Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Spinal Neoplasm Image Inpainting with Deep Convolutional Neutral Networks*

Yanru Miao[1,2], Yu Sun[1], Shibo Li[1], Peng Zhang[1,*], Yuanyuan Yang[1,*], Ying Hu[1]

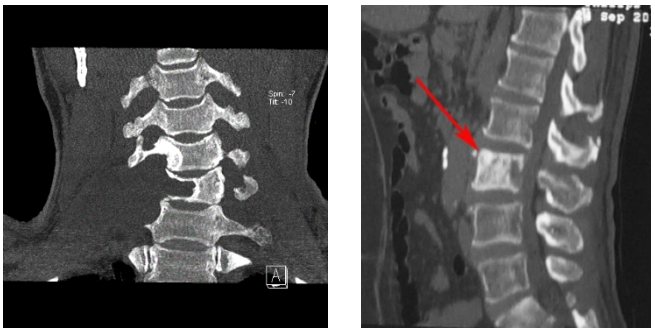
1 Shenzhen Key Laboratory of Minimally Invasive Surgical Robotics and System, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China;
2 University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: zhangpeng@siat.ac.cn (Peng Zhang), yy.yang1@siat.ac.cn (Yuanyuan Yang)

Abstract— **Spinal tumor is an abnormal mass of tissue in or surrounding the spinal cord or spinal column that is difficult to dissection for its complicated anatomic location. At present, Total En Bloc Spondylectomy (TES) is an effective treatment which can remove tumor lesions with tumor-free surgical margin and maintain the mechanical stability of the spine. This paper aims to address the spinal Computed Tomography (CT) image inpainting to reconstruct a normal and complete vertebral body used for TES by 3D printing. Firstly, the images of spine CT are segmented by neural network to obtain binary images. Then, the generative adversarial network (GAN) is trained using binary images of normal spinal CT. Next, the region of the lesion is detected from CT images with the spinal tumor by the deep convolutional neural network. Finally, the constructed GAN model is applied for the image inpainting of the spinal tumor CTs. As a result, the restored CT images are reconstructed in three dimensions to obtain a normal spine model. The experimental results prove that the proposed method is capable of reconstructing practical vertebral bodies and evolving potential applications.**

Keyword— **Spinal tumor; 3D printing; Semantic segment; Lesion detection; Image inpainting**

I. INTRODUCTION

Spinal tumors refer to those neoplasms developing in the region of spinal, including the bones, tissues, and nerves of the spine. As a spine tumor grows, it may compress the spinal cord or nerve roots, affect the blood vessels and cause life-threatening complications, even paralysis [1]. As shown in Figure 1(a), the patient's preoperative CT shows the tumor at L4 invading the vertebral body, resulting in that the vertebral body is destroyed. With the continuous development of surgical technique and medical devices in recent years, surgery is often the treatment option for tumors that can be resected with an acceptable risk of spinal cord or nerve injury damage. Total En Bloc Spondylectomy (TES) is a radical and complex surgery that can completely removes one or more affected vertebral bodies and tissues with negative margins [2]. Owing to the removal of the posterior vertebral elements and anterior parts, the continuity of the spine is interrupted and the stabilization of the spine is considerably necessary. There are many choices for reconstructing materials after TES, such as large inter-body fusion and bone substitute. Recently, the processing method of the artificial vertebral body has broken the traditional approaches and new 3D printing reconstruction materials are gradually applied for spinal surgery which has broad development prospect [3]. However, 3D printing has its limitations in clinical practice now. It is extremely difficult to completely segment the region of cancellous bone and soft tissue in CT images, so that the prosthesis produced will not fit perfectly with the patient. What's more, without sufficient medical knowledge, it is likely that the prosthesis produced by the revised 3D digital model may differ from the ideal one. Therefore, making multiple pre-operative models for discussion is often necessary which results in the increase of extra time and overall cost.



(**a**) (**b**)

Figure 1. There are two with different view showing the tumor at lumber which invades the vertebral bodies: (**a**) Coronal plane of CT; (**b**) Sagittal view of CT.

To solve the above problems, we propose a practical method restoring the CT images of the spinal tumor by generative adversarial network (GAN) [4] to reconstruct a normal and complete vertebral body for 3D-printed implant. The workflow of the method is introduced in Figure 2. The segmentation and lesion detection of spinal neoplasm image are prerequisites of image inpainting in our work which segment the spinal region and distinguish between cancer and normal areas respectively.

