



Pelvic Fracture Segmentation Using a Multi-scale Distance-Weighted Neural Network

Yanzhen Liu¹ , Sutuke Yibulayimu¹ , Yudi Sang³ , Gang Zhu³,
Yu Wang¹ , Chunpeng Zhao², and Xinbao Wu²

¹ Key Laboratory of Biomechanics and Mechanobiology, Ministry of Education, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100083, China
{[yanzhenliu, wangyu](mailto:yanzhenliu@buaa.edu.cn)}@buaa.edu.cn

² Department of Orthopaedics and Traumatology, Beijing Jishuitan Hospital, Beijing, China

³ Beijing Rossum Robot Technology Co., Ltd., Beijing, China

Abstract. Pelvic fracture is a severe type of high-energy injury. Segmentation of pelvic fractures from 3D CT images is important for trauma diagnosis, evaluation, and treatment planning. Manual delineation of the fracture surface can be done in a slice-by-slice fashion but is slow and error-prone. Automatic fracture segmentation is challenged by the complex structure of pelvic bones and the large variations in fracture types and shapes. This study proposes a deep-learning method for automatic pelvic fracture segmentation. Our approach consists of two consecutive networks. The anatomical segmentation network extracts left and right ilia and sacrum from CT scans. Then, the fracture segmentation network further isolates the fragments in each masked bone region. We design and integrate a distance-weighted loss into a 3D U-net to improve accuracy near the fracture site. In addition, multi-scale deep supervision and a smooth transition strategy are used to facilitate training. We built a dataset containing 100 CT scans with fractured pelvis and manually annotated the fractures. A five-fold cross-validation experiment shows that our method outperformed max-flow segmentation and network without distance weighting, achieving a global Dice of 99.38%, a local Dice of 93.79%, and an Hausdorff distance of 17.12 mm. We have made our dataset and source code publicly available and expect them to facilitate further pelvic research, especially reduction planning.

Keywords: CT segmentation · Pelvic fracture · Reduction planning

1 Introduction

Pelvic fracture is a severe type of high-energy injury, with a fatality rate greater than 50%, ranking the first among all complex fractures [8, 16]. Surgical planning and reduction tasks are challenged by the complex pelvic structure, as well as the surrounding muscle groups, ligaments, neurovascular and other tissues. Robotic fracture reduction surgery has been studied and put into clinical use in recent

years, and has successfully increased reduction precision and reduced radiation exposure [2]. Accurate segmentation of pelvic fracture is required in both manual and automatic reduction planning, which aim to find the optimal target location to restore the healthy morphology of pelvic bones. Segmenting pelvic fragments from CT is challenging due to the uncertain shape and irregular position of the bone fragments and the complex collision fracture surface. Therefore, surgeons typically annotate the anatomy of pelvic fractures in a semi-automatic way. First, by tuning thresholds and selecting seed points, adaptive thresholding and region-growing methods are used to extract bone regions [1, 11, 15]. Then, the fracture surfaces are manually delineated by outlining the fragments in 3D view or even modifying the masks in a slice-by-slice fashion. Usually, this tedious process can take more than 30 min, especially when the fractured fragments are collided or not completely separated.

Several studies have been proposed to provide more efficient tools for operators. A semi-automatic graph-cut method based on continuous max-flow has been proposed for pelvic fracture segmentation, but it still requires the manual selection of seed points and trail-and-error [4, 22]. Fully automatic max-flow segmentation based on graph cut and boundary enhancement filter is useful when fragments are separated, but it often fails on fragments that are collided or compressed [9, 19]. Learning-based bone segmentation has been successfully applied to various anatomy, including the pelvis, rib, skull, etc. [12, 13]. Some deep learning methods have been proposed to detect fractures [17, 18, 21], but the output from these methods cannot provide a fully automated solution for subsequent operations. In FracNet, rib fracture detection was formulated as a segmentation problem, but with a resultant Dice of 71.5%, it merely outlined the fracture site coarsely without delineating the fracture surface [7]. Learning-based methods that directly deal with fracture segmentation have rarely been studied.

