BiLuNet: A Multi-path Network for Semantic Segmentation on X-ray Images

Van Luan Tran[1], Huei-Yung Lin[1], Hsiao-Wei Liu[2], Fang-Jie Jang[3], Chun-Han Tseng[1]
[1]*Department of Electrical Engineering, National Chung Cheng University, Chiayi 621, Taiwan*
[2]*Center for Measurement Standards, Industrial Technology Research Institute, Hsinchu 300, Taiwan*
[3]*Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu 300, Taiwan*
tranvanluan07118@gmail.com, lin@ee.ccu.edu.tw, rachel_liu@itri.org.tw, funjay@itri.org.tw, chun2360646@gmail.com

***Abstract*—Semantic segmentation and shape detection of lumbar vertebrae, sacrum, and femoral heads from clinical X-ray images are important and challenging tasks. In this paper, we propose a new multi-path convolutional neural network, BiLuNet, for semantic segmentation on X-ray images. The network is capable of medical image segmentation with very limited training data. With the shape fitting of the bones, we can identify the location of the target regions very accurately for lumbar vertebra inspection. We collected our dataset and annotated by doctors for model training and performance evaluation. Compared to the state-of-the-art methods, the proposed technique provides better mIoUs and higher success rates with the same training data. The experimental results have demonstrated the feasibility of our network to perform semantic segmentation for lumbar vertebrae, sacrum, and femoral heads. Code is available at: https://github.com/LuanTran07/BiLUnet-Lumbar-Spine.**

***Index Terms*—Semantic Segmentation, X-ray image, lumbar vertebra inspection.**

I. Introduction

Recently, deep learning algorithms, especially CNNs have remarkable results in computer vision applications, including medical applications. The region segmentation is an important task for object identification in medical images [6], [10], [13]. While the X-ray inspection is of great clinical necessity, X-ray images' automatic understanding is still a very challenging task [4], [17]. It is primarily due to X-ray imaging's projective nature, which causes significant overlaps among the object shapes, blurred boundaries, and complicated composition patterns [21]. For specific medical applications such as the treatment for low back pain, the images of spine bones are acquired by X-ray [5], computed tomography [8], or magnetic resonance imaging [3], [7], to facilitate the assessment of the causes by doctors. Consequently, the detection and segmentation of vertebrae in medical images is essential for analyzing the disease severity and recovery from the treatment.

The semantic segmentation of lumbar vertebrae, sacrum, and femoral heads have a number of published research with MRI images. In [11], Lu *et al.* presented a vertebral segmentation and disc-level localization based on a U-Net architecture combined with a spine-curve fitting method in a large volume of MRI imaging. The detected sacrum (S1) does not overlap with the lumbar vertebra. All lumbar intervertebral discs can be extracted with this method. According to these criteria, the successful rate on the test set is 94% (188 out of 200 cases). In [12], Memis *et al.* proposed a method for semantic segmentation of the proximal femurs and femoral heads with deep convolutional neural networks on low-quality MRI images. The segmentation accuracy of their method was observed over 90%. In [2], Deniz *et al.* proposed an automatic proximal femur segmentation method based on deep convolutional neural networks. They tested their segmentation performance against the gold standard of manual segmentation using four-fold cross-validation. The segmentation accuracy of approximately 94% was achieved in MR images of 86 subjects. The automatic detection and segmentation of vertebrae, sacrum, and femoral heads are more challenging for X-ray images than for CT or MRI images [21]. It is due to the overlapping shadows of chest organs, including lungs, bowels, and other bony structures such as ribs [9]. Nevertheless, the use of X-ray imaging techniques is popular because of the least amount of radiation. The semantic segmentation of femoral heads is even more difficult due to the overlap of the pelvis on the captured image, especially on X-ray images.

In the past few years, deep learning progress with convolutional neural networks (CNNs) has shown a remarkable performance on biomedical image segmentation. However, the existing literature for CNN based lumbar vertebra segmentation is mostly related to the techniques which take MRI images as inputs, and is very limited for the processing of X-ray images [14]. Fully automatic segmentation on X-ray images for lumbar vertebrae still has many challenges. In [1], Arif *et al.* proposed a novel shape predictor network for vertebra segmentation with X-ray images. They modified the U-Net [16] architecture to generate a signed distance function (SDF) from the input image. The predicted SDF was then converted to the parameters represented in a shape space for loss computation. After the shape was transferred to the corresponding symbolic distance function, a principal component analysis (PCA) was performed. However, their results were still very limited since it was not easy to derive the shape information for complicated cases.

In this paper, we present a multi-path convolutional neural network for semantic segmentation of lumbar vertebrae, sacrum, and femoral heads. As shown in Fig. 1, with the identification of lumbosacral spine and hip joints with our BiLuNet network, the identified shape, and location of these objects (L1, L2, L3, L4, L5, S1, hip axis) are also derived

