

Automatic Localization of the Needle Target for Ultrasound-Guided Epidural Injections

Mehran Pesteie, Victoria Lessoway, Purang Abolmaesumi, and Robert N. Rohling

Abstract—Accurate identification of the needle target is crucial for effective epidural anesthesia. Currently, epidural needle placement is administered by a manual technique, relying on the sense of feel, which has a significant failure rate. Moreover, misleading the needle may lead to inadequate anesthesia, post dural puncture headaches, and other potential complications. Ultrasound offers guidance to the physician for identification of the needle target, but accurate interpretation and localization remain challenges. A hybrid machine learning system is proposed to automatically localize the needle target for epidural needle placement in ultrasound images of the spine. In particular, a deep network architecture along with a feature augmentation technique is proposed for automatic identification of the anatomical landmarks of the epidural space in ultrasound images. Experimental results of the target localization on planes of 3-D as well as 2-D images have been compared against an expert sonographer. When compared with the expert annotations, the average lateral and vertical errors on the planes of 3-D test data were 1 and 0.4 mm, respectively. On 2-D test data set, an average lateral error of 1.7 mm and vertical error of 0.8 mm were acquired.

Index Terms—Epidural injection, deep learning, 3D ultrasound, prepuncture scan, target localization.

I. INTRODUCTION

EPIDURAL needle injection is a commonly used procedure in obstetrics [1], [2] and chronic pain treatment [3], [4]. An epidural injection is administered by placing a needle into the epidural space between ligamentum flavum and dura mater, while the patient is either in a sitting or lying position with the back arched, in order to increase the intervertebral gap. Studies indicate that neither position is superior to the other in terms of efficacy, procedure

time and patient's comfort level [5], [6]. Traditional epidurals are either guided by palpation and a technique called loss-of-resistance [7], or under fluoroscopy [8]. The conventional loss-of-resistance technique has a failure rate within 6-20% [9]. Failure is referred to as inadequate or no pain relief after the procedure and is often due to needle misplacement. In some cases, misplacement via overshoot of the needle tip past the epidural space and puncture of the dura mater, can lead to headache and other complications. Moreover, fluoroscopy has significant drawbacks such as exposing the patient to ionizing radiation, which prohibits its application on pregnant patients for epidural anesthesia during labor and delivery.

Alternatively, ultrasound (US) has been recently used to facilitate spinal needle injections and lumbar epidural anesthesia and the clinical feasibility of using US-guided injections have been investigated [10]–[14]. However, when compared with other imaging techniques, US images are often difficult to interpret since the complexity of the anatomy and the images are usually affected by speckle noise [15], acoustic clutter [16], reverberation artifacts and shadowing [17]. Although improvements continue to be made to the beam forming and image filtering algorithms in order to improve the overall image quality, image interpretation remains a key challenge and has a steep learning curve for novices. Current US-guided injection systems therefore offer limited guidance to the physician, prior to the needle insertion for accurate localization of the needle target, and during the procedure for guiding the needle. A system is needed to automatically identify and localize the needle target within the US images of the spine. Such a guidance system has potential benefits to the patient such as increased use and effectiveness of analgesia, and a reduction in the frequency and severity of associated complications. In particular, the goal is to allow any operator, including the novice ultrasound user to correctly identify the target in all patients. Studies have shown that ultrasound guidance reduces the learning time for novices [18], [19] and the number of needle insertions needed before the target is reached [20]. Therefore, the expected benefit of this work is to give successful anesthesia to a larger number of patients through the complication of more accurate needle targeting and operator confidence and success rate. The objective of this study is to develop a machine learning technique which helps the operator with accurately localizing the needle target in US images prior to, or during needle insertion.

Manuscript received May 19, 2017; revised July 5, 2017 and August 8, 2017; accepted August 9, 2017. Date of publication August 11, 2017; date of current version December 29, 2017. This work was supported in part by the Natural Science and Engineering Research Council of Canada and in part by Canadian Institutes of Health Research. (Corresponding author: Mehran Pesteie.)

M. Pesteie and P. Abolmaesumi are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: mehranp@ece.ubc.ca).

V. Lessoway is with the Department of Ultrasound, Women's Hospital, Vancouver, BC V6H 3N1, Canada.

R. N. Rohling is with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada, and also with the Department of Mechanical Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2017.2739110

A. Related Prior Studies

Previous works on using US guidance for spinal injections can be divided into two major categories:

- 1) clinical feasibility studies, which investigate the accuracy and success rate of US-guided injections using a standard US system,
- 2) advanced computer assisted systems, which propose image processing, computer vision and machine learning techniques to provide the physician with additional information on the standard US images.

1) Clinical Feasibility Studies: US has been clinically investigated for needle guidance and the accuracy and efficacy of the US-guided injections have been compared with those of fluoroscopy/Computed Tomography (CT) guided injections [14], [21]–[23]. Obernauer *et al.* compared US-guided and CT-controlled cervical facet joint injections [24]. Their experiments show that not only US-guided injections were faster by 6 minutes, but their therapeutic effects were similar to those of the CT-guided injections. In another study, Voloshin investigated the use of four dimensional (real-time 3D) US guidance for epidural anesthesia at lower thoracic and lumbar levels [25]. His experimental results showed that the 4D ultrasound significantly improves epidural anesthesia by providing good visibility of the bone tissue and neuroaxial structures, due to enhanced spatial orientation of the operator. Additionally, Yoon *et al.* reviewed the role of US in guided interventions for a wide range of spinal injections. Their study showed that the success rate of the US-guided spinal injections highly depends on the expertise level of the sonographer, which often requires a long learning curve [26].

In another study, Grau *et al.* performed US scans of the lumbar spine and employed three different US planes, namely, the transverse, median and paramedian longitudinal planes. Their study showed that, because of providing better visibility of the anatomical features and the needle target, the paramedian plane is the optimal plane for US images of the lumbar anatomy [28]. Also, in a recent study, Srinivasan *et al.* [29] compared the conventional palpation based midline injections versus US-guided paramedian needle insertion. In their study, one hundred patients were consented and were randomized into group C (conventional) and group P (preprocedural ultrasound-guided) with 50 patients in each group. Their results showed that the average number of needle insertion attempts in group P was significantly lower than those of group C. However, landmark identification in paramedian US took a significant amount of time which slowed down the procedure.

Alongside the studies on the clinical feasibility of US-guided injections, prepuncture US scans have been also investigated to identify the vertebral levels and accordingly, identify the proper needle puncture site [30], [31]. Generally, clinical studies show that US-guided needle insertions are feasible [30]–[33], however, the success rate of such procedures depends on the experience and the skill level of the anesthesiologist [34], [35].

2) Computer-Assisted Systems: Several new US guidance systems have been proposed to visualize the spinal anatomy

during the needle insertion. Ungi *et al.* [36] proposed a navigation system using an electromagnetically tracked transducer and needle. Their proposed system records a sequence of US snapshots and registers the spatial coordinates of the tracked needle to the coordinates of the snapshot. Their experiments on cadaveric lamb model showed a significant improvement in success rate of the facet joint injections. In another study, Chen *et al.* augmented the reconstructed three dimensional (3D) ultrasound volume with a preoperative 3D CT volume, which provides the physician with a two dimensional x-ray vision of the anatomy while inserting the needle [37]. Also, Rafii-Tari *et al.* proposed an ultrasound guidance system which uses a transducer-mounted camera to detect a unique marker strip attached to the skin and accordingly, estimate the position of the US images with respect to the patient [38]. Their proposed system creates 3D panoramic volumes of the spine which can be used to visualize the spinal anatomy during needle insertions.

Statistical shape models have also been proposed to provide a better visualization of the spinal anatomy which helps the anesthesiologist with image interpretation [39]–[42]. For example, Rasoulian *et al.* [43] introduced a navigation system which registers a statistical model of the spine to a set of US images. Their proposed method uses an electromagnetically tracked transducer and needle for real-time guidance. Later, Brudfors *et al.* developed a real-time system that can continuously register the statistical model to the reconstructed 3D US volume without the need for an electromagnetic tracker [44].

Template matching along with machine learning algorithms have been investigated to help with US image interpretation. In a recent study, Yu *et al.* proposed an image classification method to identify the needle puncture site in the US images [45]. They used the transverse view of the vertebrae and template matching for identifying the target features, namely the articular process, epidural space and vertebral body. Their research group later proposed a spinal level identification system which matches a predefined template on a panoramic sagittal image of the spine [46]. Moreover, Tran and Rohling utilized a bone enhancement technique and template matching to identify the US echo from the surface of the bone and then, detect the epidural space [47]. However, since matching-based approaches use a static template for target identification, which is obtained from a subset of subjects, they have difficulty covering the full intra-patient shape variations of the spine.