Fracture segmentation is still a challenging task for the learning-based method because (1) compared to the more common organ/tumor segmentation tasks where the model can implicitly learn the shape prior of an object, it is difficult to learn the shape information of a bone fragment due to the large variations in fracture types and shapes [10]; (2) the fracture surface itself can take various forms including large space (fragments isolated and moved), small gap (fragments isolated but not moved), crease (fragments not completely isolated), compression (fragments collided), and their combinations, resulting in quite different image intensity profiles around the fracture site; and (3) the variable number of bone fragments in pelvic fracture makes it difficult to prescribe a consistent labeling strategy that applies to every type and case.

This paper proposes a deep learning-based method to segment pelvic fracture fragments from preoperative CT images automatically. Our major contribution includes three aspects. (1) We proposed a complete automatic pipeline for pelvic fractures segmentation, which is the first learning-based pelvic fracture segmentation method to the best of our knowledge. (2) We designed a novel multi-scale distance-weighted loss and integrated it into the deeply supervised training of the fracture segmentation network to boost accuracy near the

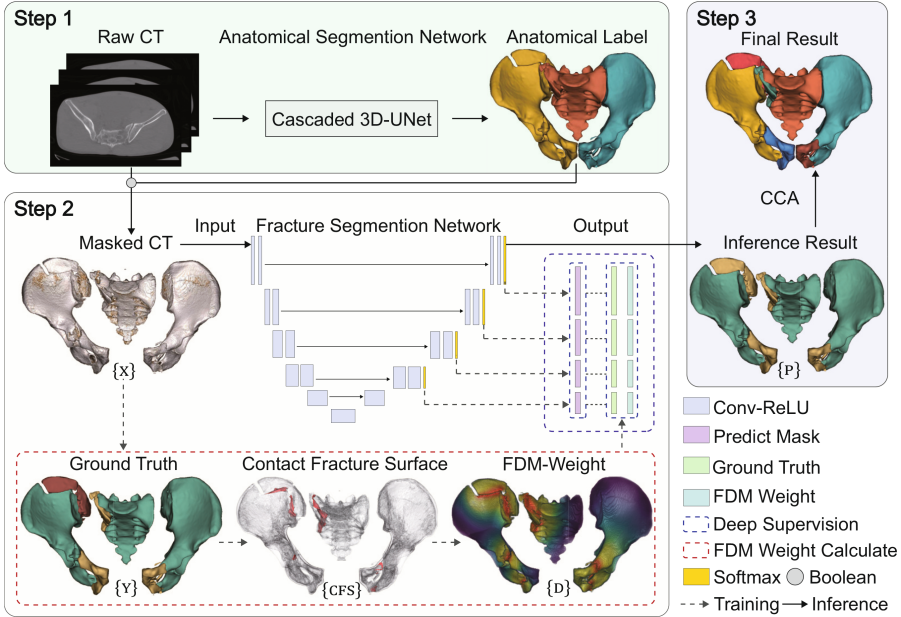


Fig. 1. Overview of the proposed pelvic fracture segmentation framework.

fracture site. (3) We established a comprehensive pelvic fracture CT dataset and provided ground-truth annotations. Our dataset and source code are publicly available at <https://github.com/YzzLiu/FracSegNet>. We expect them to facilitate further pelvis-related research, including but not limited to fracture identification, segmentation, and subsequent reduction planning.

2 Methods

2.1 The Overall Segmentation Framework

Our study aims to automatically segment the major and minor fragments of target bones (left and right ilia and sacrum) from CT scans. As illustrated in Fig. 1, our method consists of three steps. In the first step, an anatomical segmentation network is used to extract the pelvic bones from the CT scan. With a cascaded 3D nn-UNet architecture, the network is pre-trained on a set of healthy pelvic CT images [5, 13] and further refined on our fractured dataset. In the second step, a fracture segmentation network is used to separate the bone fragments from each iliac and sacral region extracted from the first step. To define a consistent labeling rule that is applicable to all fracture types, we prescribe three labels for each bone, namely the background, the main fragment, and other fragments. The main fragment is the largest fragment at the center. In the third step, isolated components are further separated and labeled, and small isolated bone fragments are removed to form the final output.