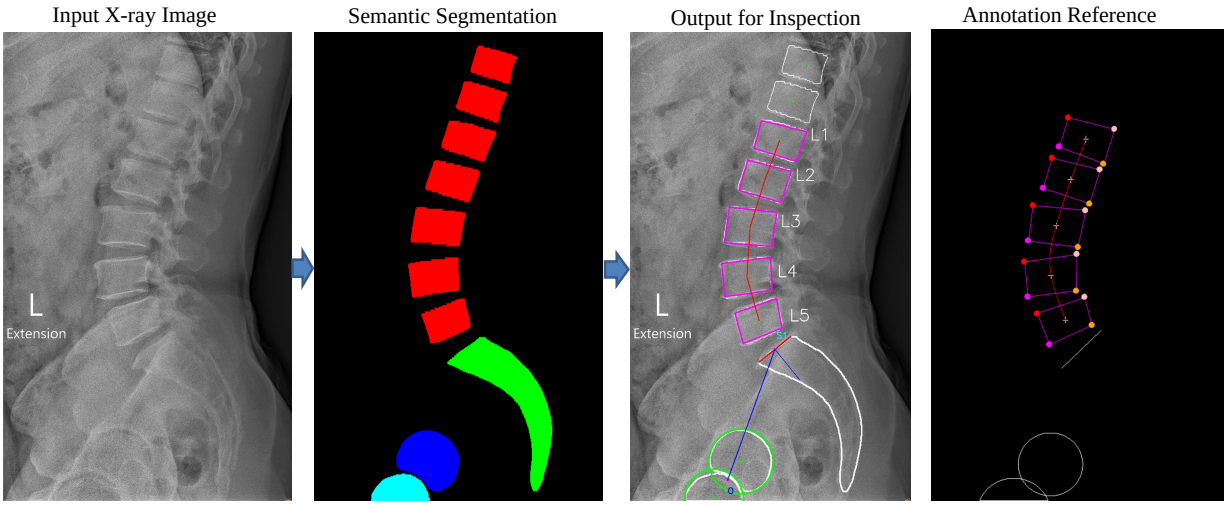


Fig. 1. An overview of lumbar vertebra inspection with X-ray images. With the input X-ray image, our BiLuNet generates the semantic segmentation result. The vertebrae, sacrum, and femoral heads are segmented with red, green, blue, and cyan color respectively. 'Output for Inspection' shows our lumbar vertebra inspection result. 'Annotation Reference' is provided by doctors.

for clinical diagnosis. The proposed BiLuNet architecture has the capability of medical image segmentation with very limited training data. By semantic segmentation, the five lumbar vertebrae locations, S1 on the sacrum, and femoral heads are obtained. The bones' shape fitting is then carried out to determine the location of five lumbar (L1 to L5), S1 on the sacrum and the hip axis. In the experiments, the training data are labeled at the pixel level manually and then annotated by doctors. The datasets are used to train our BiLuNet as well as other networks for performance evaluation and comparison. Experimental results have demonstrated the proposed network's effectiveness and feasibility for the lumbar vertebra inspection on X-ray images.

The highlight contributions of this work are as follows:

- We propose a multi-path convolutional neural network for semantic segmentation on X-ray images. The results are greatly improved, especially for the regions obscured or even not fully appeared in the image.
- The proposed network is capable of X-ray image segmentation with very limited training data, compared to the state-of-the-art techniques.
- Our method achieves high accuracy for the segmentation and shape fitting of lumbar vertebrae, sacrum, and femoral heads on X-ray images for the lumbar vertebra inspection.

II. BiLuNet for Lumbar Vertebra Segmentation

A. Network Structure

For semantic segmentation, this work is specifically concerned with an accurate segmentation method to extract the object mask. It is used to detect the locations of lumbar vertebrae, sacrum, and femoral heads. In this regard, U-Net is one of the well-known fully convolutional networks (FCNs) utilized in medical image understanding [16]. It has many variants in the literature, and most of them consist of an encoder-decoder structure. However, U-Net can hardly segment well for challenging cases or the image with a complex background. To deal with these issues, we propose a new multi-path convolutional neural network, BiLuNet, where Lu represents the shape of the network structure, and Bi denotes the bilateral segmentation network. As shown in Fig. 2, our network has two main paths, a spatial path to recover the spatial information and a downsampling path to extract the features.

In BiLuNet, Spatial Path is to preserve the spatial information and generate high-resolution features [19]. It contains three layers, and each convolutional layer has $stride = 2$, followed by batch normalization and Rectified Linear Unit (ReLU). Thus, this path extracts the output feature maps at 1/8 of the original image size. It encodes the rich spatial information from the large spatial size of the feature maps. The spatial path and downsampling path are crucial to the prediction of the detailed output to cover large objects, which leads to a rich discriminative ability.

The spatial information acquired by Spatial Path encodes most rich and detailed messages in the low levels, while the robust features extracted from the downsampling path are in the high levels. These two feature paths with different levels are fused by a Feature Fusion Module (FFM) [19]. Given the features from different levels, we first concatenate the outputs of the spatial path and the downsampling path. The batch normalization is then utilized to balance the scales of the features. Next, a global pool is used for the concatenated features to form a feature tensor, and the weight is computed by a sigmoid function. This weight tensor can re-weight the features by multiplying the concatenated features and adding to the feature selection. The details of this FFM function is shown in Fig. 2. The dimension of the output space is 5, which