Early image inpainting methods were relatively simple using the statistics of the rest of the image to fill the gaps generally. However, these methods are limited by the available image statistics and does not have the concept of visual semantics. Raymond *et al*. [5] used deep convolutional generative adversarial networks (DCGAN) to implement face image inpainting with favorable efficient. Pathak *et al*. [6]

*This work is supported by the National Natural Science Foundation of China (No. U1613224 and U1813213), Shenzhen Fundamental Research Funds (No. JCYJ20180507182415428 and JCYJ20180507182215361), in part by Shenzhen Key Laboratory Project (No. ZDSYS201707271637577)

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

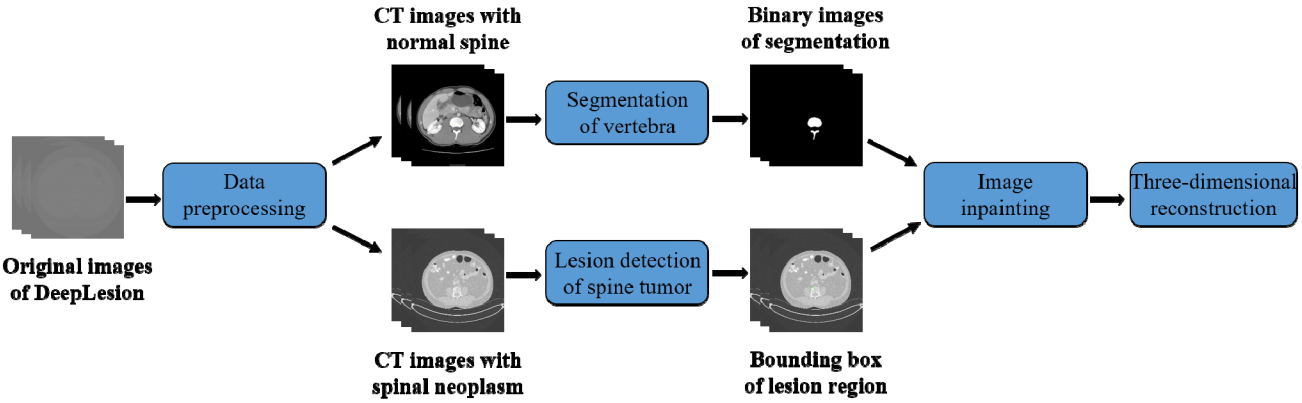


Figure 2. The workflow of the method.

proposed Context Encoder with combination of the encoder-decoder network structure and GAN. The encoder-decoder stage is used to learn the image features and generate the prediction map and the GAN is used for judging the possibility that the prediction graph comes from the training set or the prediction set. Further research, Lizuka *et al*. [7] used global and local discriminators as the adversarial loss of the GAN. The global discriminator is used to determine the semantic consistency of the entire image and the local discriminator focuses on the semantics of the small block generation region ensuring a high degree of consistency of the restored image. Liu *et al*. [8] proposed to use partial convolution for image inpainting which achieves good results aiming at some defects in the existing convolutional neural network (CNN) such as color mismatch and blur. Whereas, all the mentioned models are adopted for natural images or face images and there are rare applications of image inpainting in medical images.

In terms of image segmentation, Long *et al*. [9] proposed Fully Convolutional Networks (FCN) which extended the original CNN structure without a fully connected layer allowing the segmentation map to generate images with any size and improving the speed of training process. From this, Korez *et al*. [10] proposed a 3D FCN network structure and optimized the segmentation of vertebral bodies with deformable models which further improved the segmentation accuracy. Moeskops *et al*. [11] used three different data to train an FCN network at the same time, so that the trained network can be used for the segmentation of these three types of targets simultaneously. In 2015, U-Net [12] was proposed by Olaf *et al*. based on FCN which has been widely used for segmentation of medical images since it was released. Milletari *et al*. [13] used 3D convolutional to perform 3D image segmentation by V-Net based on U-Net subsequently. In the area of lesion detection, Roth *et al*. [14] designed a deep CNN classifier to identify sclerotic metastases following a coarse candidate generation system. Wang *et al*. [15] presented a two-stage approach including a patch-based classification and a heatmap-based processing for the identification of metastatic breast cancer. Wang *et al*. [16] developed a Siamese deep neural network for multi-resolution analysis and detection of spinal metastasis with three sub-networks. Chmelik *et al*. [17] proposed a deep CNN to automatic segment and classify spinal lesions in 3D CT images. Yan *et al*. [18] built a large-scale lesion image dataset namely DeepLesion and a unified network to detect various types of lesions at the same time. However, there is still not a state-of-art algorithms for the lesion detection of the spinal tumor.

