided on the level of subjects instead of extracted images, which follows the assumption that in the clinical setting the detailed lumbar spine structure of individuals are not given neither examined with MRI or other imaging modalities before the epidural anesthesia. Since 40 images are extracted from each video, thus there are in total 840 images in the training set, 400 images in the validation set and 480 images in the test set. The detailed statistical information of the images is listed in Table 1.

### 3.1 Performance of Neural Network

The multi-layer neural network model is trained with the training set and then validated on the validation set to get the optimal parameters. There are two parameters involved in the neural network structure with one hidden layer: the number of neurons in the hidden layer and the regularization term to avoid over-fitting. Since the BP training might get trapped in local minimum, thus the network is trained 5 times so as to get the estimation of the mean performance.

In order to get the optimal performance, the test is conducted in two steps. First, the regularization term is fixed as 1 and then get the number of hidden layer neurons. As shown in Fig 5(a) when the regularization term is set as 1, the optimal performance is obtained when the hidden neuron number is 20. Then the hidden neuron number is set as 20 and then tune the regularization term. Minimal error rate is achieved when regularization term is 1.4.

Therefore, the optimal performance on the validation set is achieved when the hidden neuron number is 20 and the regulation term is 1.4. The trained network models with the obtained parameters is further tested on the test set. The performance of the trained model is listed on Table 2.

**Table 2.** Performance of Neural Network

| | Training Set (%) | Validation Set (%) | Test Set (%) |
|---|---|---|---|
| Accuracy | 96.95 | 95.75 | 94.12 |
| Precision | 97.11 | 95.00 | 94.35 |
| Recall | 97.02 | 97.71 | 95.51 |
| F0.5 | 97.09 | 95.53 | 94.58 |

Fig 6 demonstrates the receiver operating characteristic curve (ROC) of the trained model on the test set. The area under the curve (AUC) of the neural network model is 0.981 for the test set, indicating that the neural network model is properly trained and has good predictability [21].

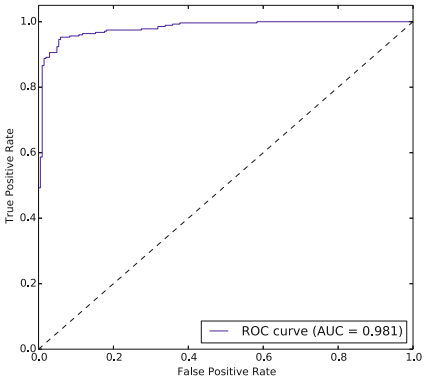


**Fig. 6.** Receiver Operating Characteristic Curve of the Trained Neural Network Model on Test Set

### 3.2 Video Processing

The trained neural network model is further tested on the ultrasound video streams collected to identify the interspinous region and bone region. In the video processing, the interspinous region is defined by the continuous appearance of more than 5 interspinous images; while for the negative detections, if it is in the interspinous region, no more than 2 bone images shall be detected by the image; vice versa for bone region. According to the definition above, the neural network model is able to identify the interspinous region and bone region correctly on all of the 43/43 video streams collected.

Table 3 lists the computation time for major operations in the pre-processing, feature extraction and classification procedure. Matlab (R2012a) was used for the implementation of the algorithm. The computation time for each frame is 67.58 ms. Given that the video is collected at the frame rate of 15 FPS, thus the computation speed is a little bit slower for real-time processing and may

result in frame loss for real time processing. However, improvement in computational speed has been realized by implementing the program using Python (with OpenCV library), which shortened the computation time to 32.85 ms per frame. Therefore, the proposed image processing procedure is applicable to real time processing.

**Table 3.** Computation Cost of Video Processing with Matlab

| Operation | Computation Cost(ms) |
|---|---|
| Preprocessing | 9.10 |
| Template Matching | 17.21 |
| Midline Detection | 17.65 |
| Symmetry Detection | 15.32 |
| Neural Network Classification | 0.08 |
| Others | 8.25 |
| Processing Time Per Frame | 67.58 |

## 4 Conclusion

In this paper, we propose a feature extraction procedure for the ultrasound images collected from lumbar spine. The important anatomical features, including epidural space, vertebra body and articular processes are extracted from the ultrasound images. Moreover, the rate of black pixels along with midline are also extracted with midline detection. Based on the features extracted from training samples and test samples, neural network is used to classify the interspinous/ bone images with maximal margin. The trained neural network model is also tested on the 43 ultrasound video streams collected from pregnant patients, and successfully identified the interspinous region / bone region on all of the videos collected.