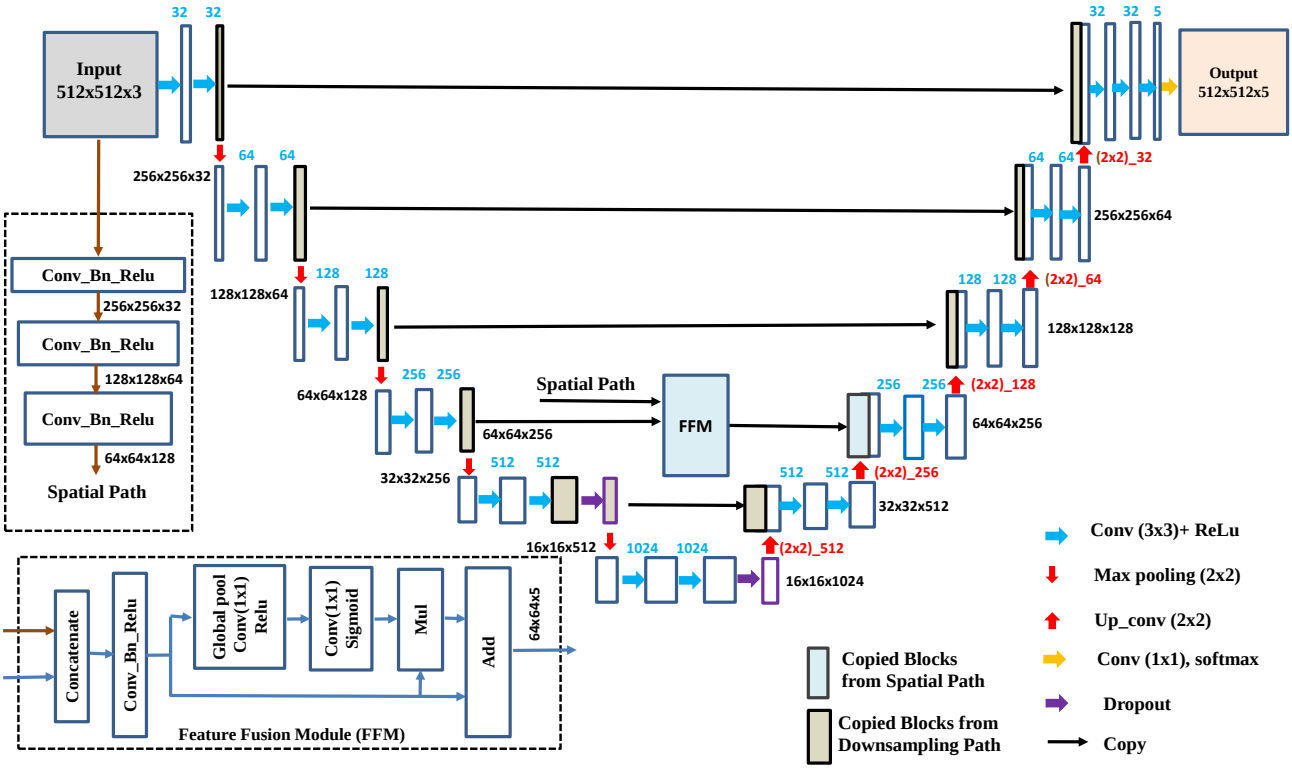


Fig. 2. The end-to-end BiLuNet architecture for object semantic segmentation. The yellow white boxes present the concatenation results of a copied feature map and an *up_conv*, an upsampled and convolutional layer. Purple boxes indicate the results of the node dropped out between two convolutional layers. BiLuNet takes a three-channel input image at the resolution of 512×512, and the output contains five channels, the classes with background, vertebra, sacrum, femoral_head1, and femoral_head2.

corresponds to the 5 classes of the segmentation requirement (background, vertebra, sacrum, femoral head1, and femoral head2).

Similar to U-Net, the proposed BiLuNet operates with both the downsampling and upsampling paths, and adds a spatial path to recover the spatial information lost in the pooling layers of the contracting path. The downsampling path decreases the feature size while increasing the number of feature maps, whereas the upsampling path increases the feature size while decreasing the number of feature maps. The latter eventually leads to a pixel-wise mask. For the upsampling operation, we modify the existing architecture to reconstruct high-resolution feature maps. The detailed network structure with the parameter settings is illustrated in Fig. 2.

To maintain the same number of channels in the encoder and decoder, skip connections are added between the corresponding blocks in the downsampling and upsampling paths. The skip connections are also used to explicitly copy features from the earlier layers to the subsequent layers. This prevents the neural network from having to learn the identity functions. Further functions are solely tasked with building on the representation. Through the skip connections, the nodes from a shallow layer are concatenated to the nodes in a deeper layer. The deeper layer is then treated as a wider layer and connected to the next layer. We use a dropout layer between these two convolutional layers to prevent overfitting and co-adaptation of the features. In the implementation, 50% of the nodes are dropped out during the forward and backward passes. The network has totally 29 convolutional layers. In the final layer, a 1×1 convolution, 5 dimensional space, and a sigmoid activation function are used to output the probability map of the semantic segmentation, which has the same size as the original 512×512 input. [1]

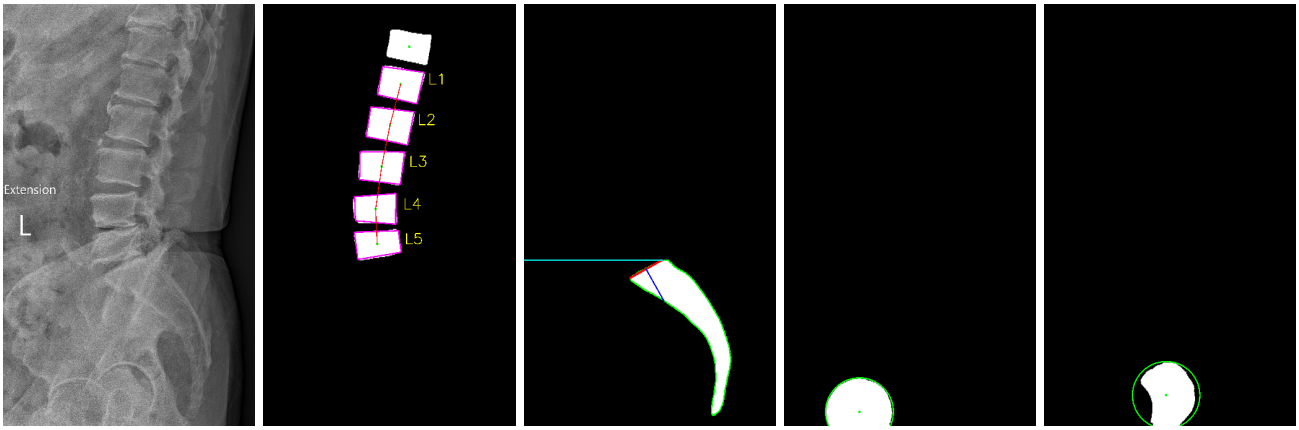
B. Loss Function

We use categorical cross-entropy loss function for model training. The categorical cross-entropy (CCE) is given by