To our knowledge, this is the first report of three-dimensional reconstruction of normal spinal model used for 3D printing by means of restoration of images with lesions. The rest of this paper is organized as follows. The proposed methods and procedures are described in Section 2, including the pre-process of data and the spinal neoplasm image inpainting. In Section 3, experiments and results are presented in detail. The discussion and analysis are made in Section 5.

II. Materials and Methods

A. Dataset and Data Preprocessing

In this paper, the spinal CT images used for neural network training are from DeepLesion which is a large-scale annotated image dataset and has been publicly released at present. There are multiple categories of lesions in this dataset, 32735 lesions in 32120 CT images specifically, including bone, liver, lung and others. Each lesion in DeepLesion is comprehensively described with the image coordinates of the two response evaluation criteria in solid tumors (RECIST) diameters of the lesion, the bounding-box of the lesion and the type of the lesion. Also, the index and age of the patient, image size and the windowing in Hounsfield unit (HU) extracted from the original DICOM file are provided. Because the type of spinal neoplasm is focused on, 282 images with spinal tumor were selected from DeepLesion to train models for lesion detection and image inpainting. Meanwhile, 3863 images without lesion in the spinal region are used for segmentation of vertebral bodies. One more thing, the images in DeepLesion have very low contrast with 16-bit which are needed to be preprocessed. Firstly, all images are converted to int32 format and subtracted 32768 from the pixel intensities to obtain the original HU values. Then, do intensity windowing on HU values so that make the intensities in a certain range to 0 and 255 for viewing. The windowed intensity is calculated by the default window of each image. Last, the windowed images are saved as an 8-bit image file. The basic architecture of the proposed method is shown in Figure 3.

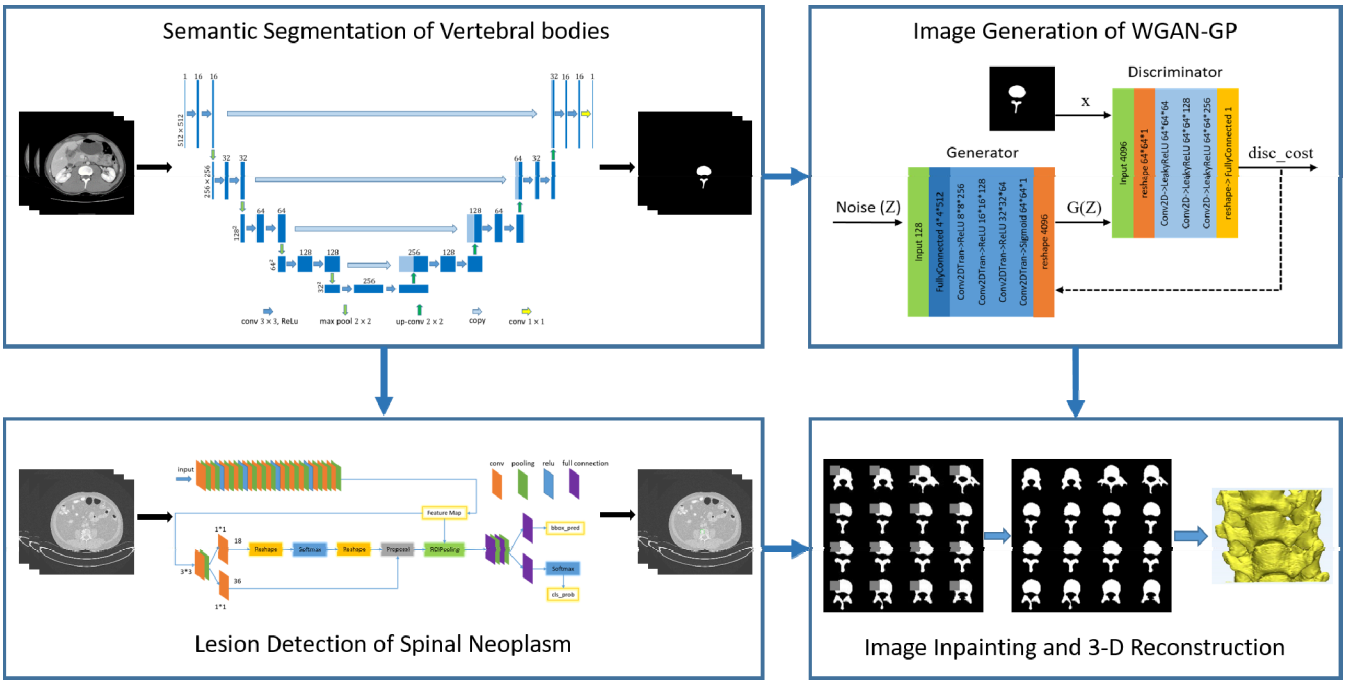


Figure 3. The architecture of the method.

