

Segmentation of Lumbar Spine MRI Images for Stenosis Detection using Patch-based Pixel Classification Neural Network

Ala S. Al Kafri¹, Sud Sudirman¹, Abir J. Hussain¹, Dhiya Al-Jumeily¹, Paul Fergus¹, Friska Natalia², Hira Meidia², Nunik Afriliana², Ali Sophian³, Mohammed Al-Jumaily⁴, Wasfi Al-Rashdan⁵ and Mohammad Bashtawi⁵

¹Liverpool John Moores University, Liverpool L3 3AF, UK
a.s.alkafri@2015.ljmu.ac.uk; {s.sudirman; a.hussain; d.aljumeily; p.fergus}@ljmu.ac.uk

²Universitas Multimedia Nusantara, Jakarta, Indonesia
{friska.natalia; hira.meidia; nunik}@umn.ac.id

³International Islamic University Malaysia, Kuala Lumpur, Malaysia
ali_sophian@iium.edu.my

⁴Dr. Sulaiman Al Habib Hospital, Dubai Healthcare City, Dubai, UAE
maljumaily@yahoo.fr

⁵Irbid Speciality Hospital, Irbid, Jordan
drwasfi1@hotmail.com; mohdbakhiet@yahoo.com

Abstract— This paper addresses the central problem of automatic segmentation of lumbar spine Magnetic Resonance Imaging (MRI) images to delineate boundaries between the anterior arch and posterior arch of the lumbar spine. This is necessary to efficiently detect the occurrence of lumbar spinal stenosis as a leading cause of Chronic Lower Back Pain. A patch-based classification neural network consisting of convolutional and fully connected layers is used to classify and label pixels in MRI images. The classifier is trained using overlapping patches of size 25x25 pixels taken from a set of cropped axial-view T2-weighted MRI images of the bottom three intervertebral discs. A set of experiment is conducted to measure the performance of the classification network in segmenting the images when either all or each of the discs separately is used. Using pixel accuracy, mean accuracy, mean Intersection over Union (IoU), and frequency weighted IoU as the performance metrics we have shown that our approach produces better segmentation results than eleven other pixel classifiers. Furthermore, our experiment result also indicates that our approach produces more accurate delineation of all important boundaries and making it best suited for the subsequent stage of lumbar spinal stenosis detection.

Keywords—Patch-based classification neural network, lumbar spine MRI, lumbar spinal stenosis

I. INTRODUCTION

Lumbar spinal stenosis is a narrowing of spinal canal caused by inflammation of bone or soft tissues, which in turn produces pressure on spinal nerve roots. This pressure could result in Chronic Lower Back Pain (CLBP) with symptoms ranging from radicular pain to atypical leg pain to neurogenic claudication [1].

CLBP is a debilitating illness that is affecting the health, social life, and employment of millions of people around the world. In the UK, the cost of treating patients with CLBP is estimated to be around £500 million annually to the National Health Service (NHS) [2]. This is on top of other economic cost

resulting from the loss of productivity and other informal care – which is estimated to reach around £10,668 million [3].

It is important to emphasise that CLBP is a chronic illness. While millions of people experience some sort of back pain at one point in their lives, only small percentage of them experience what is termed acute lower back pain. This form of back pain is often non-specific and may not necessarily caused by a serious condition. However, a small percentage of people with acute lower back pain can develop CLBP if the underlying cause is serious and left untreated. A study [4] discovered that the longer a person is disabled by back pain, the less chance he or she returns to work and the more health care cost he or she will require. It is therefore important for early identification of back pain cause to happen in order to improve the chance of patient rehabilitation.

The speediness of early diagnosis can depend on many factors including referral time from GP to the hospital, waiting time for a specialist appointment, time for an MRI scan and time for the analysis result to come out. Due to heavy demand for radiologist and clinicians, the whole process could take weeks to complete.

As a short case study, NHS England currently put a target for all imaging and radio diagnostics to be within 6 to 13 weeks. However, specialist shortages have caused many patients to wait longer. A report by NHS England [5] reveals that there is a significant number of cases where diagnostic wait time can be extended to more than 13 weeks. This problem is expected to get worse since the number of scans is historically an uptrend. A report [6] by National Office of Statistics shows that there is a 12.3% average annual growth in demand for MRI scans since 1995. This rationalises the need for a new approach to increase the efficiency and effectiveness of the imaging diagnostic process.

One way to achieve this is to utilise current advances in image analysis and machine-learning technologies to develop a computer assisted diagnostic software. Clinical studies have shown that there are strong correlations between clinical outcome and the morphological characteristics of patient’s lumbar spine as captured in the MRI images [7]. Using a combination of different image analysis and machine learning techniques, computers can analyse the images and identify regions of interests for clinicians to focus on. The techniques could also be used to provide tools for the specialists to help them come to the correct diagnosis quickly.

As with many other medical imaging processes, the aforementioned approach will apply image segmentation as one of the earliest yet important steps. To ensure the validity of this step, the process must be grounded in sound medical evidence and reasoning as well as proven methodology. This paper presents the result of segmentation of specific area of lumbar spine MRI that is important to the diagnosis of lumbar spinal stenosis using patch-based Deep Neural Network.