$$CCE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_c \cdot \log(\hat{y}_c) + (1 - y_c) \cdot \log(1 - \hat{y}_c) \quad (1)$$

where y_j and $\hat{y}_j$ indicate the predicted and ground-truth values of the object in the image, respectively. C is the number of categories, and N is the observation number. The categorical cross-entropy cost function works in tandem with a softmax activation function. It calculates the loss between network predictions and target values for multi-class classification.

[1]The full training and testing code is open source at https://github.com/LuanTran07/ BiLUnet-Lumbar-Spine.



(a) (b) (c) (d) (e)

Fig. 3. An overview of the localization of lumbar vertebra and the shape fitting of lumbar vertebrae spine, sacrum, and two femoral heads presented in four layers. (a) the input X-ray image, (b) the first layer processing the lumbar vertebra from $L1$ to $L5$ and detecting five quadrilaterals of lumbar vertebra (in purple), (c) the second layer processing the sacrum (green contour) and the sacral endplate S1 (red line), with the blue line perpendicular to the center of S1 and the cyan line aligned with the sacral endplate, (d), (e) the layers of two femoral heads (green circles).

C. Location Detection and Shape Fitting of the Bones

After training our BiLuNet, we tested our model on X-ray images. As illustrated in Fig. 3, the segmentation is carried out on four layers (i.e., vertebrae, sacrum, femoral_head1 and femoral_head2), to detect the locations and shapes of the bones.

1) Lumbar Vertebrae: In the vertebra layer, we first find the contours and the regional centers and remove the noise for semantic segmentation based on the size of the regions (see Fig. 3(b)). The contours are sorted to determine the locations of the lumbar vertebrae from $L1$ to $L5$. To identify the structure of the vertebrae, the shape fitting is performed on each bone by approximating the contour with a quadrilateral [18]. The four corner points of the quadrilateral are then used to calculate the related parameters for the diagnosis of lumbar vertebrae.

2) S1 on the Sacrum: In the sacrum layer, there is only one ROI (region of interest) to identify. We also find the contour and remove the noise based on the contour size. The most extreme points along the contour are detected [15], and the left and right extremes are used to define the sacral endplate ($S1$), i.e., the red line segment shown in Fig. 3(c). The line perpendicular to the center of the sacral endplate (the blue line segment in Fig. 3(c)) is then determined and used to derive the pelvic incidence (PI).

3) Hip Axis: In the femoral head layers as shown in Figs. 3(d) and 3(d), similar post-processing is carried out for contour detection and noise removal. We adopt minimum enclosing circles to fit the femoral heads [20]. The hip axis is then determined by the middle point of the line segment connecting the femoral head centers. In the cases where only one femoral head appears in the image, the hip axis is given by the center of the available femoral head.

III. Experiments

In this work, we collected our own image data for the experiments. The dataset consists of 650 images (with a half before surgery and half after surgery) for network training and additional 100 images for testing (also with a half before surgery and half after surgery). The training images are labeled with four categories (vertebra, sacrum, femoral_head1, and femoral_head2) and annotated by doctors. For the CPU configuration, the computer is with an Intel i7-7700 CPU at 3.6 GHz and 32GB RAM. The GPU is an NVIDIA TITAN RTX with 24GB RAM. For training parameters, the number of epochs is set to 70, and the batch size is set to 4. Our initial learning rate is set to $3e-4$ for Adam optimizator. The processing time for semantic segmentation is $31ms$ on GPU and $980ms$ on CPU, and the total computation time for lumbar vertebra inspection is $1.06s$ on GPU and $2.01s$ on CPU.

We conduct several experiments to demonstrate the effectiveness of different modules in our framework. The ablation study is carried out with three cases. In the first case, the network is implemented without Spatial Path and Feature Fusion Module. The network in the second case does not include FFM, but use a concatenation function to combine Spatial Path and the downsampling path. The third case denotes the full BiLuNet. Table I reports the mIoUs (mean intersection over union) of the three cases using 100 images for evaluation. The average mIoUs are 78.93%, 82.10% and 85.05%, respectively. It demonstrates that the spatial path designed in our network does improve the segmentation results by preserving the spatial information from the X-ray image input. The ablation analysis also shows that our network can effectively combine two feature paths fused by Feature Fusion Module.