3D US may facilitate improved identification of the spinal anatomy and the needle target in epidural anesthesia [48]. In order to visualize the spinal anatomy in paramedian 3D US while performing needle insertion, a novel needle guide called EpiGuide, was proposed by Malenfant *et al.* [49]. Single operator midline needle insertion is feasible using EpiGuide, since the US probe does not obscure the needle puncture site. Therefore, the physician has real-time visualization of the needle target during injection. However, due to the complex structure of the spinal anatomy and the surrounding tissue, which are usually affected by the US noise, image interpretation remains the key challenge in using 3D US for needle guidance. In a preliminary study, Pesteie *et al.* proposed a

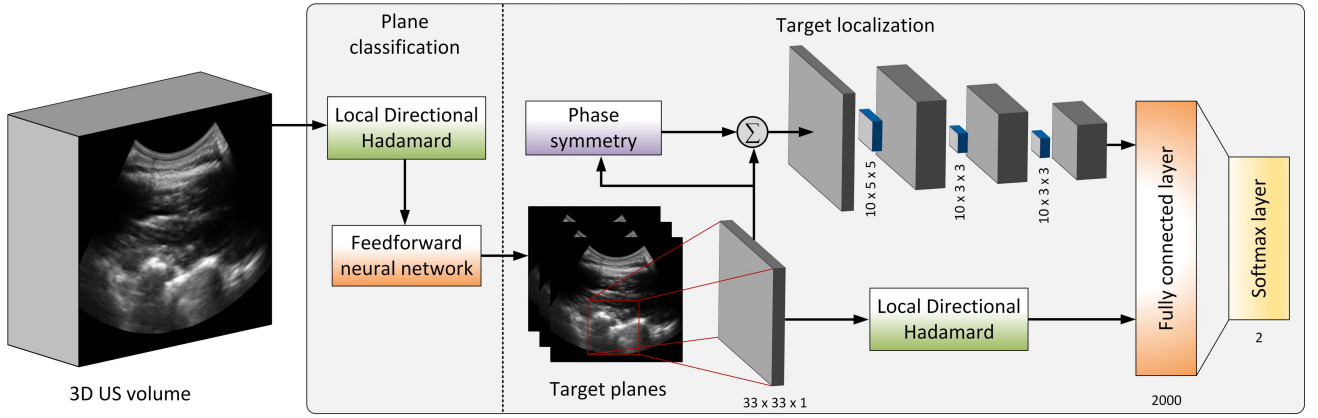


Fig. 1. Block diagram of the system. Target planes are identified from the 3D US volume using local directional Hadamard features and a fully connected neural network (from [27]). Within the target planes, the needle target is localized by classifying pixels in the image via a deep convolutional network. For each pixel of the original image, the input to the network is a patch of size 33×33 where the pixel is located at its center. Bone surfaces are enhanced in the input image patch, using a set of bandpass filters to estimate phase symmetry. The proposed network for localization consists of three convolutional layers, each followed by a 2×2 max-pooling layer. The first convolutional layer has 10×5 kernels and the other two layers have 10×3 kernels. The output of the final max-pooling layer is augmented with the local directional Hadamard features and is then given to a fully connected layer with 2000 hidden neurons, which is followed by a softmax layer.

real-time US image classification system to identify the spinal anatomy in the planes of a 3D US volume [27]. They proposed a novel feature extraction method based on the sequency-ordered Hadamard transform to identify distinct signatures of the spinal structures in the paramedian view and accordingly, identify the specific planes which contain the needle target. The operator is left with the task of localizing the target in the image and aligning with the needle target depicted on the image.

B. Contributions

Extending our previous work on automatic US image classification presented in [27], in this paper we propose a convolutional network architecture that both identifies and localizes the epidural space in paramedian US images of the lumbar spine.

We introduce a feature augmentation technique that incorporates convolutional feature maps, which are obtained by convolutional layers, with multi-scale local directional Hadamard features. Since the multi-scale Hadamard features are sensitive to the directionality of the US echoes from the surface of the vertebrae in the US image, the augmentation provides the deep network with a set of distinctive directional features from the sequency domain in addition to the feature maps automatically obtained from the spatial domain.

We demonstrate that the augmented Hadamard features are not automatically learned by a deep network with the same number of convolutional layers and kernels and hence, the augmented features are not redundant. The impact of the feature augmentation on the performance of the pixel-level classification is shown through evaluation of accuracy of the proposed network against a conventional CNN architecture with the same number of convolutional layers and kernels. Also, we demonstrate that augmenting the Hadamard features with the convolutional feature maps improves the target localization accuracy when compared to the localization results

of template matching and a state-of-the-art deep network for biomedical image segmentation.

Using features obtained from the training data along with a set of hand-crafted features for image analysis has been previously investigated by Wang et al. for breast cancer detection [50]. However, their proposed method is a cascade classifier of two sets of features. Therefore, the error of one classifier can propagate to the next level. Moreover, it does not take into account the hand-crafted features while learning the ones from training data. Hence, it is arguable that a wide range of the learned features are already extracted via the hand-engineered feature extractors.

II. METHODS AND MATERIALS

Figure 1 shows the block diagram of our proposed hybrid algorithm. We start with a brief discussion of the background of the Hadamard transform in the next section, followed by a short explanation of our feature extraction algorithm, as one of the building blocks of the target localization network, summarized from [27]. Next, we propose our Hadamard feature augmentation along with a deep convolutional network architecture for automatic target localization in US images.

A. Walsh-Hadamard Transform

As an orthogonal linear transform, the Walsh-Hadamard transform (hereafter referred to as the Hadamard transform) decomposes the input signal into rectangular wave primitives in the transform domain and is defined as [51]:

$$W(u, v) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} J(x, y) \left[(-1)^{\Psi(u, v, x, y)} \right], \quad (1)$$

where J is the image function of size $N \times N$ and Ψ indicates the parametric kernel function of the transform. In the

naturally-ordered transform Ψ is defined as:

$$\Psi(u, v, x, y) = \sum_{k=0}^{l-1} [b_k(x)b_k(u) + b_k(y)b_k(v)]. \quad (2)$$

$b_k(w)$ is the k 'th bit in the binary representation of w and l denotes the total number of bits. Since the Hadamard transform is a separable transform, one can rewrite (1) and the corresponding matrix form as:

$$\begin{aligned} W(u, v) &= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} H_N(u, x) J(x, y) H_N^T(y, v) \\ &\equiv \mathbf{W} = \mathbf{H}_N \times \mathbf{J} \times \mathbf{H}_N^T, \end{aligned} \quad (3)$$

where \mathbf{H}_N and \mathbf{H}_N^T are the Hadamard matrix of size N and its transpose, respectively, and N is an integer power of 2. \mathbf{H}_N can be defined recursively as:

$$\mathbf{H}_N = \begin{cases} \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{pmatrix} & \text{if } N \geq 2; \\ 1 & \text{if } N = 1. \end{cases} \quad (4)$$

Rows of \mathbf{H}_N correspond to basis functions, called Walsh functions, which are orthogonal and consist of binary values (± 1). Hence, the sequency domain refers to the Walsh projections of the signal via Walsh basis functions. In the sequency-ordered transform, the Walsh functions are ordered in the Hadamard matrix based on their sign change. Therefore, there is no sign change in the first row and column of the transform matrix, as opposed to the n 'th row and column which contains $n-1$ sign changes. Accordingly, the sequency-ordered transform yields a representation of the image based on the ordered variations of the pixel values along horizontal and vertical directions. Hence, the top left corner of the transform represents low variations whereas the bottom right corner contains high variations.

1) Multi-Scale and Multi-Directional Features: The laminae of the vertebrae are the prominent anatomical features of the epidural space in the paramedian US images of the spine. Because the pixel intensities of the US image is proportional to the direction of the reflected US wave from the tissue with respect to the transducer, the particular wave-like signature of the laminae of the vertebrae, can be identified via distinctive local directional features from the US echoes of the bone surfaces. Therefore using (4), we utilize a recursive partitioning (RP) algorithm, adopted from Donoho and Huo's study [52], to partition the image, compute the Walsh-Hadamard coefficients at any scale of interest and accordingly, obtain the directional features. In order to use the multi-scale features to train a classifier, we apply two slight modifications to the original RP algorithm: *a)* the dimension of the image at a given scale changes with respect to the scale of decomposition, instead of a continuum square belonging to $[0, 1]^2$, and *b)* in order to have a consistent order while partitioning the image, we use a list instead of a set. Therefore, the result of the RP algorithm represents a depth-first traversal of the partitioning tree with maximum depth δ which the root indicates the original image.

Let P_j and \mathbf{H}_N^j be a list of all image patches \mathbf{S}_m and the Hadamard transform matrix at scale j , respectively. Thus, for all $j \leq \delta$ and all $\mathbf{S}_m \in P_j$:

$$C^{P_j} = \left\{ \mathbf{H}_N^j \times \mathbf{S}_m \times \mathbf{H}_N^{jT} \right\} \quad \text{for } m = 1, \dots, |P_j|, \quad (5)$$

where C^{P_j} is the list of recursive partitioning Walsh-Hadamard coefficients (RPHC). RPHC enables one to examine the image at any level of detail in the sequency domain and also yields a local correlation between spatial and sequency domains, which is beneficial for recognizing different patterns that appear in different regions of the image.