The paper will start by describing the process of diagnosing lumbar spine stenosis, followed by the design of the network architecture, description of the data and the algorithm used, followed by experiment results and analysis.

II. DIAGNOSTIC PROCESS OF LUMBAR SPINAL STENOSIS

A. Magnetic Resonance Imaging (MRI)

MRI is the preferred method of medical scans for detecting the causes of back pains. MRI images can be used to visualise lumbar spine, slice by slice, in three view-planes namely sagittal (side), axial (top-down) and coronal (frontal) – typically only the first two are used in lumbar spine MRI.

Sagittal view of lumbar spine MRI is the easiest to understand and traditionally is considered the best view that can be used to detect certain types of pain including conditions. This can be seen from the approach used by some of the earliest computer assisted diagnosis of disc herniation [8]–[10]. On the other hand, axial view of MRI is much more difficult to read because it has much more information about the lumbar spine at each vertebrae segment. As a result, in this view, the clinicians could pinpoint the exact location of any tissue inflammations, facet cysts, disc herniation, stenosis, and fractures.

Based on the timing of radiofrequency pulse sequences used, MRI images can be of any one of two types namely T1- and T2-weighted MRI. The same tissue could have markedly different intensity levels when imaged in each type. One example is *cerebrospinal fluid* (CSF), a clear and colourless body fluid that can be found inside our brain and spinal cord. Identification of CSF tissue in MRI images has significant relevance to our case since it surrounds the many nerve roots inside the spinal cord. For any further and more detailed information on MRI and its uses as medical imaging technology, interested readers can refer to [11] or any other relevant textbooks in this area.

B. Anatomy of Lumbar Spine

The lumbar spine is made up of five lumbar vertebrae and each vertebra consists of an anterior (body) arch and posterior

arch. In a mid-sagittal view of the lumbar spine shown in Fig. 1, the two arches are separated by a long white opening. The part of this opening that is visible in this mid-sagittal cut is *thecal sac* (TS) which contains CSF, the same type of fluid that resides inside the brain. The back of the opening, which borders with the anterior of the posterior arch, is covered with *ligamentum flavum* (LF).

The anterior arch of each lumbar vertebra is labelled L1 – L5, each separated by an Intervertebral Disc (IVD) labelled D1 to D4. The last disc, D5, separates L5 and the large triangular shaped bone at the bottom of the spine, called the *sacrum*.

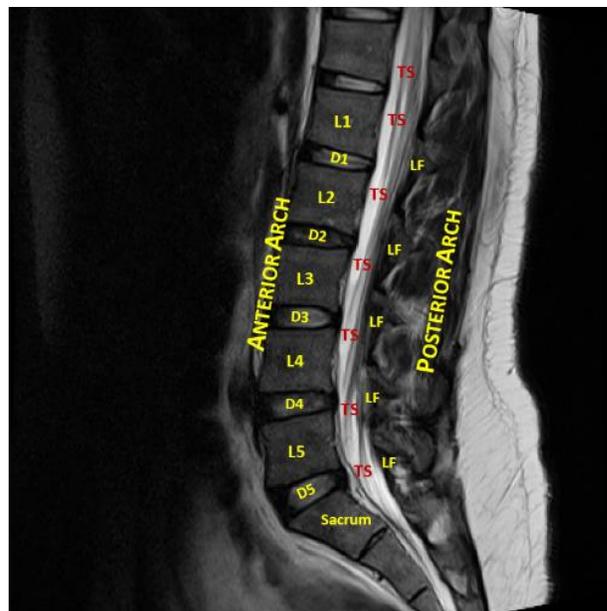


Fig. 1. T2-weighted sagittal view MRI of a lumbar spine

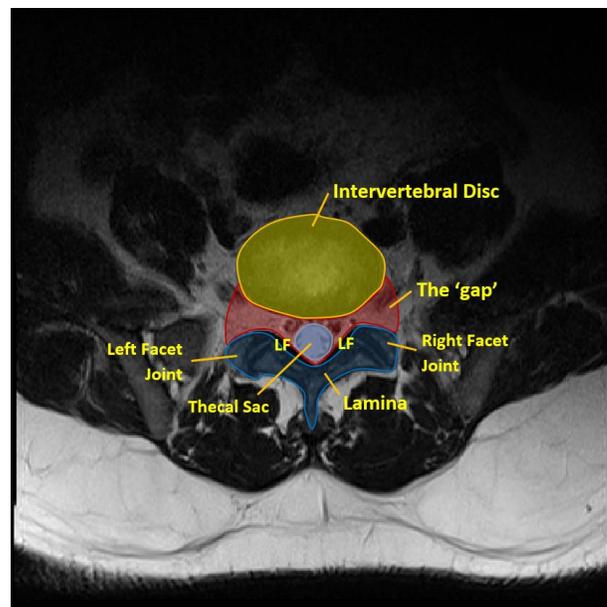


Fig. 2. T2-weighted axial view MRI of D4

The axial view of the spine manifests as different slices of the MRI images across the vertebra or intervertebral discs. Such view, as illustrated in Fig. 2, can show more information on the various tissues surrounding each IVD and a part of the posterior arch called *Lamina*. In this view, the opening (the gap) between the IVD and the lamina can be seen very clearly as a white region that contains the TS. As will be discussed later, identifying the area of this opening and its borders with both the disc and the lamina, plays an important role in detecting the occurrence of stenosis in the lumbar spine.