2.2 Fracture Segmentation Network

The contact fracture surface (CFS) is the part where the bones collide and overlap due to compression and impact, and is the most challenging part for human operators to draw. We are particularly concerned about the segmentation performance in this region. Therefore, we introduce guidance into the network training using fracture distance map (FDM). A 3D UNet is selected as the base model [6]. The model learns a non-linear mapping relationship $M : X \rightarrow Y$, where X and Y are the input and ground truth of a training sample, respectively.

Fracture Distance Map. The FDM is computed on the ground-truth segmentation of each data sample before training. This representation provides information about the boundary, shape, and position of the object to be segmented. First, CFS regions are identified by comparing the labels within each voxel's neighbourhood. Then, we calculate the distance of each foreground voxel to the nearest CFS as its distance value D_v , and divide it by the maximum.

$$D_v = \mathbf{I}(Y_v \geq 1) \min_{u \in CFS} \|v - u\|_2, \quad (1)$$

$$\hat{D}_v = \frac{D_v}{\max_{v \in V} D_v}, \quad (2)$$

where $v = (h, w, d) \in V$ is the voxel index, Y is the ground-truth segmentation, $\mathbf{I}(Y_v \geq 1)$ is the indicator function for foreground, and \hat{D} is the normalized distance. The distance is then used to calculate the FDM weight \hat{W} using the following formula:

$$W_v = \lambda_{back} + \mathbf{I}(Y_v \geq 1) \frac{1 - \lambda_{back}}{1 + e^{\lambda_{FDM} \hat{D}_v - 5}}, \quad (3)$$

$$\hat{W}_v = \frac{W_v \cdot |V|}{\sum_{v \in V} W_v}. \quad (4)$$

To ensure the equivalence of the loss among different samples, the weights are normalized so that the average is always 1.

FDM-Weighted Loss. The FDM weight \hat{W} is then used to calculate the weighted Dice and cross-entropy losses to emphasize the performance near CFS by assigning larger weights to those pixels.

$$\mathcal{L}_{dice} = 1 - \frac{2}{|L|} \sum_{l \in L} \frac{\sum_{v \in V} \hat{W}_v P_v^l Y_v^l}{\hat{W}_v P_v^l + \sum_{v \in V} \hat{W}_v Y_v^l}, \quad (5)$$

$$\mathcal{L}_{ce} = -\frac{1}{|V||L|} \sum_{v \in V} \sum_{l \in L} \hat{W}_v Y_v^l \log(P_v^l), \quad (6)$$

where L is the number of labels, P_v^l, Y_v^l are the network output prediction and the one-hot encoding form of the ground truth for the l^{th} label of the v^{th} voxel. The overall loss is their weighted sum:

$$\mathcal{L}_{total} = \mathcal{L}_{dice} + \lambda_{ce}\mathcal{L}_{ce}, \quad (7)$$

where λ_{ce} is a balancing weight.

Multi-scale Deep Supervision. We use a multi-scale deep supervision strategy in model training to learn different features more effectively [20]. The deep layers mainly capture the global features with shape/structural information, whereas the shallow layers focus more on local features that help delineate fracture surfaces. We add auxiliary losses to the decoder at different resolution levels (except the lowest resolution level). The losses are calculated using the corresponding down-sampled FDM \hat{W}_v^n , and down-sampled ground truth Y_v^n . We calculate the output of the n^{th} level \mathcal{L}^n by changing λ_{FDM} in Eq. (3). The λ_{FDM} of each layer decreases by a factor of 2 as the depth increases, i.e., $\lambda^{n+1} = \lambda^n/2$. In this way, the local CFS information are assigned more attention in the shallow layers, while the weights become more uniform in the deep layers.