The results of BiLuNet tested on the images acquired before and after the operations are shown in Figs. 4(a)–4(d) and Figs. 4(e)–4(h), respectively. Two sample test X-ray images

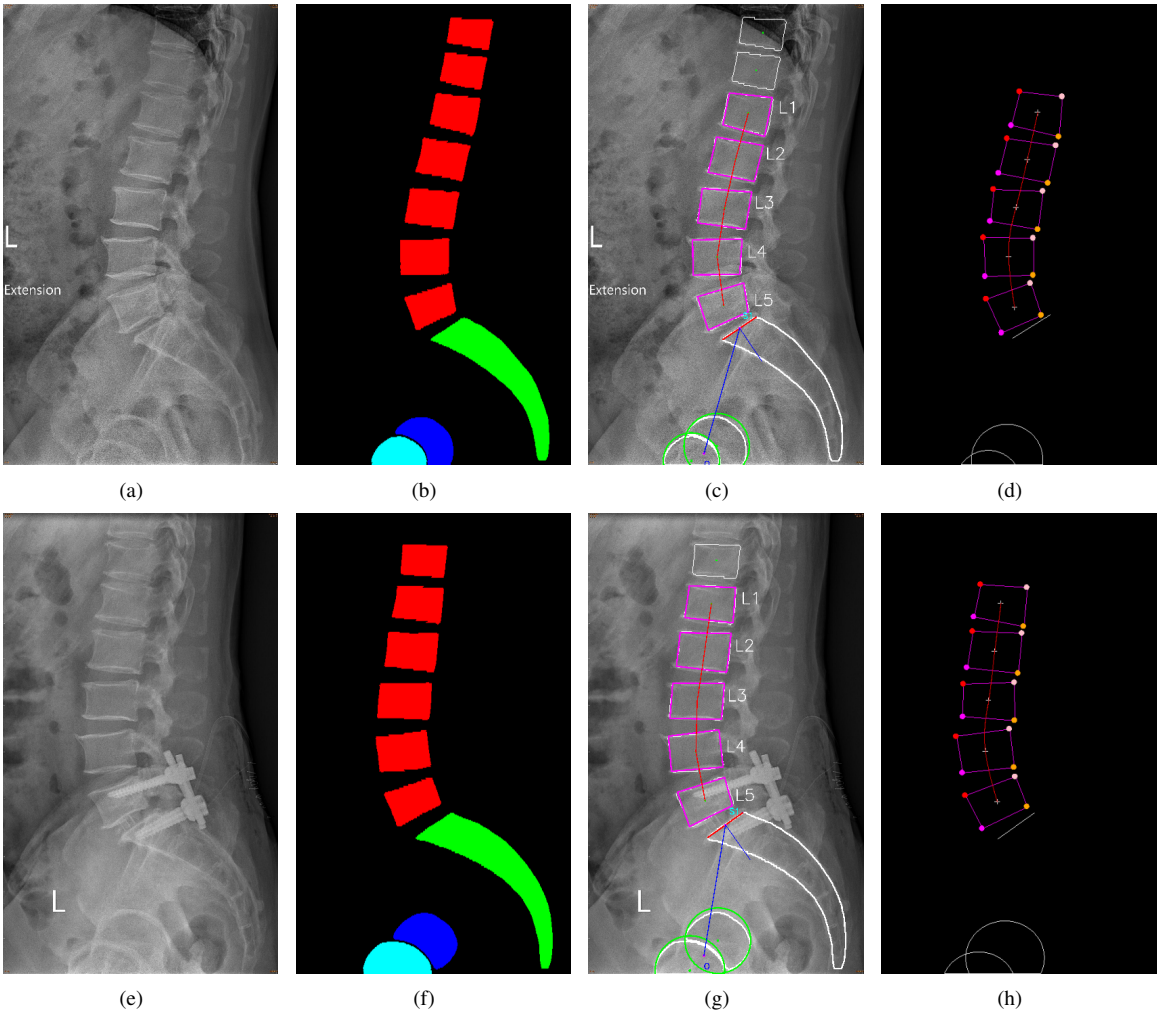


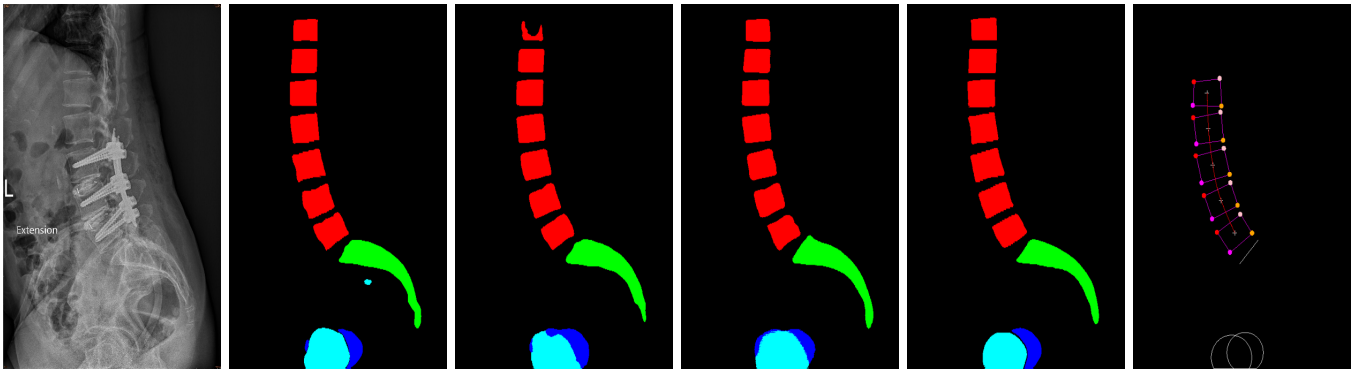
Fig. 4. Our identification of lumbosacral spine and hip joints with our BiLuNet network on X-ray images. (a, e) are the input X-ray images before and after surgery, respectively. (b, f) are the semantic segmentation of vertebrae, sacrum, left and right hip joints, in red, green, blue and cyan, respectively. (c, g) are our results with the identified shape and location of these objects (L1, L2, L3, L4, L5, S1, hip axis). (d, h) are the reference annotation by a doctor.