C. Lumbar Spinal Stenosis Detection

The gap between the anterior arch and the posterior arch of the vertebrae extends from the cervical spine down to the lumbar spine. For the lack of a better word, we refer this gap in this paper as the *gap*. The gap's width varies depending on the location in the spine where the measurement is taken. Furthermore, there have been a number of studies [12]–[14], which record the measurements of spinal canal widths in a population of different countries and the results suggest that the gap's width is also affected by the patient's ethnicity.

A lumbar spine stenosis can manifest as the narrowing of any parts of the gap and depending on where it occurs can be classified as either central stenosis or lateral stenosis. This phenomenon is associated as one of the main cause of CLBP because of both TS and NF house many nerve roots from different parts of the spine to other parts of the body. An abnormal compression of either of them would exert pressure on these roots and creates a sensation of pain.

The stenosis could occur in any part of the gap and could be caused by different types of defect such as posterior/posterolateral disc herniation, osteoarthritic thickening of the posterolateral vertebral body, or hypertrophy of *ligamentum flavum* (LF). In all of these cases, clinicians will perform manual delineation of the boundaries between the gap and the IVD, between the gap and the left and right facet joints, and between the gap and LF. These three boundaries are illustrated in Fig. 3.

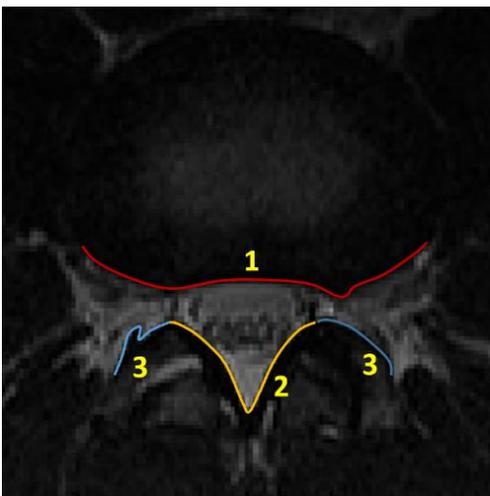


Fig. 3. The three important boundaries for stenosis detection between the gap and 1) Intervertebral Disc, 2) Ligamentum Flavum and 3) Facet Joints

A hastily done observation of the above case would suggest that the boundaries could be located by applying an edge detection algorithm. However, it should be noted that lumbar spine MRI images can have a varying wide spectrum of edge strengths with the strength of those three edges is somewhat convoluted in the middle. This makes detection of those three edges alone a challenge. To illustrate this claim, two results of applying the Canny edge detection using two sets of thresholds and sigma values to capture different levels of edge strength are shown in Fig. 4. The figure shows that the three edges cannot be reliably located without losing accuracy or having too much noise.

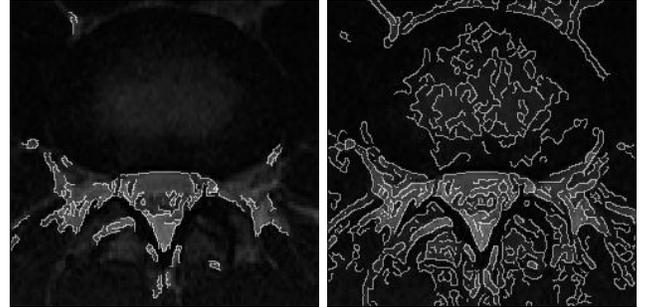


Fig. 4. The results of Canny edge detection. The left image is produced to capture strong edges only using higher threshold and sigma values than the one on the right.

This then leads to our conclusion that the three boundaries can be located by first performing segmentation on the MRI image to separate the gap region from the rest of the lumbar spine areas before applying the edge detection algorithm. It is also important to note that this approach will consider both TS and NF as one homogenous part of the gap hence T2-weighted MRI should be used.

III. PATCH-BASED IMAGE SEGMENTATION

Image segmentation is one of the fundamental processes in image analysis. Traditional approaches to image segmentation include clustering techniques based on pixel values such as k-means clustering, histogram-based clustering, and probabilistic-based model. However, current state-of-the-art image segmentation techniques are dominated by different types of Deep Neural Network to achieve semantic segmentation. There are two main approaches of semantic segmentation namely patch-based pixel classification [15]–[17] using convolutional and fully connected layers, and whole-image segmentation [18]–[20] using fully convolutional and convolution-transpose layers.