Directional features in the sequency domain can be extracted from the US images via separately examining significant vertical, horizontal and diagonal Walsh-Hadamard coefficients. Using RPHC, orientation features can be locally extracted from the image. In our previous study, we introduced three wedge filters in the sequency domain, \mathbf{F}_{vt} , \mathbf{F}_{hr} and \mathbf{F}_{dg} to identify the significant Walsh-Hadamard coefficients in vertical, horizontal and diagonal orientations, respectively, as:

$$\begin{aligned} F_{vt}(u, v, \alpha) &= \frac{1}{2} + \frac{1}{2} \frac{|v - \frac{u\alpha}{50}|}{v - \frac{u\alpha}{50}}, \\ \mathbf{F}_{hr} &= \mathbf{F}_{vt}^T, \\ \mathbf{F}_{dg} &= \mathbf{1} - (\mathbf{F}_{vt} + \mathbf{F}_{hr}), \end{aligned} \quad (6)$$

where u and v are sequency domain indices and α is the parameter which defines the percentage of the covered area by the filter ($\alpha \leq 50$). More detail about the directional filters and estimating α can be found in [27].

At a given scale, j , the local directional Hadamard (LDH) feature vector is comprised of the variances of the nonzero elements of the filter outputs:

$$\begin{aligned} \mathbf{v}^j &= \left[\text{Var}(\mathbf{C}_k^{P_j} \circ \mathbf{F}_{vt}^j), \text{Var}(\mathbf{C}_k^{P_j} \circ \mathbf{F}_{hr}^j), \text{Var}(\mathbf{C}_k^{P_j} \circ \mathbf{F}_{dg}^j) \right] \\ k &= 1, \dots, |C^{P_j}|, \end{aligned} \quad (7)$$

where $\mathbf{C}_k^{P_j}$ is the k 'th Hadamard transformed matrix in the list of RPHC C^{P_j} , \mathbf{F}^j is the orientation filter at scale j and \circ denotes the Hadamard product.

B. Target Localization

1) Target Definition: The epidural space is superior to the upper margin of the base of the spinous process at its junction with the lamina [53]. In 2D paramedian US images of adult spine, the echo from the ligamentum flavum can be seen immediately adjacent to the anterior base of the laminae of vertebrae in paramedian plane [47] (see Figure 2). Since the location of the epidural space corresponds to that of the base of the laminae of vertebrae, the anterior base of the lamina is referred to as the "needle target" in the US images, hereafter.

2) Convolutional Neural Networks: Convolutional neural networks (CNNs) have shown noticable success in numerous cognitive tasks such as object and speech recognition [54], [55] and also medical image analysis [56], [57]. In contrast to traditional machine learning methods that require hand-engineered features, CNNs learn a set of convolutional kernels in multiple layers, which are specifically tuned for the problem. This is

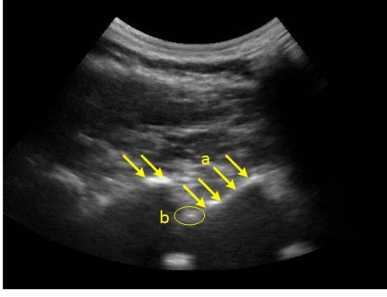


Fig. 2. Sample paramedian US image from a volunteer's (#1) laminae of the L3-L4 vertebrae at 7 cm depth. Prominent anatomical landmarks of the target for epidural insertion in the paramedian view are a) the wave-like pattern of the US echoes from the surfaces of the laminae of adjacent vertebrae and b) the echo from the ligamentum flavum, representing the target.

achieved by minimizing a loss function $\mathcal{L}(\cdot)$ of the output $f(\mathbf{X}; \mathbf{K}, \mathbf{b})$ and target t , over the convolutional kernels \mathbf{K} and biases \mathbf{b} , given input \mathbf{X} from data set \mathcal{D} :

$$\arg \min_{\mathbf{K}, \mathbf{b}} \left\{ \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} \mathcal{L} \left(f(\mathbf{X}^i; \mathbf{K}, \mathbf{b}), t^i \right) \right\}. \quad (8)$$

The hierarchical feature learning in a CNN allows it to extract representations of the data at multiple levels of abstraction. A CNN architecture may perform a downsampling operation, called “pooling,” before proceeding to the next convolutional layer, which reduces the number of weights and also provides some level of robustness to a set of transformations, including rotation and translation of the features. Therefore, a CNN can potentially cover a large variability in US appearance of the target.

3) CNN Architecture: The needle target, namely the epidural space, can be identified by 1) the particular wave-like signature of the laminae of the vertebrae in the parasagittal planes and 2) the ultrasonic echo from the ligamentum flavum (see Figure 2). A CNN architecture with feature augmentation is proposed for pixel-level classification of the US image to localize the epidural space through identification of the respective anatomical landmarks. For a given pixel p , the goal of the network is to classify the image patch centered at p , \mathbf{J}_p , and accordingly, predict whether or not p belongs to the needle target region annotated by the expert.

The proposed CNN consists of three convolutional layers. The first layer has 10 5×5 kernels and is followed by a max pooling layer with a 2×2 pool size. The next two convolutional layers have 10 3×3 kernels each and are followed by max pooling layers with the same pool size as the first pooling layer. The activation functions of the convolutional and fully connected layers are the rectifier function, which is defined as:

$$f(x) = \max(0, x). \quad (9)$$

The last convolutional layer yields feature maps which represent local dependencies of the spinal anatomy in spatial domain. However, since convolution is a spatial domain operation, multi-scale analysis of the image in the sequency domain provides a set of features which the conventional

CNN architectures with relatively small number of parameters cannot extract from the spatial domain. Therefore, augmenting convolutional feature maps with multi-scale Hadamard features provides a richer representation of the anatomical landmarks of the needle target in US images.

Let \mathbf{Y}_f be the final output feature map for image patch \mathbf{J}_p prior to the fully connected layer, i.e.:

$$\mathbf{Y}_f = \downarrow \left(\mathcal{F} \left(\sum_i \mathbf{K}_f \star \mathbf{X}^i + \mathbf{b}_f \right) \right), \quad (10)$$

where \star and \downarrow denote the 2D discrete convolution and max pooling operators, respectively, $\mathcal{F}(\cdot)$ is the non-linear activation function, \mathbf{X}^i is the i 'th feature map obtained from the previous layer and \mathbf{K}_f and \mathbf{b}_f are the trainable kernel and the bias vector, respectively. Also, let \mathbf{v}^m be the $1 \times d$ local sequency domain filter outputs of the image patch \mathbf{J}_p . We define shift matrix $\mathbf{T}_{d \times |P|d}$ as:

$$\mathbf{T}_m = [t_m(p, q)] = \begin{cases} 1 & \text{if } p - d(m-1) = q \text{ for } m \leq |P| \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where P is the list of the partitioned images. The multi-scale filter outputs are aggregated as:

$$\mathbf{f}_{LDH} = \sum_{i=1}^{|P|} \mathbf{v}^i \mathbf{T}_i. \quad (12)$$

In order to augment the convolutional feature maps $\mathbf{Y}_f \in \mathbb{R}^{M \times N \times L}$ with the sequency domain multi-scale filter outputs \mathbf{f}_{LDH} , we introduce matrices $\mathbf{\Omega} \in \{0, 1\}^{|P|d \times MNL + |P|d}$ and $\mathbf{\Lambda} \in \{0, 1\}^{MNL \times MNL + |P|d}$ as:

$$\mathbf{\Omega} = [\mathbf{I}_{(|P|d \times |P|d)} | \mathbf{0}_{(|P|d \times MNL)}], \quad (13)$$

$$\mathbf{\Lambda} = [\mathbf{0}_{(MNL \times |P|d)} | \mathbf{I}_{(MNL \times MNL)}]. \quad (14)$$

The final input to the fully connected layer is:

$$\mathbf{y} = \mathbf{f}_{LDH} \mathbf{\Omega} + \text{vec}(\mathbf{Y}_f) \mathbf{\Lambda}, \quad (15)$$

$\mathbf{I}_{N \times N}$ and $\mathbf{0}_{M \times N}$ are the identity matrix of size N and zero matrix of size $M \times N$, respectively. Also, $\text{vec}(\cdot)$ denotes the vectorization operation on a matrix.

In order to classify the augmented feature maps, a fully connected layer with 2000 hidden neurons follows the last pooling layer. As indicated in Figure 1, in the proposed architecture, the input to the fully connected layer is comprised of a) the convolutional features which are obtained via the convolution kernels from the previous layers (\mathbf{Y}_f) and b) the multi-scale and multi-directional filter outputs in the sequency domain (\mathbf{f}_{LDH}) from each image patch \mathbf{J}_p in the training set.