Smooth Transition. To stabilize network training, we use a smooth transition strategy to maintain the model’s attention on global features at the early stage of training and gradually shift the attention towards the fracture site as the model evolves [14]. The smooth transition dynamically adjusts the proportion of the FDM in the overall weight matrix based on the number of training iterations. The dynamic weight is calculated using the following formula:

$$W_{st} = \mathbf{I}(t < \tau) \left(\frac{1}{1+\delta} \mathcal{J} + \frac{\delta}{1+\delta} \hat{W} \right) + \mathbf{I}(t \geq \tau) \hat{W}, \quad (8)$$

where $\delta = -\ln(1 - \frac{t}{\tau})$, \mathcal{J} is an all-ones matrix with the same size as the input volume, t is the current iteration number, and τ is a hyper-parameter. The dynamic weight W_{st} is adjusted by controlling the relative proportion of \mathcal{J} and \hat{W} . The transition terminates when the epoch reaches τ .

2.3 Post-processing

Connected component analysis (CCA) has been widely used in segmentation [3], but is usually unsuitable for fracture segmentation because of the collision between fragments. However, after identifying and removing the main central fragment from the previous step, other fragments are naturally separated. Therefore, in the post-processing step, we further isolate the remaining other fragments by CCA. The isolated components are then assigned different labels. In addition, we remove fragments smaller than a certain threshold, which has no significant impact on planning and robotic surgery.

2.4 Data and Annotation

Although large-scale datasets on pelvic segmentation have been studied in some research [13], to the best of our knowledge, currently there is no well-annotated fractured pelvic dataset publicly available. Therefore, we curated a dataset of 100 preoperative CT scans covering all common types of pelvic fractures. These data is collected from 100 patients (aged 18–74 years, 41 females) who were to undergo pelvic reduction surgery at Beijing Jishuitan Hospital between 2018 and 2022, under IRB approval (202009-04). The CT scans were acquired on a Toshiba Aquilion scanner. The average voxel spacing is $0.82 \times 0.82 \times 0.94 \text{ mm}^3$. The average image shape is $480 \times 397 \times 310$.

To generate ground-truth labels for bone fragments, a pre-trained segmentation network was used to create initial segmentations for the ilium and sacrum [13]. Then, these labels were further modified and annotated by two annotators and checked by a senior expert.

3 Experiments and Results

3.1 Implementation

We compared the proposed method (FDMSS-UNet) against the network without smooth transition and deep supervision (FDM-UNet) and the network without distance weighting (UNet) in an ablation study. In addition, we also compared the traditional max-flow segmentation method. In the five-fold cross-validation, each model was trained for 2000 epochs per fold. The network input was augmented eight times by mirror flip. The learning rate in ADAM optimizer was set to 0.0001. λ_{back} was set to 0.2. λ_{ce} was set to 1. The initial λ_{FDM} was set to 16. The termination number for smooth transition τ was set to 500 epochs.

The models were implemented in PyTorch 1.12. The experiments were performed on an Intel Xeon CPU with 40 cores, a 256 GB memory, and a Quadro RTX 5000 GPU. The comprehensive code and pertinent details are provided at <https://github.com/YzzLiu/FracSegNet>.

3.2 Evaluation

We calculate the Dice similarity coefficient (DSC) and the 95th percentile of Hausdorff Distance (HD) to evaluation the performance. In addition, we evaluated the local Dice (LDSC) within the 10 mm range near the CFS to assess the performance in the critical areas. We reported the performance on iliac main fragment (I-main), iliac other fragment(s) (I-other), sacral main fragment (S-main), sacral other fragment(s) (S-other), and all together.