In this paper, we utilise a patch-based classification approach to image segmentation to separate the gap area from other tissues in lumbar spine MRI. For each pixel, p , in the input image, a patch centred at p is considered. The patch dimension, w , is set to an odd number to enforce symmetry. The network will classify each pixel based on the information contained in the patch into one of two possible classes namely NF (0) and non-NF (1). The process flow of image segmentation using this architecture is illustrated in Fig. 5.

A. Deep Neural Network Architecture

Our neural network shares a similar architecture to other patch-based deep neural networks. It includes an Image Input Layer, a cascade of Convolution Layers that will produce multiscale classification features and a Fully Connected Layer for classification of these features.

The Image Input Layer contains a 2D array of nodes that act as a receptor to each pixel in the patch. The dimension of this layer is identical to the dimension of the patch. A data normalisation step is applied to the pixel values by subtracting the mean pixel value before generating the output of this layer.

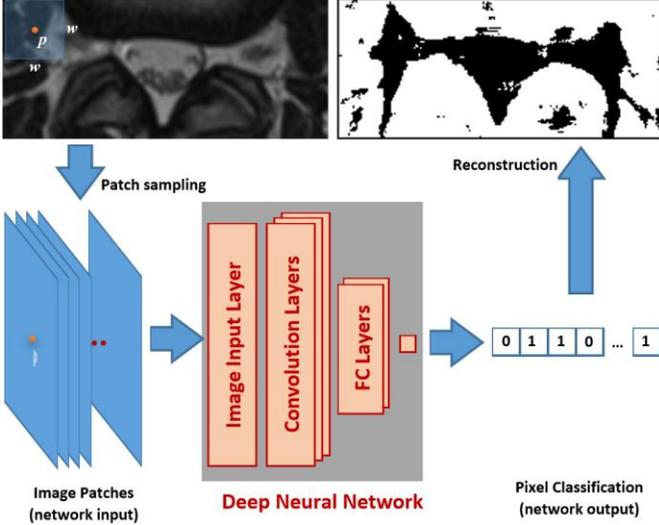


Fig. 5. The process flow of image segmentation using patch-based pixel classification approach.

The output of the Image Input Layer is fed to a cascade of Convolution Layers. In our network, each convolution layer has a fixed number of equal-sized kernels. The kernel size chosen is relatively small since the input of the network is already a small subset of the entire image. Each kernel corresponds to each classification feature trained within the layer. Each of the output of these layers is calculated as the dot product of the input and the kernel. As the signal is passed across these layers, it gets down sampled hence the output of each layer covers an increasing support area of the input patch. Each layer, however, is set to have an increased number of kernels compared to the last. This decision is made to allow more features to be trained as the support area of the convolution layer increases.

To further speed up the training process and reduce the sensitivity of the network to its initial weight values, we opt to apply batch normalisation [21] [22] to the convolution layer output. This approach works by subtracting the output of the previous convolution layer by the batch mean and dividing the result by the batch standard deviation. The increase in training speed and stability are achieved since the batch normalisation process reduces the variability of the input signal into two trainable parameters namely, the batch means and batch standard deviation.

As with other deep neural network architectures, we opt to use Rectified Linear Unit (ReLU) activation function [23]. This decision is based on the function’s advantages over others, such as Sigmoid function, that include faster training speed due to a reduced likelihood of vanishing gradient. ReLU activation function also has been proven to train networks with sparser, and hence considered better, weights representation.

We set the convolution stride to one and apply an appropriate number of pixel padding to its input to ensure the length of input and output of each convolution layer remain the same. A down sampling process is applied at the end of each convolution layer by taking the maximum of the activation signals inside a 2x2 non-overlapping region to produce a signal that has half the dimension. In total, our network uses three convolution layers to perform the feature weight training.

After the end of the last convolution layer, we use two Fully Connected Layers to learn all high-level combinations of the features learned by the earlier layers. Since the network will classify the central pixel in each patch into one of the two possible classes, the last layer will have two neurons. For classification, our network uses the soft max classifier. This classifier produces the probability of the input to belong to each class and can produce a more representative loss value when evaluating the network [24].

For ease of reference, throughout the remainder of this paper, we will refer to this network as PALMSNet which is an acronym for Patch-based Axial-view Lumbar-spine MRI Segmentation Neural Network. The overall architecture of PALMSNet is summarised in Table I.

TABLE I. THE ARCHITECTURE OF PALMSNET

Layer	Type	Number of Neurons	Kernel Size	Number of Kernels
0	Image Input	25x25		
1	Convolution	25x25	3x3	16
2	Batch Normalisation	25x25		
3	ReLU	25x25		
4	Max Pooling	25x25		
5	Convolution	12x12	3x3	32
6	Batch Normalisation	12x12		
7	ReLU	12x12		
8	Max Pooling	12x12		
9	Convolution	6x6	3x3	64
10	Batch Normalisation	6x6		
11	ReLU	6x6		
12	Max Pooling	6x6		
13	Fully Connected	2304		
14	Fully Connected	2		
15	Soft Max	1		

B. Training

The training data consists of 400,680 image patches of size 25x25 pixels. They are taken from T2-weighted axial lumbar spine MRI images of the last three IVDs (D3-D5) of seven patients. The images are cropped to provide better focus on the areas around the IVD, the gap, and the lamina. Ground truth images were created with the assistance of physicians in this field by creating a binary image marking the area of interest. An

example of the MRI and ground-truth (label) image pair is shown in Fig. 6.