A training criterion for the network is to maximize the probability of the training labels or equivalently, to minimize the negative log-likelihood as the loss function. Therefore, the model learns a distribution of the convolutional and multi-scale Hadamard feature maps over the probability of the training labels. A softmax layer is used in the output to produce the class likelihood of the input.

Since the rectifier function was used as the activation of the convolutional and fully connected layers, all of the weights of the model were initialized using the approach proposed in [58].

4) *Training*: The conventional gradient descent algorithm along with the chain rule with a slight modification can be used to update the parameters of the network and accordingly, train the model. Let \mathbf{o} be the output of the softmax activation and \mathbf{z} be the input to the fully connected layer. Therefore in order to update the weights of the model using the gradient descent algorithm, the error can be back propagated via applying the chain rule to the derivative of the cost function with respect to the parameters:

$$\frac{\partial \mathcal{L}}{\partial \Theta} = \frac{\partial \mathcal{L}}{\partial \mathbf{o}} \times \frac{\partial \mathbf{o}}{\partial \mathbf{z}} \times \frac{\partial \mathbf{z}}{\partial \mathbf{y}} \times \frac{\partial \mathbf{y}}{\partial \Theta}, \quad (16)$$

where Θ is the parameter set of the model before the fully connected layer. Hence, Θ can be decomposed into two sets: $\Theta = [\Theta_{LDH}; \Theta_{CNN}]$ where Θ_{LDH} and Θ_{CNN} correspond to the parameter set which is associated with the local directional Hadamard and convolutional features, respectively. (16) can be rewritten based on (15) as:

$$\frac{\partial \mathcal{L}}{\partial \Theta} = \frac{\partial \mathcal{L}}{\partial o_i} \times \frac{\partial o_i}{\partial \mathbf{z}} \times \frac{\partial \mathbf{z}}{\partial \mathbf{y}} \times \frac{\partial (\mathbf{f}_{LDH}\Omega + \text{vec}(\mathbf{Y}_f)\Lambda)}{\partial \Theta}. \quad (17)$$

Finally, since only Θ_{LDH} and Θ_{CNN} contribute to the error originated from \mathbf{f}_{LDH} and \mathbf{Y}_f , respectively, we have:

$$\frac{\partial (\mathbf{f}_{LDH}\Omega + \text{vec}(\mathbf{Y}_f)\Lambda)}{\partial \Theta} = \frac{\partial \mathbf{f}_{LDH}\Omega}{\partial \Theta_{LDH}} + \frac{\partial \text{vec}(\mathbf{Y}_f)\Lambda}{\partial \Theta_{CNN}} \quad (18)$$

The backpropagation algorithm continues to update the convolutional kernels and biases by applying the chain rule to $\frac{\partial \text{vec}(\mathbf{Y}_f)\Lambda}{\partial \Theta_{CNN}}$, as in a conventional CNN.

Due to high capacity, CNNs are often vulnerable to overfitting where there is not enough variability in the training data. Accordingly, regularization is critical in obtaining a generalizable model. In the proposed CNN architecture, the absolute value of the weights were bounded by the ℓ_2 norm to prevent overfitting. Moreover, the dropout technique was used to stochastically add noise in the computation of the network [59]. Dropout encourages the network to learn features that are independently useful, since each neuron activation is masked to zero by a certain probability. During training, the ℓ_2 regularization parameter was set to 0.002 and after each convolutional layer a dropout layer with $P = 0.25$ was added.

C. Image Enhancement

In order to enhance the bone surfaces and accordingly, the wave-like pattern of the laminae, a set of quadrature band-pass filters were used, as proposed by Anas *et al.* [60], to estimate the phase symmetry, which is maximized at bone locations. Also to exploit the echo from the ligamentum flavum, the enhanced image was augmented by a weighted sum of the original US image, which contains the echo, and the enhanced image as:

$$\text{augmented_image} = \alpha \times \text{enhanced} + (1 - \alpha) \times \text{US},$$

where α is the weight of the bone enhanced image and was set to 0.7 based on trial and error. Figure 3 shows a sample US target plane and the bone enhanced image.

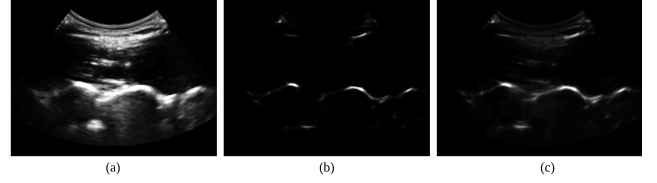


Fig. 3. Original paramedian US image of a randomly selected volunteer's (#2) laminae of the L3-L4 vertebrae at 7 cm depth (a), the bone enhanced image using phase symmetry (b) and the weighted sum of the enhanced and the original image (c).

D. Data Collection

Following written consent, three separate paramedian US data sets of the lumbar spine were collected by an expert sonographer using a SonixTOUCH US machine (Analogic Corp. Richmond, Canada):

- (A) one hundred and four 3D US volumes from thirteen volunteers (8 volumes from each side of the spine with subtle movement of the US probe) at 7 cm depth. The gain and field of view were set to 50% and 43.9°, respectively (ethics certificate number: H07-0691),
- (B) twenty 3D US volumes from 10 different volunteers (one volume from each side of the spine) with various depth, gain and field of view optimized by the sonographer for each volunteer (ethics certificate number: H15-01310) and
- (C) ten sequences of 2D US images from 10 different volunteers 7 cm depth and 50% gain (ethics certificate number: H15-01310).

All of the participants were aged between 20 and 35 years whose body mass indices (BMI) were less than 30 with no records of spinal pathology or surgery.

3D US data sets (A and B) were collected with a 5 MHz 3D motorized transducer from L3-L4 vertebrae of the spine and the transducer was placed laterally within 10 mm of the midline and angled by 5° to 10°. Therefore the paramedian volumes contained the vertebral structures and particularly, the laminae of vertebrae which is the prominent anatomical landmark of the epidural space, was present in the volumes. The L3/L4 vertebral level was scanned by identifying the sacrum and counting up to the L3 and L4 vertebrae. Each US plane was annotated by an expert sonographer as either a target plane for epidural needle insertion or nontarget plane. The labeling procedure was repeated 36 weeks after the initial annotation by the same sonographer. The average Dice similarity coefficient [61] of the corresponding labels was 0.92. Based on the sonographer's annotations, 545 and 150 paramedian target planes were extracted from data set A and B, respectively.

2D US sequences in (C) were captured with a C5-2/60 curvilinear transducer at 3.3 MHz from the laminae of the vertebrae. The sequences started from the L4 and ended at the L3 vertebra. Due to high correlation of the subsequent frames, the captured 2D sequences were sub-sampled every 5 frames. In total, 300 images were extracted from the sequences (30 from each participant).

TABLE I
SUMMARY OF THE EVALUATION METRICS AND THEIR DESCRIPTION

Category	Metric	Description
Pixel-level classification	Accuracy	Correct classification rate on target and non-target pixels
	Sensitivity	Correct classification rate on target pixels
	Specificity	Correct classification rate on non-target pixels
	Normalized cross-correlation	Similarity between two image patches
Target localization	Lateral error	Error in x direction (mm)
	Vertical error	Error in y direction (mm)

E. Data Preparation

The proposed network was trained and validated on data set *A* and tested on data sets *B* and *C*. For pixel-level classification, image patches of size $M \times M$ centered at the expert's annotation, were extracted from the US images. The patch size should be chosen with care, in order to take into account the contextual information and local dependencies of the anatomical landmarks i.e. relative location of the target with respect to the laminae of the vertebrae. Following the definition of the needle target in Section II-B.1, the location of the bases of the laminae were annotated on the image as the needle target. Target annotation was repeated by the same sonographer. The average distance between the corresponding annotations were 0.4 mm in lateral and 0.3 mm in vertical directions. In order to generate sufficient training samples, the annotated targets was transformed from a point to a region via a box of the size 25×40 pixels. Pixels of the target and non-target regions were used to train the network.

Originally, the images contain more than 170,000 pixels each, therefore, training the network on the original planes would be computationally intensive and practically impossible. Thus, after augmenting the native planes with the bone enhanced images, they were resized down to 100×100 . The sliding window size M was set to 33 to guarantee an overlap of at least $\frac{1}{3}$ of the image size between the window and the image. Also, the image was zero padded to let the sliding window be located at the borders of the image. The CNN training data set consisted of 501,298 image patches including equal number of target and non-target images, extracted from the optimal planes. The network was also validated on the balanced leave-one-out cross-validation data.

F. Implementation

The proposed algorithm for target localization was implemented in Theano [62] and trained on a GeForce(R) GTX 980 Ti GPU with 6 GB of memory, hosted by a machine running Ubuntu 14.04 operating system on a 3.4 GHz Intel(R) CoreTM i7 CPU with 16 GB of memory.