TABLE I
ABLATION STUDY ON DIFFERENT CASES

Methods	Vertebrae	Sacrum	Femoral head1	Femoral head2	Background	Average mIoU
Case 1	82.98%	81.95%	72.89%	60.91%	95.90%	78.93%
Case 2	83.15%	82.71%	77.26%	69.42%	97.96%	82.64%
BiLuNet	84.10%	83.67%	80.35%	78.80%	98.48%	85.05%

with the resolution of 2035×3408 acquired before and after surgery are shown in Figs. 4(a) and 4(e), respectively. The semantic segmentation is carried out with four target classes (vertebra, sacrum, femoral_head1, and femoral_head2), which are represented by red, green, blue and cyan, respectively (see Figs. 4(b) and 4(f)). After the detection and segmentation of vertebrae, sacrum, and femoral heads, the shape and location of these objects (L1, L2, L3, L4, L5, S1, hip axis) are identified. As shown in Figs. 4(c) and 4(g), our inspection results are very close to the ground-truth (Figs. 4(b) and 4(f)) as well as the annotation provided by doctors (Figs. 4(d) and 4(h)).

To evaluate the identified shape and location of the objects (L1, L2, L3, L4, L5, sacrum, femoral_head1, and femoral_head2), 100 X-ray images are used to test our algorithm. We define the success rate as the percentage of successful shape fitting of the bones and location detection of the objects on the testing images. The ordered lumbar vertebra locations, the locations of S1 on the sacrum, and femoral_head1 are detected with the success rate of 100%. The location of femoral_head2 is detected with the success rate of 91%. They are some cases in the challenging X-ray images with the location of the femoral_head2 region that is not fully in X-ray images or obscured as reported in Fig. 7.



(a) Original (b) U-Net (c) PSPNet (d) UNet++ (e) BiLuNet (f) GT

Fig. 5. An example of lumbar vertebra segmentation output on a challenging X-ray image. Our result shows a significant improvement compared to those derived from U-Net, PSPNet, and UNet++.

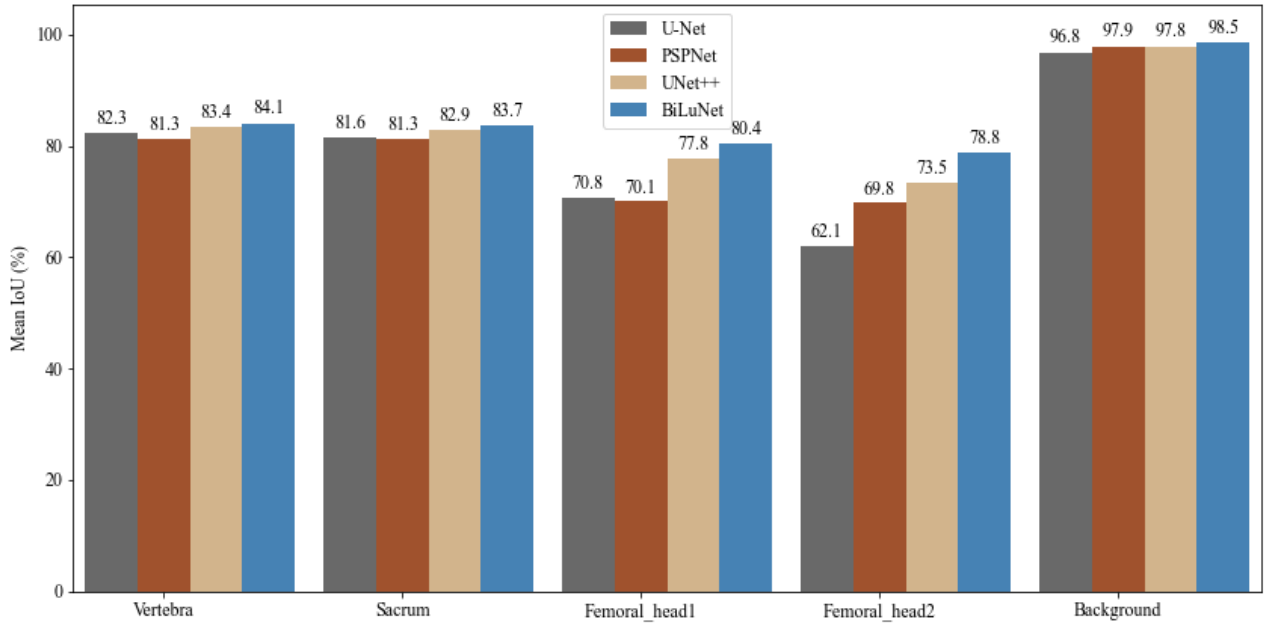


Fig. 6. The performance comparison on our dataset for different algorithms, U-Net [16], PSPNet [22], UNet++ [23], and our BiLuNet. It shows the IoU (intersection over union) and mean IoU for the evaluation between the semantic segmentation testing results and ground-truth labels.

The testing dataset is also used to evaluate other methods for performance comparison. Fig. 6 illustrates the mIoU (mean intersection over union) of U-Net, PSPNet, UNet++, and our BiLuNet. In 100 testing images, the proposed network has 84.10%, 83.67%, 80.35%, 78.80% and 98.48% mIoUs for vertebrae, sacrum, femoral_head1, femoral_head2 and background, respectively. The proposed BiLuNet outperforms the state-of-the-art methods in all categories. Fig. 5 also shows the higher accuracy semantic segmentation results compared to others. Our model can cover the most region of input X-ray images by the feature fusion between the spatial path and the downsampling path. The spatial path designed in our network preserves the spatial information from the original X-ray image. As shown in the figure, very precise results of lumbar vertebrae, sacrum, and femoral heads are obtained. More specifically, we are able to derive the accurate shapes of femoral_head1 and femoral_head2, which are failed by all other methods.