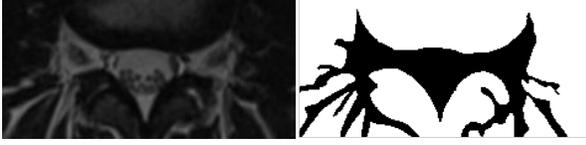


Fig. 6. An example of MRI and ground-truth image pair used for training.

Furthermore, we exploit the fact that all axial view of lumbar spine MRI images are always taken from the same orientation and similar scale to avoid having to augment the image patches synthetically. However, this approach may not be suitable when using the network to segment other types of medical images.

The network is trained using the popular Stochastic Gradient Descent with Momentum algorithm [25] to update the network weights and biases. The algorithm works by taking small steps in the direction of the negative gradient of the loss function to minimize the error function. The size of the step is modulated by a learning rate parameter that is set to 0.001.

The training is performed on a personal computer in Windows 10 with i7-7700 CPU @ 3.60GHz, 64 GB RAM, and two NVIDIA Titan X GPUs. The training process takes 2,102 seconds and 9.5×10^4 iterations to complete.

IV. EXPERIMENT RESULTS

Once trained, the network is used to classify pixels from test images. The test data consists of 171,720 image patches of size 25x25 pixels extracted from T2-weighted axial lumbar spine MRI images of D3-D5 of three patients. They are different patients from the seven whose MRI images are used for training.

We carried out four different experiments to measure the performance of PALMSNet. The first experiment uses the images from all three IVDs for training and testing whereas the other three use only those from each disc separately. This experiment setup is used based on our hypothesis made during observation and analysis of the MRI images, that separation of the disc images during training will improve the classification results.

We will use four performance metrics that are common in other semantic segmentation evaluations [19]. They are pixel accuracy (a_p), mean accuracy (a_m), mean intersection over union (iou_m) and frequency weighted intersection over union (iou_{fw}). The formula to calculate these metrics are shown in Eq. (1) to (4).

$$a_p = \frac{\sum_i n_{ii}}{\sum_i t_i} \quad (1)$$

$$a_m = \frac{\sum_i \frac{n_{ii}}{t_i}}{n_{cl}} \quad (2)$$

$$iou_m = \frac{\sum_i \frac{n_{ii}}{t_i + \sum_i n_{ii} - n_{ii}}}{n_{cl}} \quad (3)$$

$$iou_{fw} = \frac{\sum_i \frac{t_i n_{ii}}{t_i + \sum_i n_{ii} - n_{ii}}}{\sum_i t_i} \quad (4)$$

Where n_{ij} is the number of pixels of class i predicted to belong to class j , and n_{cl} is the number of classes and $t_i = \sum_j n_{ij}$ is the total number of pixels of class i .

For comparison purposes, we perform another image segmentation using SegNet [20], a fully convolutional neural net, and using ten other, but more conventional, image segmentation algorithms. Four of these ten algorithms are variants of the decision tree (complex/medium [26]–[28], and ensemble [29], [30]). A further four are based on the modified k-nearest neighbour (medium, weighted, coarse and cubic [31]–[34]) and the other two are based on modified SVM (fine and medium Gaussian [35]). The features used in these ten conventional algorithms are the raw pixel values and their corresponding locations. Pixel locations have been used in conjunction with pixel values to provide better spatial coherence to the segmentation result [36].

TABLE II. PERFORMANCE USING ALL INTERVERTEBRAL DISCS

Methods	a_p	a_m	iou_m	iou_{fw}
Fine Gaussian SVM	0.78	0.77	0.62	0.64
Medium Gaussian SVM	0.79	0.78	0.64	0.66
Cubic k-nearest neighbour	0.79	0.78	0.64	0.66
Medium k-nearest neighbour	0.79	0.78	0.64	0.66
Weighted k-nearest neighbour	0.77	0.77	0.62	0.64
Coarse k-nearest neighbour	0.79	0.79	0.64	0.66
Complex Tree	0.75	0.75	0.59	0.61
Medium Tree	0.75	0.75	0.59	0.61
Boosted Tree Ensemble	0.76	0.76	0.60	0.62
Bagged Tree Ensemble	0.75	0.75	0.59	0.61
SegNet	0.75	0.67	0.53	0.58
PALMSNet	0.86	0.85	0.75	0.76

TABLE III. PERFORMANCE WHEN TRAINED AND TESTED USING D3 ONLY

Methods	a_p	a_m	iou_m	iou_{fw}
Fine Gaussian SVM	0.81	0.77	0.62	0.70
Medium Gaussian SVM	0.79	0.78	0.60	0.68
Cubic k-nearest neighbour	0.79	0.78	0.60	0.68
Medium k-nearest neighbour	0.79	0.78	0.60	0.68
Weighted k-nearest neighbour	0.79	0.78	0.61	0.69
Coarse k-nearest neighbour	0.80	0.78	0.61	0.69
Complex Tree	0.74	0.74	0.54	0.62
Medium Tree	0.76	0.75	0.56	0.64
Boosted Tree Ensemble	0.71	0.74	0.52	0.59
Bagged Tree Ensemble	0.79	0.78	0.60	0.68
SegNet	0.79	0.73	0.58	0.68
PALMSNet	0.93	0.92	0.85	0.87