III. EXPERIMENTS AND RESULTS

Multiple experiments were conducted to evaluate the performance of the proposed method. As an intermediate step to target localization, the pixel-level classification performance was investigated on equal number of target and randomly selected non-target patches. Then, target localization accuracy was measured by comparing the expert annotations against the detected target obtained by classifying all of the pixels of the

image in validation/test data sets. Table I indicates a summary of the evaluation metrics which have been used throughout the experiments.

A. Pixel-Level Classification Performance

For each image patch, LDH features were extracted and aggregated using (12). The final feature map of the convolutional layer was augmented with the corresponding LDH features using (15). Due to the large size of the data set, the training phase was repeated 10 times on 50,000 image patches which were randomly selected from the data set. Classification accuracy is defined as $\frac{T_t}{N_t}$ where T_t and N_t correspond to total number of correct classifications and total number of pixels, respectively. The average classification accuracy of the proposed CNN with feature augmentation was 96% in the final training step.

Different experiments were conducted to evaluate the effect of the augmented features and the performance of the proposed network. First, in order to verify whether or not the proposed network can generalize to cover the inter-patient variabilities in shape and size of the vertebrae, a leave-one-out cross-validation was performed on the training data (data set *A*). The average pixel classification accuracy during the leave-one-out cross-validation was 94% with 93% sensitivity and 98% specificity. Sensitivity and specificity are defined as $\frac{T_p}{N_p}$ and $\frac{T_n}{N_n}$, respectively, where T_p is the number of correctly classified target pixels, T_n is the number of correctly classified non-target pixels, N_p is the number of target pixels and N_n is the number of non-target pixels. Therefore, sensitivity corresponds to the classification performance of the method on target pixels whereas specificity corresponds to the classification performance on non-target pixels.

B. Performance Evaluation of the Augmented LDH Features

In order to investigate the impact of the LDH features on the accuracy of the pixel-level classification, another CNN with the same number of convolutional layers, kernels and hidden neurons in the fully connected layer with no feature augmentation was trained on the same data set (experiment 1). Experimental results indicated that the CNN without feature augmentation successfully classified 90% of the pixels during leave-one-out cross-validation. In order to investigate the correlation of the CNN feature maps and the LDH features, pairwise correlations between the LDH features and CNN feature maps were calculated at all of the convolutional layers (experiment 2). Figure 4 shows the histograms of

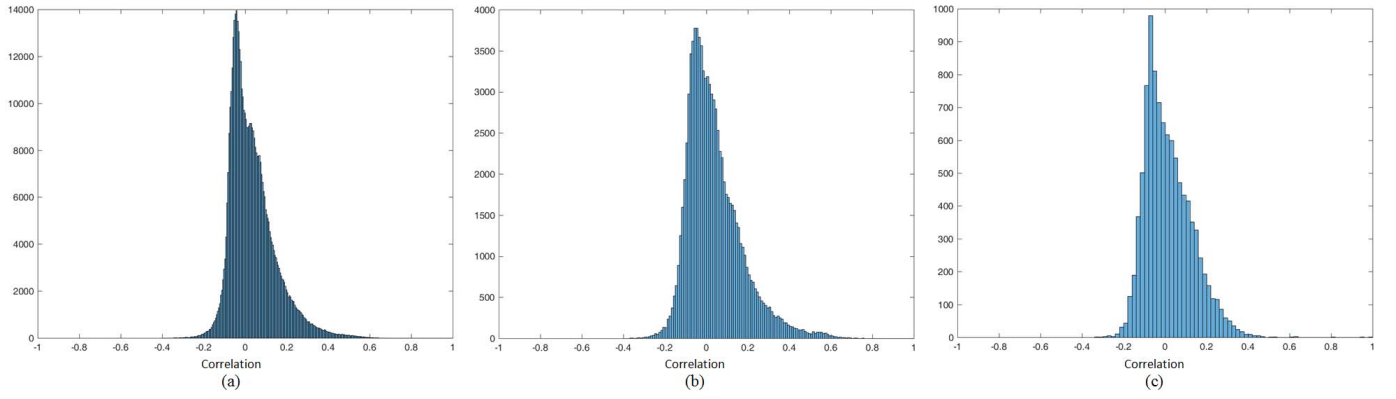


Fig. 4. Histogram of the correlation between LDH features and a) first (mean 0.03 ± 0.11) b) second (mean 0.04 ± 0.13) and c) third (mean 0.01 ± 0.11) convolutional feature maps of a similar CNN architecture with no feature augmentation after training (t-test p -value < 0.05).

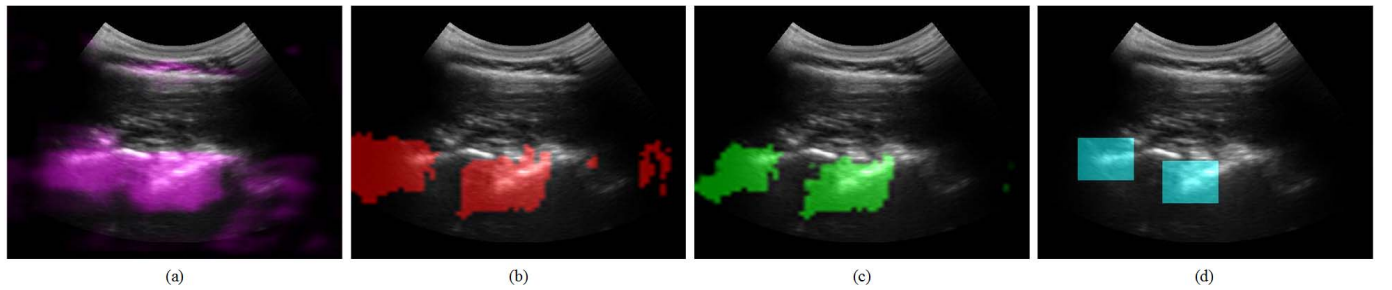


Fig. 5. Pixel-level classifications of a) LDH features, b) CNN with no augmentation c) CNN augmented with the LDH features at the fully connected layer and d) ground truth relative to the original paramedian US image of the laminae of the L3-L4 vertebrae of a randomly selected volunteer (#3) at 7 cm depth. The LDH features are not distinctive enough alone to distinguish the area between target regions. However, they contribute to the CNN by removing some outliers in (c).

the pairwise correlations, which indicate no significant linear correlation between the CNN feature maps and the LDH features, since the means are all smaller than 0.05 (t-test p -value < 0.05). Therefore, the CNN has not learned any variation of feature maps which can replicate the behavior of the LDH. However, since the classification performance has been improved with the LDH features, there is a nonlinear correlation between the LDH and CNN feature maps which has been extracted by the nonlinearities in the fully connected layer of the proposed network.

In the next experiment, the weights of the fully connected layer which correspond to the LDH features were compared against those which correspond to the CNN feature maps (experiment 3). The average dot product of the weights of the fully connected layer and their associated feature maps (LDH or CNN) were monitored. After 200 epochs, the average dot products were 0.018 for CNN and 0.0175 for LDH features, which implies that the network has not neglected the LDH features during training. However, the LDH features were not distinctive enough for pixel-level classification. This has been investigated through training a fully connected neural network with the same number of neurons in the hidden layer as in the proposed CNN, on the LDH features (experiment 4). The average classification accuracy of the LDH features during leave-one-out cross-validation was 89% and the corresponding sensitivity and specificity was 83% and 90%, respectively. Figure 5 shows sample pixel-level

classification results of the CNN with feature augmentation, without augmentation, only with the LDH features and the gold standard. As shown in Figure 5, multi-scale and multi-directional features contribute to the classification accuracy with removing some outliers. This was verified on the leave-one-out cross-validation data for all volunteers (see Figure 6).

C. Target Localization Accuracy

The expert annotations were used to evaluate the target localization accuracy of the proposed method. Since the annotated points were transformed to regions of interest i.e., a box centered at the needle target annotated by the expert (see Section II-E), the CNN tries to obtain the target regions from the images. The localization error was measured by first, classifying all of the pixels of the image and obtaining target regions, accordingly. Then, the centroids of the regions were extracted and the distance between the expert annotations and the extracted targets was measured for all of the images in validation/test data sets. The centroid of the region which is close to the midline of the image was considered as the clinically significant target for epidural needle insertion. The average error was 0.75 mm and 0.38 mm in lateral and vertical directions, respectively for the leave-one-out cross-validation.