B. *Semantic Segmentation of Vertebral Bodies*

Medical image segmentation is based on certain features that segment the part of the 2D or 3D medical image with a specific region, such as grayscale, texture, frequency domain features, etc. In this paper, V-Net is used for 2D spinal images segmentation to get binary images for image inpainting and three-dimensional reconstruction subsequently. V-Net is a 3D version of U-Net aiming at the problem of 3D image segmentation and it uses the structure of encoder-decoder similar to U-Net. The downsampling process is performed by five sets of convolution operations to capture context information. The operations of convolution are used with a kernel size of 2*2 with a stride of 2 instead of pooling which can reduce memory footprint. The upsampling process uses de-convolution operations which double the image size each time, combining with the information of the layers in the downsampling process to restore the details. After the upsampling process, a feature map with the size of 512*512*16 is obtained, and finally a 1*1 convolution kernel is used to reduce the number of channels to be 1. In the process of downsampling, the kernels in the five sets of convolution operations are with the size of 3*3 and padding. The advantage of this is that the size of feature maps is the same with the ones in the corresponding layer of the upsampling process and the work of cropping images is omitted.

C. *Lesion Detection of Spinal Neoplasm*

In terms of a large number of diagnostic tasks, deep learning has achieved doctor-level accuracy which is attributed to CNN's performance comparable to humans in image classification and target detection tasks to a great extent. The framework of Faster R-CNN [19] is used to detect the lesion of spinal neoplasm in this paper. Firstly, the feature maps of original images are extracted by a pre-trained CNN for the image classification task. The excellent networks include VGG-16, ResNet, MobileNet, DenseNet, etc. and VGG-16 are applied in this paper. The length and width of feature maps decrease with pooling between the convolutional layers while the depth increases with the number of filters. The image of M*N becomes a feature map of (M/16)*(N/16) which is the output of the conv5 layer of VGG-16. Then, the region proposal network (RPN) takes the feature map as input to propose the region of interest (ROI) with improved generation speed of candidate compared with selective search method. For each sliding-window of each feature map after the first convolutional layer of RPN, K anchors are generated with three scales and three aspect ratios so that K equals 9. Subsequently, two parallel processes are the following. One is to determine the probability of background and foreground while the other is bounding box regression. According to the probability of the anchors as the foreground sorted from large to small, the first L anchors are selected. Based on the degree of overlap between anchors, nonmaximum suppression (NMS) is used to remove redundant ones leaving the final P anchors. Next, because the fully connected layer needs proposals with the same size, ROI pooling is used to make a fixed-size feature map. The region-based CNN classifies the feature map and outputs a probability for each possible target class finally.

D. *Spinal Neoplasm Image Inpainting*

Image inpainting is a very intuitive problem which involves other information to fill a hole in an image and makes it plausible for human vision. To solve the problem of spinal neoplasm image inpainting, an algorithm is utilized using WGAN-GP [20]. WGAN-GP is the improvement of WGAN [21] from the perspective of loss function which has faster convergence speed and can generate higher-quality samples. In order to deal with the Lipschitz constraints in WGAN, the weights are clipped to a certain range and most of the weights have only two possible numbers. Therefore, WGAN-GP adds

gradient penalty to original loss function with Wassertein distance in WGAN. The final loss function is:

$$L = W\left(P_r, P_g\right) + \lambda \underset{\hat{x}\sim P_{\hat{x}}}{E}\left[\left(\left\|\nabla_{\hat{x}} D(\hat{x})\right\|_2 - 1\right)^2\right] \quad (1)$$

As is shown in Equation (1), the first part is the Wassertein distance in WGAN and the second part is the additional gradient penalty.