The results of the segmentation as measured using the four performance metrics are shown in Table II-V. The results shown in the Table II-IV indicate that PALMSNet performs significantly better than the other eleven. The other DNN-based algorithm, however, has a similar level of pixel accuracy than the more conventional methods but performs slightly worse

using the other three metrics. It is important to note from these tables that the training and testing on D3 or D4 separately yield significantly better results than on all of them together.

TABLE IV. PERFORMANCE WHEN TRAINED AND TESTED USING D4 ONLY

Methods	a_p	a_m	iou_m	iou_{fw}
Fine Gaussian SVM	0.82	0.78	0.63	0.71
Medium Gaussian SVM	0.81	0.78	0.62	0.70
Cubic k-nearest neighbour	0.80	0.78	0.61	0.68
Medium k-nearest neighbour	0.80	0.78	0.61	0.69
Weighted k-nearest neighbour	0.80	0.78	0.61	0.69
Coarse k-nearest neighbour	0.81	0.78	0.63	0.70
Complex Tree	0.76	0.76	0.57	0.65
Medium Tree	0.77	0.75	0.57	0.65
Boosted Tree Ensemble	0.79	0.77	0.59	0.67
Bagged Tree Ensemble	0.78	0.77	0.59	0.67
SegNet	0.73	0.69	0.53	0.57
PALMSNet	0.92	0.91	0.84	0.85

The last experiment results are shown in Table V also indicate the superiority of PALMSNet over the other eleven methods. However, the gaps between them are less pronounced than the previous three results. For example, only 0.02% difference in mean accuracy between the best of the eleven and PALMSNet in Table V compared to 0.06% in Table I.

TABLE V. PERFORMANCE WHEN TRAINED AND TESTED USING D5 ONLY

Methods	a_p	a_m	iou_m	iou_{fw}
Fine Gaussian SVM	0.68	0.71	0.48	0.55
Medium Gaussian SVM	0.68	0.71	0.49	0.56
Cubic k-nearest neighbour	0.66	0.70	0.47	0.53
Medium k-nearest neighbour	0.66	0.70	0.47	0.53
Weighted k-nearest neighbour	0.67	0.69	0.47	0.54
Coarse k-nearest neighbour	0.69	0.72	0.49	0.56
Complex Tree	0.66	0.68	0.46	0.53
Medium Tree	0.65	0.65	0.44	0.52
Boosted Tree Ensemble	0.64	0.68	0.45	0.51
Bagged Tree Ensemble	0.65	0.68	0.46	0.52
SegNet	0.60	0.62	0.43	0.43
PALMSNet	0.74	0.74	0.59	0.59

The results provide a mixed degree of validation to our initial hypothesis. While the decision to separate the disc yields much better results for D3 and D4, the same cannot be said for D5. As can be seen in Table VI the net trained using data from all discs still yields better performance on D5 test data than the network that is trained using only D5 data.

TABLE VI. PALMSNET PERFORMANCE WHEN TESTED ON D5

Methods	a_p	a_m	iou_m	iou_{fw}
PALMSNet trained using all discs	0.80	0.80	0.66	0.66
PALMSNet trained using D5 only	0.74	0.74	0.59	0.59

V. DISCUSSION AND ANALYSIS

The experiment results presented in the previous section clearly show that PALMSNet can produce more accurate segmentation than the other eleven techniques. However, while the segmentation metrics that are used provide us with a measure

of how well each algorithm is in segmenting the MRI images, they do not provide a good visual indicator on how each of the algorithm fares in providing accurate delineation of the MRI images. This is because if we were to use the result of this segmentation to assist detection of lumbar spinal stenosis, the next step will rely heavily on the accuracy of the edges found around the areas shown in Fig. 3, rather than the entire segmented image.

To provide better visual cues on the suitability of the network for this purpose, we show the result of superimposing the resulting edges with the ground truth. In the interest of conciseness of the discussion, we only show the result using PALMSNet trained on all discs. These are shown in Fig. 7.

The images illustrated in Fig. 7 confirm our finding in the previous section that the results are markedly worse when applied to D5 compared to the other two discs.