To evaluate the target localization of the trained network on unseen data with different acquisition parameters, the trained network was tested on the target planes of data set B. Also,

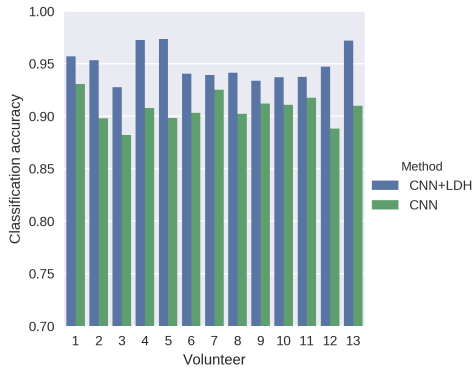


Fig. 6. Pixel-level classification performance of the proposed method (CNN + LDH) and the network without feature augmentation (CNN) on the leave-one-out cross-validation balanced data. LDH features contribute to the classification accuracy for all of the cases.

in order to investigate the application of the proposed network on standard 2D US systems, the trained network was tested on data set *C*. The average localization error for planes of 3D volumes was 1 mm and 0.4 mm in lateral and vertical orientations, respectively (data set *B*). Also, the average localization error on the test data set of 2D images was 1.7 mm in lateral and 0.8 mm in vertical directions (data set *C*). Moreover, target localization accuracy of a similar CNN without feature augmentation was investigated. The average localization error for data set *B* was 8 mm and 6.2 mm and for data set *C* was 2.6 mm and 4.5 mm in lateral and vertical directions, respectively.

D. Comparative Study

The proposed method was compared against two other previously proposed approaches:

- 1) a static template matching and image processing approach, proposed by Tran and Rohling [47] and,
- 2) a state-of-the-art deep convolutional network for biomedical image segmentation, called “U-net,” proposed by Ronneberger *et al.* [63].

For template matching, the bone surfaces of the US images were enhanced using a set of bandpass filters in frequency domain, as suggested by Tran and Rohling. Later, the LF template was generated using the same parameters as in their study, and the region with maximum cross correlation score was annotated as the LF on the image. The results of template matching (TM) were compared against the expert annotations. Next, the U-net architecture was implemented using Keras deep learning framework¹ and was trained on the same training data i.e., bone-enhanced augmented images from data set *A*. The weights of the U-net were initialized based on He’s method, as suggested by Ronneberger *et al.* [63]. The same approach was also used to initialize the weights of the proposed method. Since the U-net architecture is designed for end to end segmentation, its performance was evaluated using the Sørensen Dice coefficient. U-net achieved 81% dice coefficient on leave-one-out cross-validation. Figure 7 shows a

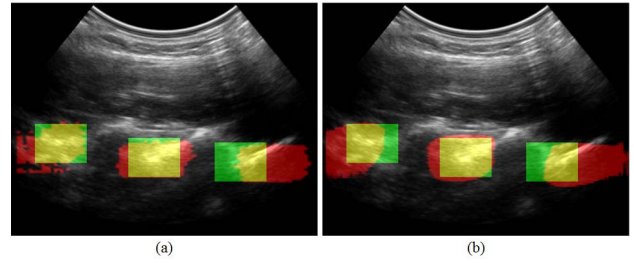


Fig. 7. Sample segmentation result of a) the proposed network and b) U-net on a randomly selected volunteer’s (#4) image from data set *B* at 8 cm depth. The yellow and red regions correspond to the correct and incorrect segmentations, respectively. The green area indicates parts of the ground truth which have not been covered by the network.

sample segmentation result of U-net along with the proposed method. Target localization error of template matching and U-net were investigated on data sets *B* and *C*. Template matching on the bone-enhanced images indicated an average error of 2.6 mm (lateral) and 2.1 mm (vertical) on the 3D test data set (*B*) and 7.8 mm (lateral) and 3 mm (vertical) on the 2D test data (*C*), whereas U-net localized the target with the average error of 1.4 mm and 0.5 mm for lateral and vertical directions, respectively on the test data set *B*. Also, on the 2D test data set, U-net target localization error was 1 mm in both lateral and vertical directions. Target localization errors on data sets *B* and *C* in lateral and vertical directions are summarized in Table II. Since the group variances are non-homogeneous, Kruskal-Wallis analysis of variance along with Nemenyi post hoc tests were performed on target localization errors, in order to test the statistical significance of the localization results. The Kruskal-Wallis H-test on both 3D and 2D data sets indicated that the population median of all of the groups are significantly different ($p < 0.05$). Moreover, the Nemenyi post hoc tests on 2D data set indicated that the localization errors are significantly different between all pairwise combinations. On 3D test data, the Nemenyi test failed to reject the null hypothesis on the errors of U-net and the proposed method with $p < 0.05$. However, the test indicated there is a significant difference between the errors of the proposed method and the original CNN without feature augmentation on 3D test data. Therefore, the augmented LDH features significantly contribute to the accuracy of target localization.

E. Execution Time

The execution time of the algorithm on the same environment described in Section II-F was measured for 100 images, which were randomly selected from data sets *B* and *C* (50 from each set). On the described setup, the average localization time per image was 1.34 seconds.

IV. DISCUSSION

Although the target localization error is correlated with pixel-level classification accuracy, accurate target localization can still be achieved without 100% pixel classification accuracy, if the misclassified pixels do not change the location of the centroid of the target region. Therefore, another metric which shows the performance of the proposed CNN is

¹<https://github.com/fchollet/keras>

TABLE II

COMPARISON OF THE TARGET LOCALIZATION ERRORS (MM) OF TEMPLATE MATCHING APPROACH (TM), CONVOLUTIONAL NEURAL NETWORK WITH NO FEATURE AUGMENTATION (CNN), STATE-OF-THE-ART DEEP NETWORK FOR BIOMEDICAL IMAGE SEGMENTATION (U-NET) AND THE PROPOSED METHOD (CNN + LDH) ON TEST DATA SETS. THE ERRORS OF THE PROPOSED METHOD ARE HIGHLIGHTED VIA ASTERISK WHERE THEY ARE SIGNIFICANTLY LOWER COMPARED TO OTHER APPROACHES (p - Value < 0.05). MINIMUM AND MAXIMUM ERRORS HAVE BEEN INDICATED IN BRACKETS. THE CLINICALLY ACCEPTABLE LATERAL AND VERTICAL ERRORS ARE 7 mm AND 5 mm

Dataset	Direction	TM	CNN	U-net	CNN+LDH
B	Lateral	$2.6 \pm 2.2 \in [0, 9.2]$	$8 \pm 6.6 \in [0, 36.9]$	$1.4 \pm 0.7 \in [0, 3.3]$	$1 \pm 0.7^* \in [0, 3.4]$
	Vertical	$2.1 \pm 2 \in [0, 8.7]$	$6.2 \pm 6 \in [0, 24]$	$0.5 \pm 0.4 \in [0, 2.2]$	$0.4 \pm 0.3 \in [0, 2]$
C	Lateral	$7.8 \pm 4 \in [0.3, 23.8]$	$2.6 \pm 2.6 \in [0, 22.8]$	$1 \pm 0.7 \in [0, 3.7]$	$1.7 \pm 1 \in [0, 5.2]$
	Vertical	$3 \pm 2.3 \in [0, 12]$	$4.5 \pm 4.7 \in [0, 23.5]$	$1 \pm 0.8 \in [0, 3.6]$	$0.8 \pm 0.6^* \in [0, 3.2]$

the normalized cross-correlation of the misclassified image patches and the closest patch which belongs to the region of the needle target, which was 81% on average for both the 2D and the planes of the 3D data set. This indicates a high similarity between the misclassified patches and the closest patch belonging to the target region. Therefore, the majority of the misclassified patches are located close to the border of the regions.

The clinically accepted error for epidural needle placement is defined by the width and the thickness of the ligamentum flavum to avoid overshoot. According to the literature, the width of the ligamentum flavum is approximately 7 mm [64]. The thickness of the ligamentum flavum is approximately 5 mm [65] but thickness decreases as the ligament is stretched, e.g. thickness was 2.6 mm on average in B-mode ultrasound images of parturients with spinal flexion [47]. Because of the complications of the needle overshoot, minimizing the vertical localization error is clinically more significant compared to the lateral error. The vertical localization error of the proposed method and U-net is below 5 mm in both data sets *B* and *C* for all of the test images. Also, both U-net and the proposed method successfully localized the epidural space with less than 2.6 mm error in all of the planes of the 3D data set (*B*). However, the vertical localization errors of U-net and the proposed method were above 2.6 mm (max. 3.6 mm vs. max. 3.2 mm) in 34 and 14 images out of 300 in *C* (11% vs. 5%), respectively. Therefore, not only the proposed algorithm has fewer outliers in localizing the needle target compared to U-net, but also the maximum vertical error of the proposed method is slightly less than U-net. Moreover, the proposed method outperforms U-net and the template matching approach in estimating the depth of the needle target in all of the test data, since it indicates smaller mean vertical error in data sets *B* and *C* (see Table II).

Since efficient implementation of the proposed method has not been the focus of this work, the current implementation does not run in real time (see Section III-E). However, because classifying a given pixel of an image is independent of classifying other pixels, a concurrent parallel implementation of the method would highly reduce the execution time. Alongside concurrency, utilizing additional techniques, such as cropping the image based on anatomy can reduce the target localization time. Therefore, the proposed method can potentially be used in real-time applications.

Lastly, US image interpretation and localization of the needle target is just one of the steps of an effective anesthesia.