This research is part of a bigger project which aims to insert the needle for the epidural anesthesia procedure automatically under the guidance of ultrasound imaging. This paper fulfills the purpose of automatic interpretation and identification of ultrasound images, so that anesthetists are relieved from reading raw ultrasound images. It also proves that the proposed algorithms are fast enough for real-time video processing. In a future work of this research, the algorithm will be implemented in real-time manner to detect epidural needle insertion site.

## References

1. Rawal, N.: Reg. Anesth. Pain Med. 37(3), 310–317 (2012)
2. Osterman, M.J.K., Martin, J.: Epidural and Spinal Anesthesia Use During Labor: 27-State Reporting Area, Centers for Disease Control and Preventnion (2008)
3. Le Coq, G., Ducot, B., Benhamou, D.: Risk factors of inadequate pain relief during epidural analgesia for labour and delivery. Can. J. Anaesth. 45(8), 719–723 (1998)

4. Paech, M.J., Godkin, R., Webster, S.: Complications of obstetric epidural analgesia and anaesthesia: a prospective analysis of 10,995 cases. Int. J. Obstet. Anesth. 7(1), 5–11 (1998)
5. La Grange, P., Foster, P.A., Pretorius, L.K.: Application of the Doppler ultrasound bloodflow detector in supraclavicular brachial plexus block. Br. J. Anaesth. 50(9), 965–967 (1978)
6. Grau, T., Leipold, R.W., Conradi, R., Martin, E., Motsch, J.: Efficacy of ultrasound imaging in obstetric epidural anesthesia. J. Clin. Anesth. 14(3), 169–175 (2002)
7. Ecimovic, P., Loughrey, J.: Ultrasound in obstetric anaesthesia: a review of current applications. Int. J. Obstet. Anesth. 19(3), 320–326 (2010)
8. Noble, J.A., Navab, N., Becher, H.: Ultrasonic image analysis and image-guided interventions. Interface Focus 1(4), 673–685 (2011)
9. Tran, D., Rohling, R.: Automatic detection of lumbar anatomy in ultrasound images of human subjects. IEEE Trans. Biomed. Eng. 57(9), 2248–2256 (2010)
10. Kerby, B., Rohling, R., Nair, V., Abolmaesumi, P.: Automatic identification of lumbar level with ultrasound. In: Conf Proc. IEEE Eng. Med. Biol. Soc., pp. 2980–2983 (2008)
11. Al-Deen Ashab, H., Lessoway, V.A., Khallaghi, S., Cheng, A., Rohling, R., Abolmaesumi, P.: An augmented reality system for epidural anesthesia (AREA): prepuncture identification of vertebrae. IEEE Trans. Biomed. Eng. 60(9), 2636–2644 (2013)
12. Yu, S., Tan, K.K., Shen, C.Y., Sia, A.: Ultrasound Guided Automatic localization of needle insertion site for epidural anesthesia. In: Proceeding of IEEE International Conference on Mechatronics and Automation, pp. 985–990 (2013)
13. Lee, Y., Tanaka, M., Carvalho, J.: Sonoanatomy of the lumbar spine in patients with previous unintentional dural punctures during labour epidurals. Reg. Anesth. Pain. Med. 33(3), 266–270 (2008)
14. Carvalho, J.C.: Ultrasound-facilitated epidurals and spinals in obstetrics. Anesthesiol Clin. 26(1), 145–158 (2008)
15. Yu, S., Tan, K.K., Sng, B.L., Li, S.J., Sia, A.: Automatic identification of needle insertion site in epidural anesthesia with a cascading classifier. Ultrasound Med. Biol. (in press)
16. Maiorov, V., Pinkus, A.: Lower bounds for approximation by MLP neural networks. Neurocomputing 25(1), 81–91 (1999)
17. Hecht-Nielsen, R.: Theory of the backpropagation neural network. In: International Joint Conference on Neural Networks, IJCNN, pp. 593–605. IEEE (1989)
18. Haykin, S.: Neural networks and learning machines (vol. 3). Pearson Education, Upper Saddle River (2009)
19. Ding, S., Li, H., Su, C., Yu, J., Jin, F.: Evolutionary artificial neural networks: a review. Artif. Intell. Rev. 39, 251–260 (2013)
20. Alba, E., Chicano, J.F.: Training Neural Networks with GA Hybrid Algorithms. In: Deb, K., Tari, Z. (eds.) GECCO 2004. LNCS, vol. 3102, pp. 852–863. Springer, Heidelberg (2004)
21. David, J., Goadrich, M.: The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning, pp. 233–240. ACM (2006)