The generator (G) and discriminator (D) are two simple CNNs without pooling layer in WGAN-GP which provides a good network topology. In the part of generator, four transposed convolution operations are used to convert the input noise into a false image. Accordingly, the discriminator downsamples the input images by four convolution operations to distinguish whether an input image is real or not. The size of the convolution kernel is 5*5 throughout the whole framework. An inconvenient truth is that the introduction of the gradient penalty makes it conflict with the batch normalization (BN). Therefore, BN is not used in the process of WGAN-GP training.

After training the WGAN-GP with normal spinal CT images, the model that can convert random noise into a spinal image is obtained. Then, the model is used to accomplish the spinal neoplasm image inpainting. M denotes a binary mask with the same size of the image F which needs to be restored. The value of M is only zero or one. Zero means the region in the image to be repaired while one means the part to be retained. Thus, the repaired image F' with an image $G(z)$ from the generator of WGAN-GP are defined as:

$$F' = M \odot F + (1 - M) \odot G(z). \quad (2)$$

As can be seen, the first part of Equation (2) is Hadamard product which is constant and the second part changes with $G(z)$. In order to calculate the missing region of an image preferably, the information of context and perception are considerably important. The acquisition of them is mainly by the loss function. To ensure the same context information of F', the pixels in F is ought to be as similar as possible to the corresponding pixel of $G(z)$. Therefore, it is necessary to penalize the generation of dissimilar pixels. The loss function of context is:

$$L_{contextual} = \left\|M \odot G(z) - M \odot F\right\|_1. \quad (3)$$

Ideally, all the pixel values of F and $G(z)$ are the same and the value of the loss function is zero. To obtain a high-quality repaired image, the discriminator needs to have the ability to correctly distinguish the real image. The perceptual loss function is:

$$L_{perceptual} = \log\left(1 - D(G(z))\right) \quad (4)$$

So that the final loss function is:

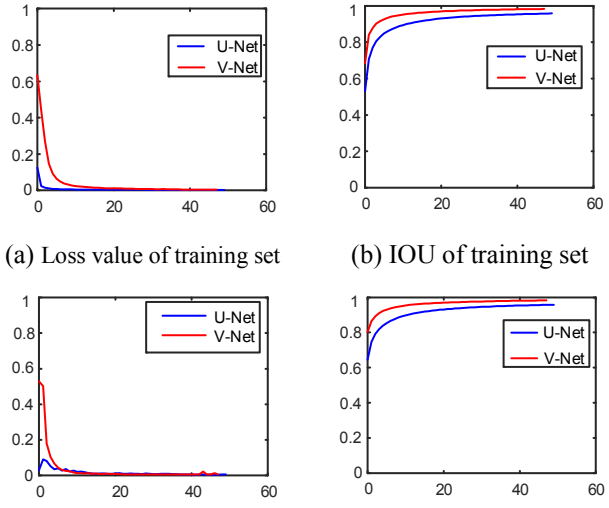
$$L = L_{contextual} + \lambda L_{perceptual} \quad (5)$$

λ is a hyperparameter used to control the respective importance of the two parts.

III. Experiments and Results

A. Segmentation Results

U-Net and V-Net are implemented through the Keras framework using 3863 CT images with normal vertebral bodies. Nevertheless, the total parameters of V-Net are about 35 times that of U-Net. In the paper, ten-fold cross-validation is used to divide the training set into ten parts, take nine parts as training data and one part as the validation set. The sizes of all the input and output in the two networks are 512*512. Figure 4 shows the evaluation indicators of the U-Net and V-Net on the training set and validation set which include loss and IOU. The X axis represents the number of epochs, and Y axis stands for the value of evaluation index. The blue solid line denotes U-Net, while the red signifies V-Net. In short, IOU is the accuracy of segmentation, which is the ratio of intersection of the network segmentation result and ground truth to the union of them. The closer the value of IOU is to 1, the better the segmentation result is. The specific values are demonstrated in Table 1.



(a) Loss value of training set (b) IOU of training set

(c) Loss of verification set (d) IOU of verification set

Figure 4. Training result of training set and verification set.