3.3 Results

Figure 2 shows a qualitative comparison among different methods. Max-flow was able to generate reasonable segmentation only when the CFS is clear and mostly

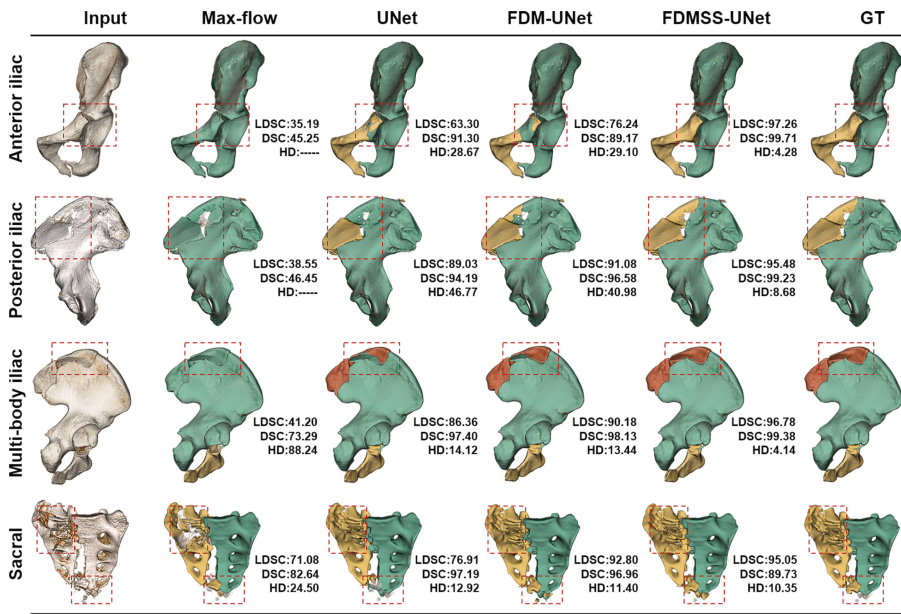


Fig. 2. Pelvic fracture segmentation results from different methods.

non-contact. UNet correctly identified the fracture fragments, but was often confused by the complicated CFS regions, resulting in imprecise fracture lines. The introduction of FDM weighting and deep supervision with smooth transition successfully improved the performance near the CFS, and achieved the overall best result.

Table 1 shows the quantitative results. The results on the main fragments were better than other fragments due to the larger proportion. The deep learning methods had much higher success rates in identifying the fragments, resulting in significantly better results in minor fragments than max-flow, with $p < 0.05$ in paired t-test. Introducing the FDM significantly improved the prediction accuracy in the CFS region ($p < 0.05$). Although FDM-UNet achieved the best LDSC results in several parts, it compromised the global performance of DSC and HD significantly, compared to FDMSS-UNet. The deep supervision and smooth transition strategies stabilized the training, and achieved the overall best results, with balanced local and global performance.

The average inference time for the fracture segmentation network was 12s. The overall running time for processing a pelvic CT was 0.5 to 2 min, depending on the image size and the number of fractured bones.

Table 1. Quantitative segmentation results using different methods. (L)DSC is reported in percentage value, and HD is in mm. Best values are shown in bold. Statistically significant differences compared to FDMSS-UNet are indicated by * ($p < 0.05$).