IV. Conclusion

In this paper, we propose a multi-path convolutional neural network, BiLuNet, for semantic segmentation on X-ray images. The solutions for shape fitting of the bones and location detection of five lumbar vertebrae, S1 on the sacrum and hip axis are also presented. It is able to achieve high accuracy for the automated segmentation of the lumbar vertebrae, sacrum, and femoral heads on X-ray images. With the bones' shape fitting, we can identify the location of the target regions (L1, L2, L3, L4, L5, sacrum, femoral_head1, and femoral_head2) with a high success rate. Our model with the global pooling

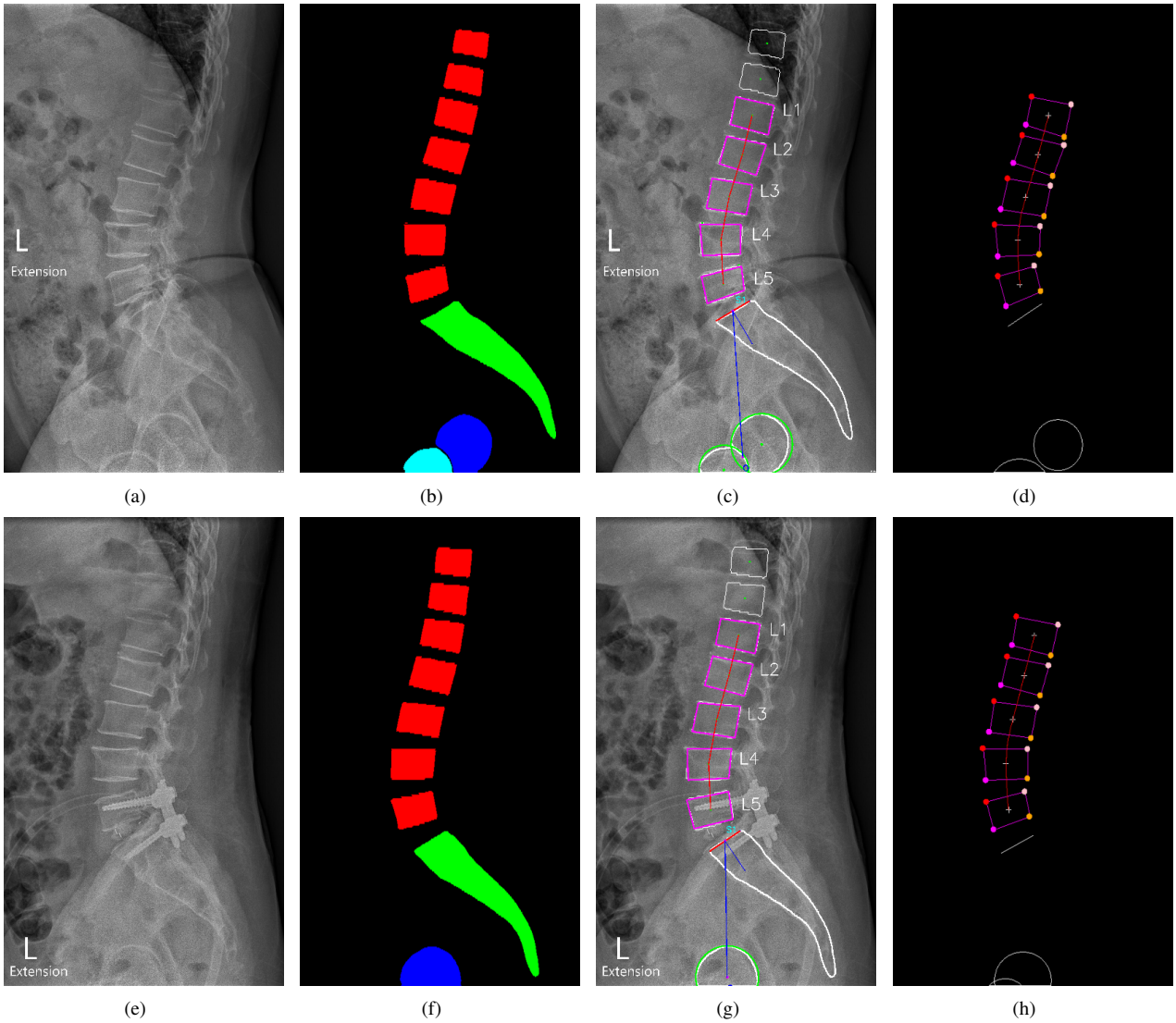


Fig. 7. Our identification of lumbosacral spine and hip joints with our BiLuNet network on the challenging X-ray image. (a, e) are an input X-ray images before surgery and after surgery. (b, f) are semantic segmentation of the vertebrae, sacrum, left hip joint, and right hip joint, which is segmented with red, green, blue, and cyan color respectively. (c, g) are our results with the identified shape and location of these objects (L1, L2, L3, L4, L5, S1, hip axis). (d, h) are the annotated results of experts and doctors.

and spatial path outperforms other state-of-the-art semantic segmentation methods in the performance evaluation. The future work will focus on developing an end-to-end network that takes X-ray images and surgical measurements as training data for lumbar vertebra inspection.