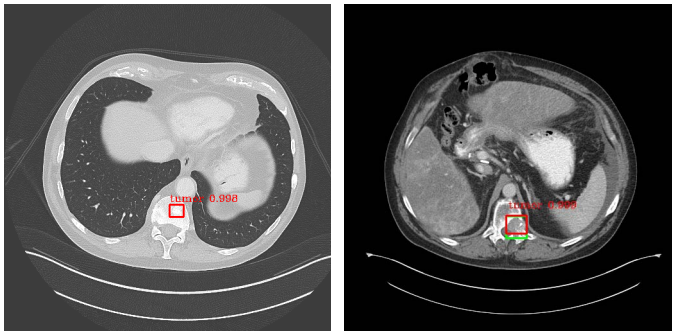
TABLE I. The Values of Loss and IOU

	U-Net	V-Net
Loss value of training set	0.0030	0.0044
Loss value of verification set	0.0055	0.0056
IOU of training set	0.9582	0.9829
IOU of verification set	0.9584	0.9831

B. Lesion Detection Results

There are 282 annotated images selected from DeepLesion for spinal tumor detection. In order to improve the accuracy and the generalization ability of the lesion detector, it is necessary to extend the amount of available data through data augmentation. By random rotation, flip, crop and adding noise to images, the original image can be converted to multiple images. The backbone of the adopted Faster R-CNN is VGG16 and average precision (AP) and Frame per second

(FPS) are calculated on the same platform to evaluate their performance. Figure 5 shows the examples of detection results which contain sclerotic and lytic lesion of test set with high confidence. The ground truth and predicted bounding boxes are green and red, respectively. AP is the average precision of spinal tumor detection with all test images and FPS is the number of images which can be recognized in a second. In terms of AP, Faster R-CNN with the backbone of VGG-16 is 0.7688. The value of FPS is 60.33.



(**a**) (**b**)

Figure 5. Detection results of test set. The ground truth and predicted bounding boxes are green and red, respectively. (**a**) Sclerotic lesion; (**b**) Lytic lesion.

C. Image Generation of GANs

Through the above segmentation network, binary images of the spine can be obtained. Since the spinal region is too small compared to the background region, the images need to be cropped and scaled from the initial size of 512*512 to a size of 64*64 for GAN training and the generation of just-as-good fake images. In this paper, DCGAN, WGAN-GP, PGGAN [22] are applied to generate images. Although DCGAN has a good network architecture, it does not fundamentally solve the problem of collapse mode, and it is still necessary to carefully balance the training process of the generator and the discriminator during training. The quality of the images generated by WGAN-GP and PGGAN is better by comparison, but PGGAN takes more time than WGAN-PG on the same platform. Moreover, WGAN-GP with a lighter architecture has fewer parameters than the other two. What's more, the process of the training is considerably stable so that the adjustment of parameters and the balance of the training level of the generator and discriminator are not required.

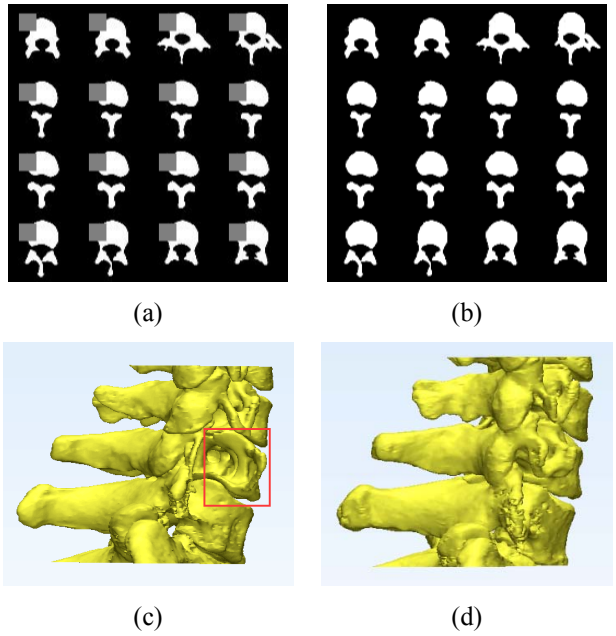
TABLE II. THE VALUES OF LOSS AND IOU

Model	DCGAN	WGAN-GP	PGGAN
Activation Function of Generator	ReLU & tanh	ReLU & sigmoid	LeakyReLU
Activation Function of Discriminator	LeakyReLU & Sigmoid	LeakyReLU	LeakyReLU & sigmoid
Size of Kernel	5*5	5*5	3*3
Batchnormal	Included	None	None
Progressive Training	No	No	Yes

D. Results of Image Inpainting

Through the above-mentioned lesion detection of the vertebral body, the CT images of the vertebral body in which the tumor is located can be obtained. Next, the abnormal spine CT images are segmented and clipped to the size of 64*64. Then, the binary masks to them according to the above Hadamard product are added, corresponding to the first part of Equation (2). After the training process of GANs with normal spine CT images, the masked area is restored by the images generated by the trained GANs. With the network constant iterating, the loss value of Equation (5) decrease and the image of the mask region is more consistent with other regions. To verify the restoration ability of above GANs, normal spine CT images are used first. As is shown in Figure 6(a), every image is covered with a binary mask of 16*16 and the final image inpainting results are shown in Figure 6(b). At the same time, the spine model is reconstructed with lytic lesion marked by the red bounding box using marching cube algorithm [23] in Figure 6(c) and the restored model is in Figure 6(d).