The final steps involve accurate guidance of the needle towards the target and delivering the anesthetic, while trying to avoid extra needle insertion attempts. Therefore, although the proposed method helps the operator with image interpretation and target localization, accurate guidance of the needle is also required for a successful epidural anesthesia. Several guidance systems or techniques are feasible, including the commercial electromagnetic needle tracking system for freehand needle guidance (SonixGPS, Analogic, Richmod, Canada) or EpiGuide [49], which is designed and optimized specifically for epidurals i.e., can be used along with the loss-of-resistance technique with zero force needle release while providing the needle trajectory as an overlay. Given this paper's focus on 3D US, EpiGuide could be used directly with the proposed target localization method so that the multi-planar reformatted view used by EpiGuide shows the target. Moreover, 2D US could also be used with a standard needle guide [66] in the paramedian plane. Therefore, the proposed method can be combined with a guidance system for accurate target localization and needle guidance in both 2D and 3D US-guided epidurals. The output of the proposed method should be regarded as "additional information" to increase the confidence level of the operator for target identification and localization. As operator experience with the system and interpretation of the target appearance in US grows, the operator is expected to be able to identify the outliers from the algorithm and simply recognize the target.

V. SUMMARY AND CONCLUSION

In this paper, we proposed a hybrid machine learning algorithm to localize the epidural space in the paramedian US images of the lumbar spine. In particular, we proposed a convolutional network architecture along with a feature augmentation technique to localize the needle target in the target planes. This is achieved through pixel-level classification of the target planes. We showed there is no linear correlation between the augmented LDH features obtained from the sequency domain and the convolutional feature maps. Our experimental results showed the LDH features from the sequency domain provide multi-scale representations of the data which are non-linearly correlated with the convolutional feature maps, hence they contribute to the performance of the pixel-level classification. We validated the accuracy of our proposed method using leave-one-out cross-validation and also tested on different unseen data sets.

ACKNOWLEDGMENT

The authors would like to thank Emran Abu Anas, Ilker Hacihaliloglu and Denis Tran for generously sharing the implementations of their methods. They also thank Jorden Hetherington for his kind and generous help in data collection.

REFERENCES

- [1] D. Agrawal, B. Makhija, and S. Fotedar, "Impact of epidural analgesia on labour: A review," *Ann. Woman Child Health*, vol. 1, no. 1, pp. 1–9, 2015.
- [2] H. K. Eltzschig, E. S. Lieberman, and W. R. Camann, "Regional anesthesia and analgesia for labor and delivery," *New England J. Med.*, vol. 348, no. 4, pp. 319–332, 2003.
- [3] M. H. Andreae and D. Andreae, "Regional anaesthesia to prevent chronic pain after surgery: A cochrane systematic review and meta-analysis," *Brit. J. Anaesthesia*, vol. 111, no. 5, pp. 711–720, Nov. 2013.
- [4] L. Manchikanti, A. Kaye, K. Manchikanti, M. Boswell, V. Pampati, and J. Hirsch, "Efficacy of epidural injections in the treatment of lumbar central spinal stenosis: A systematic review," *Anesthesiol. Pain Med.*, vol. 5, no. 1, p. e23139, 2015.
- [5] R. Russell, M. Popat, E. Richards, and J. Burry, "Combined spinal epidural anaesthesia for caesarean section: A randomised comparison of Oxford, lateral and sitting positions," *Int. J. Obstetric Anesthesia*, vol. 11, no. 3, pp. 190–195, Jul. 2002.
- [6] R. D. Vincent and D. H. Chestnut, "Which position is more comfortable for the parturient during identification of the epidural space?" *Int. J. Obstetric Anesthesia*, vol. 1, no. 1, pp. 9–11, Sep. 1991.
- [7] H. Q. De Tran, A. P. González, F. Bernucci, and R. J. Finlayson, "Confirmation of loss-of-resistance for epidural analgesia," *Regional Anesthesia Pain Med.*, vol. 40, no. 2, pp. 166–173, Mar./Apr. 2015.
- [8] A. D. Franklin and E. M. Hughes, "Fluoroscopically guided tunneled trans-caudal epidural catheter technique for opioid-free neonatal epidural analgesia," *J. Anesthesia*, vol. 30, no. 3, pp. 493–497, Jun. 2016.
- [9] G. Le Coq, B. Ducot, and D. Benhamou, "Risk factors of inadequate pain relief during epidural analgesia for labour and delivery," *Can. J. Anaesthesia*, vol. 45, no. 8, pp. 719–723, Aug. 1998.
- [10] P. H. Conroy, C. Luyet, C. J. McCartney, and P. G. McHardy, "Real-time ultrasound-guided spinal anaesthesia: A prospective observational study of a new approach," *Anesthesiol. Res. Pract.*, vol. 2013, 2013, Art. no. 525818.
- [11] M. K. Karmakar, X. Li, A. M.-H. Ho, W. H. Kwok, and P. T. Chui, "Real-time ultrasound-guided paramedian epidural access: Evaluation of a novel in-plane technique," *Brit. J. Anaesthesia*, vol. 102, no. 6, pp. 845–854, Jun. 2009.
- [12] A. Lee, "Ultrasound in obstetric anesthesia," *Seminars Perinatol.*, vol. 38, no. 6, pp. 349–358, 2014.
- [13] T.-L. Lin, C.-T. Chung, H. H.-C. Lan, and H.-M. Sheen, "Ultrasound-guided facet joint injection to treat a spinal cyst," *J. Chin. Med. Assoc.*, vol. 77, no. 4, pp. 213–216, Apr. 2014.
- [14] A. E. Q. Santiago *et al.*, "Ultrasound-guided facet block to low back pain: A case report," *Brazilian J. Anesthesiol.*, vol. 64, no. 4, pp. 278–280, Jul./Aug. 2014.
- [15] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Trans. Sonics Ultrason.*, vol. SU-30, no. 3, pp. 156–163, May 1983.
- [16] M. A. Lediju, M. J. Pihl, J. J. Dahl, and G. E. Trahey, "Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter," *Ultrason. Imag.*, vol. 30, no. 3, pp. 151–168, 2008.
- [17] L. T. Kim, "Principles of ultrasound and applied ultrasound physics relevant for advanced sonographers," in *Advanced Thyroid and Parathyroid Ultrasound*. Cham, Switzerland: Springer, 2017, pp. 37–47.
- [18] A. U. Niazi, N. Haldipur, A. G. Prasad, and V. W. Chan, "Ultrasound-guided regional anesthesia performance in the early learning period: Effect of simulation training," *Regional Anesthesia Pain Med.*, vol. 37, no. 1, pp. 51–54, Jan./Feb. 2012.
- [19] B. D. Sites, B. C. Spence, J. D. Gallagher, C. W. Wiley, M. L. Bertrand, and G. T. Blike, "Characterizing novice behavior associated with learning ultrasound-guided peripheral regional anesthesia," *Regional Anesthesia Pain Med.*, vol. 32, no. 2, pp. 107–115, Mar. 2007.
- [20] T. Grau, E. Bartschke, R. Conradi, E. Martin, and J. Motsch, "Ultrasound imaging improves learning curves in obstetric epidural anesthesia: A preliminary study," *Can. J. Anesthesia*, vol. 50, no. 10, pp. 1047–1050, Dec. 2003.
- [21] T. Wu, W.-H. Zhao, Y. Dong, H.-X. Song, and J.-H. Li, "Effectiveness of ultrasound-guided versus fluoroscopy or computed tomography scanning guidance in lumbar facet joint injections in adults with facet joint syndrome: A meta-analysis of controlled trials," *Arch. Phys. Med. Rehabil.*, vol. 97, no. 9, pp. 1558–1563, Sep. 2016.
- [22] G. Yang *et al.*, "Ultrasound-guided versus fluoroscopy-controlled lumbar transforaminal epidural injections: A prospective randomized clinical trial," *Clin. J. Pain*, vol. 32, no. 2, pp. 103–108, Feb. 2015.
- [23] I. Evansa, I. Logina, I. Vanags, and A. Borgeat, "Ultrasound versus fluoroscopic-guided epidural steroid injections in patients with degenerative spinal diseases: A randomised study," *Eur. J. Anaesthesiol.*, vol. 32, no. 4, pp. 262–268, Apr. 2015.
- [24] J. Obernauer *et al.*, "Ultrasound-guided versus computed tomography-controlled facet joint injections in the middle and lower cervical spine: A prospective randomized clinical trial," *Med. Ultrasonogr. J.*, vol. 15, no. 1, pp. 5–10, 2013.
- [25] A. G. Voloshin, "Four-dimensional ultrasound guidance during epidural anaesthesia," *J. Ultrasound*, vol. 18, no. 2, pp. 135–142, Jun. 2015.
- [26] S. H. Yoon, S. L. O'Brien, and M. Tran, "Ultrasound guided spine injections: Advancement over fluoroscopic guidance?" *Current Phys. Med. Rehabil. Rep.*, vol. 1, no. 2, pp. 104–113, Jun. 2013.
- [27] M. Pesteie *et al.*, "Real-time ultrasound image classification for spine anesthesia using local directional Hadamard features," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 6, pp. 901–912, Jun. 2015.
- [28] T. Grau, R. W. Leipold, J. Horter, R. Conradi, E. O. Martin, and J. Motsch, "Paramedian access to the epidural space: The optimum window for ultrasound imaging," *J. Clin. Anesthesia*, vol. 13, no. 3, pp. 213–217, May 2001.
- [29] K. K. Srinivasan, G. Iohom, F. Loughnane, and P. J. Lee, "Conventional landmark-guided midline versus preprocedure ultrasound-guided paramedian techniques in spinal anesthesia," *Anesthesia Analgesia*, vol. 121, no. 4, pp. 1089–1096, Oct. 2015.
- [30] C. Darrieurtort-Laffite, G. Bart, L. Planche, J. Glemarec, Y. Maugars, and B. Le Goff, "Usefulness of a pre-procedure ultrasound scanning of the lumbar spine before epidural injection in patients with a presumed difficult puncture: A randomized controlled trial," *Joint Bone Spine*, vol. 82, no. 5, pp. 356–361, Oct. 2015.
- [31] A. Perlas, L. E. Chaparro, and K. J. Chin, "Lumbar neuraxial ultrasound for spinal and epidural anesthesia: A systematic review and meta-analysis," *Regional Anesthesia Pain Med.*, vol. 41, no. 2, pp. 251–260, 2016.
- [32] L. Luo, J. Ni, L. Wu, and D. Luo, "Ultrasound-guided epidural anesthesia for a parturient with severe malformations of the skeletal system undergoing cesarean delivery: A case report," *Local Regional Anesthesia*, vol. 8, pp. 7–10, May 2015.
- [33] K. D. Park, T. K. Kim, W. Y. Lee, J. Ahn, S. H. Koh, and Y. Park, "Ultrasound-guided versus fluoroscopy-guided caudal epidural steroid injection for the treatment of unilateral lower lumbar radicular pain: Case-controlled, retrospective, comparative study," *Medicine*, vol. 94, no. 50, p. e2261, 2015.
- [34] A. U. Niazi, K. J. Chin, R. Jin, and V. W. Chan, "Real-time ultrasound-guided spinal anesthesia using the sonixgps ultrasound guidance system: A feasibility study," *Acta Anaesthesiol. Scand.*, vol. 58, no. 7, pp. 875–881, Aug. 2014.
- [35] J. E. Johnson, A. Perlas, and K. J. Chin, "Ultrasound guidance for central neuraxial anesthesia and analgesia," *Adv. Anesthesia*, vol. 32, no. 1, pp. 149–169, 2014.
- [36] T. Ungi *et al.*, "Spinal needle navigation by tracked ultrasound snapshots," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2766–2772, Oct. 2012.
- [37] E. C. Chen, P. Mousavi, S. Gill, G. Fichtinger, and P. Abolmaesumi, "Ultrasound guided spine needle insertion," *Proc. SPIE*, vol. 7625, p. 762538, Aug. 2010.
- [38] H. Rafii-Tari, V. A. Lessoway, A. A. Kamani, P. Abolmaesumi, and R. Rohling, "Panorama ultrasound for navigation and guidance of epidural anesthesia," *Ultrasound Med. Biol.*, vol. 41, no. 8, pp. 2220–2231, Aug. 2015.
- [39] J. Boisvert, F. Cherié, X. Pennec, H. Labelle, and N. Ayache, "Articulated spine models for 3-D reconstruction from partial radiographic data," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 11, pp. 2565–2574, Nov. 2008.
- [40] S. Khallaghi *et al.*, "Registration of a statistical shape model of the lumbar spine to 3D ultrasound images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 6362, Cham, Switzerland: Springer, 2010, pp. 68–75.