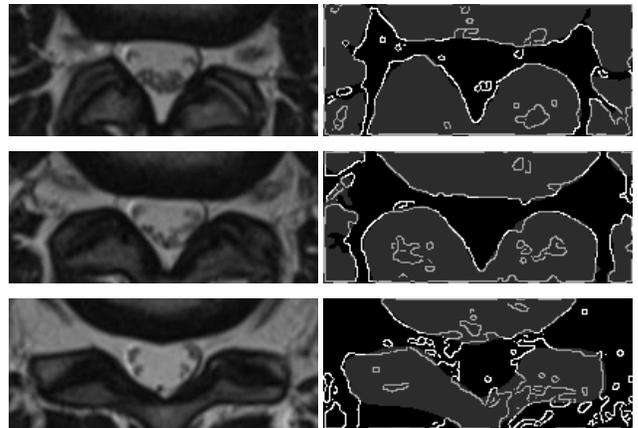


Fig. 7. The result of superimposing the edges of the segmentation result on to the ground truth (right), and their corresponding original MRI images (left). The top images are taken from D3, middle from D4, and bottom from D5.

The simulation result indicated that PALMSNet produces significantly better segmentation results than SegNet, another deep network architecture. Further investigation into the low metric score for SegNet reveals that this network suffers greatly from overfitting to the training data set which may be attributed to the high number of layers that the network has. The network overfitting is evidenced by the significantly high accuracy of the prediction on the training data (0.95) compared to that on the test data (0.75). As illustrated in Fig. 8.

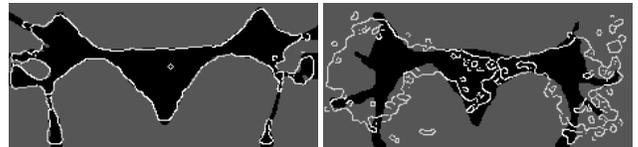


Fig. 8. Visual evidence of overfitting of SegNet in this experiment. The figure shows prediction on (left) a training image and (right) on a test image. The bright pixels outline the edges of the segmented image.

VI. CONCLUSION AND FURTHER WORK

In this paper, we have outlined the rationale of performing image segmentation as an important step to assist the delineation process to detect lumbar spinal stenosis in axial-view MRI. We did this by first describing the anatomy of the lumbar spine and the process carried out by clinicians when analysing lumbar spine MRI images. The contribution of this paper can be summarised as follows:

1. The architecture design of PALMSNet, a Deep Neural Network suitable to use for image segmentation of lumbar spine MRI. Our network is based on the patch-based image segmentation approach using Fully Connected Neural Network.
2. Experiment results trained using 400,680 image patches and tested on 171,720 image patches of size 25x25 pixels taken from T2-weighted axial lumbar spine MRI images of the last three intervertebral discs (D3-D5) of ten patients in total. Analysis of the results shows that the performance of PALMSNet exceeds those of another DNN and ten other more conventional image segmentation classifiers.

In future, we will use PALMSNet on significantly more data. We have collected lumbar spine MRI images of over 500 subjects. We will use the images for better training and validation of the network performance. We will also employ appropriate post-processing algorithms on the resulting segmentation to make it more suitable for medical image analysis purposes. The general plan is to use PALMSNet in our framework on computer-assisted detection of chronic lower back pain which was detailed in our previous publication [37].

ACKNOWLEDGMENT

This work is supported in part by PKLN funding from Indonesian Ministry of Research and Higher Education.

REFERENCES

- [1] K. P. Botwin and R. D. Gruber, "Lumbar spinal stenosis: anatomy and pathogenesis," *Phys. Med. Rehabil. Clin.*, vol. 14, no. 1, pp. 1–15, Jan. 2003.
- [2] G. J. Macfarlane, E. Thomas, P. R. Croft, A. C. Papageorgiou, M. I. V. Jayson, and A. J. Silman, "Predictors of early improvement in low back pain amongst consultants to general practice: the influence of pre-morbid and episode-related factors," *Pain*, vol. 80, no. 1, pp. 113–119, 1999.
- [3] N. Maniadakis and A. Gray, "The economic burden of back pain in the UK," *Pain*, vol. 84, no. 1, pp. 95–103, Jan. 2000.
- [4] G. Waddell, *The Back Pain Revolution*, 2nd ed. Churchill Livingstone (Elsevier), 2004.
- [5] NHS, "Quarterly Diagnostic Waiting Times," *NHS England*, 2017.
- [6] National Office for Statistics, "NHS Imaging and Radiodiagnostic activity in England 2014," no. August, 2013.
- [7] H. Jiang, W. Qi, Q. Liao, H. Zhao, W. Lei, L. Guo, and H. Lu, "Quantitative evaluation of lumbar disc herniation based on MRI image," in *Abdominal Imaging. Computational and Clinical Applications*, Springer, 2012, pp. 91–98.
- [8] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Automatic diagnosis of lumbar disc herniation with shape and appearance features from MRI," in *SPIE Medical Imaging*, 2010, p. 76241A--76241A.
- [9] J. Koh, R. S. Alomari, V. Chaudhary, and G. Dhillon, "Lumbar Spinal Stenosis CAD from ClinicalMRMand MRI Based on Inter- and Intra-Context Features with a Two-Level Classifier," vol. 7963, pp. 796304-796304–8, 2011.
- [10] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Lumbar spine disc herniation diagnosis with a joint shape model," in *Computational Methods and Clinical Applications for Spine Imaging*, Springer, 2014, pp. 87–98.
- [11] S. W. Atlas, *Magnetic resonance imaging of the brain and spine*, 5th ed., vol. 1. Lippincott Williams & Wilkins, 2016.
- [12] T. A. Aly and O. Amin, "Geometrical Dimensions and Morphological Study of the Lumbar Spinal Canal in the Normal Egyptian Population," *Orthopedics*, vol. 36, 2013.
- [13] T. Visuri, J. Ulaska, M. Eskelin, and P. Pulkkinen, "Narrowing of Lumbar Spinal Canal Predicts Chronic Low Back Pain More Accurately than Intervertebral Disc Degeneration: A Magnetic Resonance Imaging Study in Young Finnish Male Conscripts," *Mil. Med.*, vol. 170, no. 11, pp. 926–930, 2005.
- [14] M. Z. Janjua and F. Muhammad, "Measurements of the normal adult lumbar spinal canal," *J. Pak. Med. Assoc.*, vol. 39, no. 10, p. 264–268, Oct. 1989.
- [15] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," *Neuroimage*, vol. 54, no. 2, pp. 940–954, 2011.
- [16] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Advances in neural information processing systems*, 2012, pp. 2843–2851.
- [17] M. Freiman, H. Nickisch, H. Schmitt, P. Maurovich-Horvat, P. Donnelly, M. Vembar, and L. Goshen, "Learning a sparse database for patch-based medical image segmentation," in *International Workshop on Patch-based Techniques in Medical Imaging*, 2017, pp. 47–54.
- [18] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," *CoRR*, vol. abs/1706.0, 2017.
- [19] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *CoRR*, vol. abs/1411.4, 2014.
- [20] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
- [21] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *CoRR*, vol. abs/1502.0, 2015.