Due to there is not a meaningful evaluation metrics for the image inpainting result, the Kernel Maximum Mean Difference (KMMD) is used to evaluate the quality of image inpainting in this paper. It was first proposed for the problem of two-sample test to determine whether the two distributions are the same. The smaller the value, the more the two distributions are considered to be the same. The KMMDs of the original images and restored images are calculated to determine the similarity between the two distributions. The result of WGAN-GP is 0.001084.



(a) (b)

(c) (d)

Figure 6. Results of image inpainting. (a) image with binary mask; (b) image inpainting results; (c) spinal model with lytic lesion; (d) restored model

IV. CONCLUSION

Using the images selected from DeepLesion, the semantic segmentation of vertebral bodies and lesion detection of spinal neoplasm are achieved with good performance. By the trained generative model, defective spinal images are restored to the

original form and preliminary reconstructed to a complete spinal three-dimensional model which can be the foundation of successive 3-D printing. The output of the spinal repaired model shows that the proposed method is feasible and effective. However, for the Computer Aided Detection (CAD) techniques, a key limitation is that the diagnosis is performed only with medical images. Clinically, in order to provide an accurate diagnosis, doctors would access not only the medical images but also other necessary medical data. Deep learning techniques may provide more precise predictions with the ability to effectively integrate different types of data. Another limitation is whether there is a large enough annotation data set like DeepLesion. Small annotated data sets for specific tasks are easier to collect, but the performance of the algorithm on newer data is poorer. Limited to the amount of data, the CAD techniques are mainly concentrated in the study of diseases with high incidence while less research in some rare diseases. Data augmentation can improve the generalization of the model and building large annotated medical data set need more concern. With the continuous breakthrough of technology, artificial intelligence is expected to assist physicians as a collaborative means to reduce the burden of repetitive and monotonous tasks. Deeper and complex exploration of the spinal neoplasm image analysis is encouraged.

REFERENCES

[1] Valesin Filho, E. S.; de Abreu, L. C.; Lima, G. H.; de Cubero, D. I.; Ueno, F. H.; Figueiredo, G. S.; Valenti, V. E.; Monteiro, C. B. de M.; Wajnsztejn, R.; Fujiki, E. N.; et al. Pain and Quality of Life in Patients Undergoing Radiotherapy for Spinal Metastatic Disease Treatment. Int Arch Med 2013, 6, 6. https://doi.org/10.1186/1755-7682-6-6.

[2] Delgado-López, P. D.; Rodríguez-Salazar, A.; Martín-Velasco, V.; Castilla-Díez, J. M.; Martín-Alonso, J.; Galacho-Harriero, A.; Gil-Polo, C.; Araus-Galdós, E. Total En Bloc Spondylectomy for Spinal Tumours: Technical Aspects and Surgical Details. Neurocirugía 2017, 28 (2), 51–66. https://doi.org/10.1016/j.neucie.2016.07.002.

[3] Wei, R.; Guo, W.; Ji, T.; Zhang, Y.; Liang, H. One-Step Reconstruction with a 3D-Printed, Custom-Made Prosthesis after Total En Bloc Sacrectomy: A Technical Note. Eur Spine J 2017, 26 (7), 1902–1909. https://doi.org/10.1007/s00586-016-4871-z.

[4] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Advances in Neural Information Processing Systems 27; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., Weinberger, K. Q., Eds.; Curran Associates, Inc., 2014; pp 2672–2680.

[5] Yeh, R. A.; Chen, C.; Lim, T. Y.; Schwing, A. G.; Hasegawa-Johnson, M.; Do, M. N. Semantic Image Inpainting with Deep Generative Models. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE: Honolulu, HI, 2017; pp 6882–6890. https://doi.org/10.1109/CVPR.2017.728.

[6] Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A. A. Context Encoders: Feature Learning by Inpainting. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE: Las Vegas, NV, USA, 2016; pp 2536–2544. https://doi.org/10.1109/CVPR.2016.278.