- [41] A. Rasoulia, R. N. Rohling, and P. Abolmaesumi, "Augmentation of paramedian 3D ultrasound images of the spine," in *Information Processing in Computer-Assisted Interventions. IPCAI* (Lecture Notes in Computer Science), vol. 7915. Berlin, Germany: Springer, 2013, pp. 51–60.
- [42] M. G. Roberts, T. F. Coates, E. Pacheco, T. Oh, and J. E. Adams, "Segmentation of lumbar vertebrae using part-based graphs and active appearance models," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 5762. Berlin, Germany: Springer, 2009, pp. 1017–1024.
- [43] A. Rasoulia *et al.*, "Ultrasound-guided spinal injections: A feasibility study of a guidance system," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 9, pp. 1417–1425, Sep. 2015.
- [44] M. Brudfors *et al.*, "Towards real-time, tracker-less 3D ultrasound guidance for spine anaesthesia," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 6, pp. 855–865, Jun. 2015.
- [45] S. Yu, K. K. Tan, B. L. Sng, S. Li, and A. T. H. Sia, "Lumbar ultrasound image feature extraction and classification with support vector machine," *Ultrasound Med. Biol.*, vol. 41, no. 10, pp. 2677–2689, Oct. 2015.
- [46] S. Yu, K. K. Tan, B. L. Sng, S. Li, and A. T. H. Sia, "Real-time automatic spinal level identification with ultrasound image processing," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 243–246.
- [47] D. Tran and R. N. Rohling, "Automatic detection of lumbar anatomy in ultrasound images of human subjects," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 9, pp. 2248–2256, Sep. 2010.
- [48] D. Belavy, M. J. Ruitenber, and R. B. Brijball, "Feasibility study of real-time three-/four-dimensional ultrasound for epidural catheter insertion," *Brit. J. Anaesthesia*, vol. 107, no. 3, pp. 438–445, Sep. 2011.
- [49] P. Malenfant, V. Gunka, P. Beigi, A. Rasoulia, R. Rohling, and A. Dube, "Accuracy of 3D ultrasound for identification of epidural needle skin insertion point in parturients; A prospective observational study," in *Proc. 46th Annu. Meet. Soc. Obstetric Anesthesia Perinatol. (SOAP)*, Toronto, ON, Canada, 2014, p. 308.
- [50] H. Wang *et al.*, "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," *J. Med. Imag.*, vol. 1, no. 3, p. 034003, 2014.
- [51] K. G. Beauchamp, *Applications of Walsh and Related Functions*. San Diego, CA, USA: Academic, 1984.
- [52] D. L. Donoho and X. Huo, "Beamlet pyramids: A new form of multiresolution analysis, suited for extracting lines, curves, and objects from very noisy image data," *Proc. SPIE*, vol. 4119, pp. 434–444, Jul. 2000.
- [53] D. Tran, A. A. Kamani, V. A. Lessoway, C. Peterson, K. W. Hor, and R. N. Rohling, "Preinsertion paramedian ultrasound guidance for epidural anesthesia," *Anesthesia Analgesia*, vol. 109, no. 2, pp. 661–667, Aug. 2009. [Online]. Available: <https://doi.org/10.1213/ane.0b013e3181a94c75>
- [54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [55] T. N. Sainath *et al.*, "Deep convolutional neural networks for large-scale speech tasks," *Neural Netw.*, vol. 64, pp. 39–48, Apr. 2015.
- [56] J. Kleesiek *et al.*, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016.
- [57] J. M. Wolterink, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, and I. Išgum, "Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks," *Med. Image Anal.*, vol. 34, pp. 123–136, Dec. 2016.
- [58] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034. [Online]. Available: <https://doi.org/10.1109/iccv.2015.123>
- [59] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [60] E. M. A. Anas *et al.*, "Bone enhancement in ultrasound using local spectrum variations for guiding percutaneous scaphoid fracture fixation procedures," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 6, pp. 959–969, Jun. 2015.
- [61] T. J. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons," *Kongelige Danske Videnskaberne Selskab; Biologiske Skrifter*, vol. 5, no. 4, pp. 1–34, 1948.
- [62] R. Al-Rfou *et al.* (May 2016). "Theano: A Python framework for fast computation of mathematical expressions." [Online]. Available: <https://arxiv.org/abs/1605.02688>
- [63] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.
- [64] R. W. D. Nickalls and M. S. Kokri, "The width of the posterior epidural space in obstetric patients," *Anaesthesia*, vol. 41, no. 4, pp. 432–433, Apr. 1986.
- [65] V. L. Hoffmann, M. P. Vercauteren, J. P. Vreugde, G. H. Hans, H. C. Coppejans, and H. A. Adriaensen, "Posterior epidural space depth: Safety of the loss of resistance and hanging drop techniques," *Brit. J. Anaesthesia*, vol. 83, no. 5, pp. 807–809, Nov. 1999.
- [66] D. Tran, A. A. Kamani, E. Al-Attas, V. A. Lessoway, S. Massey, and R. N. Rohling, "Single-operator real-time ultrasound-guidance to aim and insert a lumbar epidural needle," *Can. J. Anesthesia*, vol. 57, no. 4, pp. 313–321, Apr. 2010.