		I-main	I-other	S-main	S-other	All
Max-flow	DSC	94.52 \pm 0.56*	24.67 \pm 3.72*	89.83 \pm 1.89*	23.00 \pm 3.18*	69.21 \pm 4.01*
	LDSC	74.53 \pm 0.99*	24.24 \pm 3.26*	62.06 \pm 1.91*	27.31 \pm 3.40*	54.11 \pm 3.20*
	HD95	68.56*	77.45*	45.16*	90.76*	73.95*
UNet	DSC	99.38 \pm 0.13*	92.88 \pm 1.30*	98.66 \pm 0.33*	87.91 \pm 1.55*	96.60 \pm 0.89*
	LDSC	90.68 \pm 1.22*	86.58 \pm 1.69*	84.44 \pm 1.77*	78.50 \pm 2.50*	86.86 \pm 1.70*
	HD95	43.42*	50.21*	27.14*	76.69*	46.11*
FDM-UNet	DSC	99.37 \pm 0.13*	96.15 \pm 0.79*	98.73 \pm 0.27*	92.20 \pm 0.74*	97.88 \pm 0.51*
	LDSC	94.43 \pm 0.34	93.94 \pm 0.36	92.88 \pm 0.39*	91.80 \pm 0.51	93.81 \pm 0.38
	HD95	34.75*	52.03*	28.99*	65.26*	45.79*
FDMSS-UNet	DSC	99.78 \pm 0.03	99.05 \pm 0.10	99.57 \pm 0.04	97.87 \pm 0.17	99.38 \pm 0.09
	LDSC	94.79 \pm 0.51	93.75 \pm 0.56	91.02 \pm 1.62	93.65 \pm 0.70	93.79 \pm 0.78
	HD95	13.24	19.43	14.85	19.89	17.12

4 Discussion and Conclusion

We have introduced a pelvic fracture CT segmentation method based on deep convolutional networks. A fracture segmentation network was trained with a distance-weighted loss and multi-scale deep supervision to improve fracture surface delineation. We evaluated our method on 100 pelvic fracture CT scans and made our dataset and ground truth publicly available. The experiments demonstrated the method’s effectiveness on various types of pelvic fractures. The FDM weighted loss, along with multi-scale deep supervision and smooth transition, improved the segmentation performance significantly, especially in the areas near fracture lines. Our method provides a convenient tool for pelvis-related research and clinical applications, and has the potential to support subsequent automatic fracture reduction planning.

One obstacle for deep learning-based fracture segmentation is the variable number of bone fragments in different cases. The ultimate goal of this study is to perform automatic fracture reduction planning for robotic surgery, where the main bone fragment is held and moved to the planned location by a robotic arm, whereas minor fragments are either moved by the surgeons’ hands or simply ignored. In such a scenario, we found isolating the minor fragments usually unnecessary. Therefore, to define a consistent labeling strategy in annotation, we restrict the number of fragments of each bone to three. This rule of labelling applies to all 100 cases we encountered. Minor fragments within each label can be further isolated by CCA or handcrafting when needed by other tasks.

We utilize a multi-scale distance-weighted loss to guide the network to learn features near the fracture site more effectively, boosting the local accuracy without compromising the overall performance. In semi-automatic pipelines where human operators are allowed to modify and refine the network predictions, an

accurate initial segmentation near the fracture site is highly desirable because the fracture surface itself is much more complicated, often intertwined and hard to draw by manual operations. Therefore, with the emphasis on fracture surface, even when the prediction from the network is inaccurate, manual operations on 3D view can suffice for most modifications, eliminating the need for the inefficient slice-by-slice handcrafting. In future studies, we plan to integrate the proposed method into an interactive segmentation and reduction planning software and evaluate the overall performance.

Acknowledgements. This work was supported by the Beijing Science and Technology Project (Grants No. Z221100003522007 and Z201100005420033).