[7] Iizuka, S.; Simo-Serra, E.; Ishikawa, H. Globally and Locally Consistent Image Completion. ACM Trans. Graph. 2017, 36 (4), 1–14. https://doi.org/10.1145/3072959.3073659.

[8] Liu, G.; Reda, F. A.; Shih, K. J.; Wang, T.-C.; Tao, A.; Catanzaro, B. Image Inpainting for Irregular Holes Using Partial Convolutions. In Computer Vision – ECCV 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer International Publishing: Cham, 2018; Vol. 11215, pp 89–105. https://doi.org/10.1007/978-3-030-01252-6_6.

[9] Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE: Boston, MA, USA, 2015; pp 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.

[10] Korez, R.; Likar, B.; Pernuš, F.; Vrtovec, T. Model-Based Segmentation of Vertebral Bodies from MR Images with 3D CNNs. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016; Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., Wells, W., Eds.; Springer International Publishing: Cham, 2016; Vol. 9901, pp 433–441. https://doi.org/10.1007/978-3-319-46723-8_50.

[11] Moeskops, P.; Wolterink, J. M.; van der Velden, B. H. M.; Gilhuijs, K. G. A.; Leiner, T.; Viergever, M. A.; Išgum, I. Deep Learning for Multi-Task Medical Image Segmentation in Multiple Modalities. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016; Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., Wells, W., Eds.; Springer International Publishing: Cham, 2016; Vol. 9901, pp 478–486. https://doi.org/10.1007/978-3-319-46723-8_55.

[12] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Navab, N., Hornegger, J., Wells, W. M., Frangi, A. F., Eds.; Lecture Notes in Computer Science; Springer International Publishing, 2015; pp 234–241.

[13] Milletari, F.; Navab, N.; Ahmadi, S.-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In 2016 Fourth International Conference on 3D Vision (3DV); IEEE: Stanford, CA, USA, 2016; pp 565–571. https://doi.org/10.1109/3DV.2016.79.

[14] Roth, H. R.; Yao, J.; Lu, L.; Stieger, J.; Burns, J. E.; Summers, R. M. Detection of Sclerotic Spine Metastases via Random Aggregation of Deep Convolutional Neural Network Classifications. In Recent Advances in Computational Methods and Clinical Applications for Spine Imaging; Yao, J., Glocker, B., Klinder, T., Li, S., Eds.; Springer International Publishing: Cham, 2015; Vol. 20, pp 3–12. https://doi.org/10.1007/978-3-319-14148-0_1.

[15] Wang, D.; Khosla, A.; Gargeya, R.; Irshad, H.; Beck, A. H. Deep Learning for Identifying Metastatic Breast Cancer. arXiv: 1606.05718 [cs, q-bio] 2016.

[16] Wang, J.; Fang, Z.; Lang, N.; Yuan, H.; Su, M.-Y.; Baldi, P. A Multi-Resolution Approach for Spinal Metastasis Detection Using Deep Siamese Neural Networks. Computers in Biology and Medicine 2017, 84, 137–146. https://doi.org/10.1016/j.compbiomed.2017.03.024.

[17] Chmelik, J.; Jakubicek, R.; Walek, P.; Jan, J.; Ourednicek, P.; Lambert, L.; Amadori, E.; Gavelli, G. Deep Convolutional Neural Network-Based Segmentation and Classification of Difficult to Define Metastatic Spinal Lesions in 3D CT Data. Medical Image Analysis 2018, 49, 76–88. https://doi.org/10.1016/j.media.2018.07.008.

[18] Yan, K.; Wang, X.; Lu, L.; Summers, R. M. DeepLesion: Automated Mining of Large-Scale Lesion Annotations and Universal Lesion Detection with Deep Learning. J. Med. Imag. 2018, 5 (03), 1. https://doi.org/10.1117/1.JMI.5.3.036501.

[19] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39 (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

[20] Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. C. Improved Training of Wasserstein GANs. In Advances in Neural Information Processing Systems 30; Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc., 2017; pp 5767–5777.

[21] Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN. arXiv: 1701.07875 [cs, stat] 2017.

[22] Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. Progressive Growing of GANs for Improved Quality, Stability, and Variation. arXiv: 1710.10196 [cs, stat] 2017.

[23] Lorensen, W. E.; Cline, H. E. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. In Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques; SIGGRAPH '87; ACM: New York, NY, USA, 1987; pp 163–169. https://doi.org/10.1145/37401.37422.