- [22] M. Khalaf, A. J. Hussain, R. Keight, D. Al-Jumeily, R. Keenan, C. Chalmers, P. Fergus, W. Salih, D. H. Abd, and I. O. Idowu, "Recurrent Neural Network Architectures for Analysing Biomedical Data Sets," *2017 10th Int. Conf. Dev. eSystems Eng.*, pp. 232–237, 2017.
- [23] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010, pp. 807–814.
- [24] A. de Brébisson and P. Vincent, "An Exploration of Softmax Alternatives Belonging to the Spherical Loss Family," *CoRR*, vol. abs/1511.0, 2015.
- [25] S. Ruder, "An overview of gradient descent optimization algorithms," *CoRR*, vol. abs/1609.0, 2016.
- [26] P.-N. Tan, M. Steinbach, and V. Kumar, "Classification : Basic Concepts , Decision Trees , and," *Introd. to Data Min.*, vol. 67, no. 17, pp. 145–205, 2006.
- [27] P. Fergus, S. Iram, D. Al-Jumeily, M. Randles, and A. Attwood, "Home-Based Health Monitoring and Measurement for Personalised Healthcare," *J. Med. Imaging Heal. Informatics*, vol. 2, pp. 35–43, 2012.
- [28] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.
- [29] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "RUSBoost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst. Man, Cybern. Part A Systems Humans*, vol. 40, no. 1, pp. 185–197, 2010.
- [30] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and a. Napolitano, "RUSBoost: Improving classification performance when training data is skewed," *2008 19th Int. Conf. Pattern Recognit.*, no. March 2016, pp. 8–11, 2008.
- [31] N. Bhatia and C. Author, "Survey of Nearest Neighbor Techniques," *IJCSIS) Int. J. Comput. Sci. Inf. Secur.*, vol. 8, no. 2, pp. 302–305, 2010.
- [32] D. Al-Jumeily, S. Iram, F.-B. Vialatte, P. Fergus, and A. Hussain, "A Novel Method of Early Diagnosis of Alzheimer's Disease Based on EEG Signals," *Sci. World J.*, vol. 2015, pp. 1–11, 2015.
- [33] M. Khalaf, A. J. Hussain, D. Al-jumeily, R. Keight, A. S. Al Kafri, and I. O. Idowu, "A Performance Evaluation of Systematic Analysis for Combining Multi-class Models for Sickle Cell Disorder Data Sets," vol. 10362, pp. 115–121, 2017.
- [34] X. Wu, V. Kumar, Q. J. Ross, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z. H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2008.
- [35] C. C. Aggarwal, *Data classification : algorithms and applications*. 2014.
- [36] A. S. Al-Kafri, S. Sudirman, A. J. Hussain, P. Fergus, D. Al-Jumeily, H. Al Smadi, M. Khalaf, M. Al-Jumaily, W. Al-Rashdan, M. Bashtawi, and J. Mustafina, *Lumbar spine discs labeling using axial view MRI based on the pixels coordinate and gray level features*, vol. 10363 LNAI. 2017.
- [37] A. S. Al-Kafri, S. Sudirman, A. Hussain, P. J. Fergus, D. Al-Jumeily, M. Al-Jumaily, and H. Al-Askar, "A Framework on a Computer Assisted and Systematic Methodology for Detection of Chronic Lower Back Pain Using Artificial Intelligence and Computer Graphics Technologies," in *Lecture Notes in Computer Science*, 2016, pp. 843–854.