References

1. Fornaro, J., Székely, G., Harders, M.: Semi-automatic segmentation of fractured pelvic bones for surgical planning. In: Bello, F., Cotin, S. (eds.) ISBMS 2010. LNCS, vol. 5958, pp. 82–89. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-11615-5_9
2. Ge, Y., Zhao, C., Wang, Y., Wu, X.: Robot-assisted autonomous reduction of a displaced pelvic fracture: a case report and brief literature review. *J. Clin. Med.* **11**(6), 1598 (2022)
3. Ghimire, K., Chen, Q., Feng, X.: Head and neck tumor segmentation with deeply-supervised 3D UNet and progression-free survival prediction with linear model. In: Andrearczyk, V., Oreiller, V., Hatt, M., Depeursinge, A. (eds.) HECKTOR 2021. LNCS, vol. 13209, pp. 141–149. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98253-9_13
4. Han, R., et al.: Fracture reduction planning and guidance in orthopaedic trauma surgery via multi-body image registration. *Med. Image Anal.* **68**, 101917 (2021)
5. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
6. Isensee, F., Jäger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: Automated design of deep learning methods for biomedical image segmentation. *arXiv preprint arXiv:1904.08128* (2019)
7. Jin, L., et al.: Deep-learning-assisted detection and segmentation of rib fractures from CT scans: development and validation of FracNet. *EBioMedicine* **62**, 103106 (2020)
8. Kowal, J., Langlotz, F., Nolte, L.P.: Basics of computer-assisted orthopaedic surgery. In: Stiehl, J.B., Konermann, W.H., Haaker, R.G., DiGioia, A.M. (eds.) Navigation and MIS in Orthopedic Surgery, pp. 2–8. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-36691-1_1
9. Krčah, M., Székely, G., Blanc, R.: Fully automatic and fast segmentation of the femur bone from 3D-CT images with no shape prior. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 2087–2090. IEEE (2011)
10. Kuiper, R.J., et al.: Efficient cascaded V-Net optimization for lower extremity CT segmentation validated using bone morphology assessment. *J. Orthop. Res.* **40**(12), 2894–2907 (2022)

11. Lai, J.Y., Essomba, T., Lee, P.Y.: Algorithm for segmentation and reduction of fractured bones in computer-aided preoperative surgery. In: Proceedings of the 3rd International Conference on Biomedical and Bioinformatics Engineering, pp. 12–18 (2016)
12. Liu, J., Xing, F., Shaikh, A., Linguraru, M.G., Porras, A.R.: Learning with context encoding for single-stage cranial bone labeling and landmark localization. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13438, pp. 286–296. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16452-1_28
13. Liu, P., et al.: Deep learning to segment pelvic bones: large-scale CT datasets and baseline models. *Int. J. Comput. Assist. Radiol. Surg.* **16**, 749–756 (2021)
14. Qamar, S., Jin, H., Zheng, R., Ahmad, P., Usama, M.: A variant form of 3D-UNet for infant brain segmentation. *Future Gener. Comput. Syst.* **108**, 613–623 (2020)
15. Ruikar, D.D., Santosh, K., Hegadi, R.S.: Automated fractured bone segmentation and labeling from CT images. *J. Med. Syst.* **43**, 1–13 (2019)
16. Sugano, N.: Computer-assisted orthopaedic surgery and robotic surgery in total hip arthroplasty. *Clin. Orthop. Surg.* **5**(1), 1–9 (2013)
17. Tomita, N., Cheung, Y.Y., Hassanpour, S.: Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. *Comput. Biol. Med.* **98**, 8–15 (2018)
18. Ukai, K., et al.: Detecting pelvic fracture on 3D-CT using deep convolutional neural networks with multi-orientated slab images. *Sci. Rep.* **11**(1), 1–11 (2021)
19. Wang, D., Yu, K., Feng, C., Zhao, D., Min, X., Li, W.: Graph cuts and shape constraint based automatic femoral head segmentation in CT images. In: Proceedings of the Third International Symposium on Image Computing and Digital Medicine, pp. 1–6 (2019)
20. Wang, J., Zhang, X., Guo, L., Shi, C., Tamura, S.: Multi-scale attention and deep supervision-based 3D UNet for automatic liver segmentation from CT. *Math. Biosci. Eng. MBE* **20**(1), 1297–1316 (2023)
21. Yamamoto, N., et al.: An automated fracture detection from pelvic CT images with 3-D convolutional neural networks. In: 2020 International Symposium on Community-Centric Systems (CcS), pp. 1–6. IEEE (2020)
22. Yuan, J., Bae, E., Tai, X.C., Boykov, Y.: A spatially continuous max-flow and min-cut framework for binary labeling problems. *Numer. Math.* **126**, 559–587 (2014)