Acknowledgment

The support of this work in part by Dr. Meng-Huang Wu from Taipei Medical University, Taiwan, for providing the clinical experience and datasets is gratefully acknowledged. The support of this work in part by the Ministry of Science and Technology of Taiwan under Grant MOST 106-2221-E-194-004 and the Advanced Institute of Manufacturing with High-tech Innovations (AIM-HI) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan is gratefully acknowledged.

References

[1] S. M. M. R. Al-Arif, K. Knapp, and G. G. Slabaugh. Spnet: Shape prediction using a fully convolutional neural network. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, pages 430–439, 2018.

[2] C. M. Deniz, S. Hallyburton, A. Welbeck, S. Honig, K. Cho, and G. Chang. Segmentation of the proximal femur from MR images using deep convolutional neural networks. *CoRR*, abs/1704.06176, 2017.

[3] M. Fu, R. Buerba, W. III, D. Blizzard, A. Lischuk, A. Haims, and J. Grauer. Inter-rater and intra-rater agreement of magnetic resonance imaging findings in the lumbar spine: Significant variability across degenerative conditions. *The Spine Journal*, 14, 10 2014.

[4] Z. Han, B. Wei, S. Leung, J. Chung, and S. Li. Towards automatic report generation in spine radiology using weakly supervised framework. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV*, pages 185–193, 2018.

[5] M. Horng, C. Kuok, M. Fu, C. Lin, and Y. Sun. Cobb angle measurement of spine from x-ray images using convolutional neural network. *Comp. Math. Methods in Medicine*, 2019:6357171:1–6357171:18, 2019.

[6] K. Huang, H. Cheng, Y. Zhang, B. Zhang, P. Xing, and C. Ning. Medical knowledge constrained semantic breast ultrasound image segmentation. In *24th International Conference on Pattern Recognition, ICPR 2018, Beijing, China, August 20-24, 2018*, pages 1193–1198. IEEE Computer Society, 2018.

[7] A. Jamaludin, M. Lootus, T. Kadir, A. Zisserman, J. Urban, M. Batti, J. Fairbank, and I. McCall. Automation of reading of radiological features from magnetic resonance images (mris) of the lumbar spine without human intervention is comparable with an expert radiologist. *European Spine Journal*, 26, 02 2017.

[8] L. Kalichman, D. H. Kim, L. Li, A. Guermazi, and D. J. Hunter. Computed tomography-evaluated features of spinal degeneration: prevalence, intercorrelation, and association with self-reported low back pain. *The spine journal : official journal of the North American Spine Society*, 10 3:200–8, 2010.

[9] K. C. Kim, H. C. Cho, T. J. Jang, J. M. Choi, and J. K. Seo. Automatic detection and segmentation of lumbar vertebra from x-ray images for compression fracture evaluation. 2019.

[10] Z. Liu, Y. Song, C. Maere, Q. Liu, Y. Zhu, H. Lu, and D. Yuan. A method for PET-CT lung cancer segmentation based on improved random walk. In *24th International Conference on Pattern Recognition, ICPR 2018, Beijing, China, August 20-24, 2018*, pages 1187–1192. IEEE Computer Society, 2018.

[11] J. Lu, S. Pedemonte, B. Bizzo, S. Doyle, K. P. Andriole, M. H. Michalski, R. G. Gonzalez, and S. R. Pomerantz. Deep spine: Automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading using deep learning. In *Proceedings of the Machine Learning for Healthcare Conference, MLHC 2018, 17-18 August 2018, Palo Alto, California*, pages 403–419, 2018.

[12] A. Memis, S. Varli, and F. Bilgili. Semantic segmentation of the multiform proximal femur and femoral head bones with the deep convolutional neural networks in low quality MRI sections acquired in different MRI protocols. *Comput. Medical Imaging Graph.*, 81:101715, 2020.

[13] S. A. Nadeem, E. A. Hoffman, and P. K. Saha. An automated airway segmentation algorithm for CT images using topological leakage detection and volume freezing. In *24th International Conference on Pattern Recognition, ICPR 2018, Beijing, China, August 20-24, 2018*, pages 1181–1186. IEEE Computer Society, 2018.

[14] D. Nie, Y. Gao, L. Wang, and D. Shen. Asdnet: Attention based semi-supervised deep networks for medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV*, pages 370–378, 2018.

[15] A. Pulagam, G. Kande, E. Krishna Rao, and R. Inampudi. Automated lung segmentation from hrct scans with diffuse parenchymal lung diseases. *Journal of Digital Imaging*, 29, 03 2016.

[16] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III*, pages 234–241, 2015.

[17] A. Sekuboyina, M. Rempfler, J. Kukacka, G. Tetteh, A. Valentinitsch, J. S. Kirschke, and B. H. Menze. Btrfly net: Vertebrae labelling with energy-based adversarial learning of local spine prior. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV*, pages 649–657, 2018.

[18] X. Song, C. Cheng, C. Zhou, and D. Zhu. Gestalt-based douglas-peucker algorithm to keep shape similarity and area consistency of polygons. *Sensor Letters*, 11:1015–1021, 06 2013.

[19] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang. Bisenet: Bilateral segmentation network for real-time semantic segmentation. *CoRR*, abs/1808.00897, 2018.

[20] H. Zangana. A new algorithm for shape detection. 19:71–76, 05 2017.

[21] Y. Zhang, S. Miao, T. Mansi, and R. Liao. Task driven generative modeling for unsupervised domain adaptation: Application to x-ray image segmentation. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II*, pages 599–607, 2018.

[22] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. Pyramid scene parsing network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 6230–6239, 2017.

[23] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang. Unet++: A nested u-net architecture for medical image segmentation. *CoRR*, abs/1807.10165, 2018.