



Research article

Automatic segmentation of knee CT images of tibial plateau fractures based on three-dimensional U-Net: Assisting junior physicians with Schatzker classification

Die Cai^{a,1}, Yu Zhou^{b,1}, Wenjie He^a, Jichun Yuan^a, Chenyuan Liu^c, Rui Li^a, Yi Wang^{b,2,*}, Jun Xia^{a,2,**}

^a Department of Radiology, The First Affiliated Hospital of Shenzhen University, Shenzhen University, Shenzhen Second People's Hospital, 3002 SunGang Road West, Shenzhen 518035, Guangdong Province, China

^b Smart Medical Imaging, Learning and Engineering (SMILE) Lab, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen 518060, China

^c Five-year Clinical Medicine, Xiangya School of Medicine, Central South University, Changsha 410083, Hunan, China

ARTICLE INFO

Keywords:

Artificial Intelligence
Deep learning
Image segmentation
Knee
Tibial plateau fracture
Computed tomography

ABSTRACT

Purpose: This study aimed to automatically segment knee computed tomography (CT) images of tibial plateau fractures using a three-dimensional (3D) U-net-based method, accurately construct 3D maps of tibial plateau fractures, and examine their usefulness for Schatzker classification in clinical practice.

Methods: We retrospectively enrolled 234 cases with tibial plateau fractures from our hospital in this study. The four constituent bones of the knee were manually annotated using ITK-SNAP software. Finally, image features were extracted using deep learning. The usefulness of the results for Schatzker classification was examined by an orthopaedic and a radiology resident.

Results: On average, our model required < 40 s to process a 3D CT scan of the knee. The average Dice coefficient for all four knee bones was higher than 0.950, and highly accurate 3D maps of the tibia were produced. With the aid of the results of our model, the accuracy, sensitivity, and specificity of the Schatzker classification of both residents improved.

Conclusions: The proposed method can rapidly and accurately segment knee CT images of tibial plateau fractures and assist residents with Schatzker classification, which can help improve diagnostic efficiency and reduce the workload of junior doctors in clinical practice.

1. Introduction

Tibial plateau fractures (TPF) are a common type of fracture that occurs near the knee at the proximal end of the tibia [1,2]. They account for approximately 1.7 % of all human fractures [3,4] and 18.6 % of tibial fractures [5]. Within 10 years of TPF, 7 % of individuals undergo total knee replacement [6]. TPF can also lead to other complications, including osteofascial compartment syndrome, peroneal nerve injury,

pain, axial dislocation, knee instability, and post-traumatic osteoarthritis [7–11]. Delayed diagnosis and unreliable preoperative planning increase the risk of these complications, with incorrect Schatzker classification often contributing to these problems.

The Schatzker classification [12] is used for clinical assessment, management planning, and prognosis prediction in TPF patients [13]. However, misclassification of the Schatzker classification can cause misleading communication, inaccurate prognostic assessments, unclear

Abbreviations: TPF, tibial plateau fractures; CT, computed tomography; 3D, three-dimensional; AI, artificial intelligence; MRI, magnetic resonance imaging; ASD, mean surface distance; HD, Hausdorff distance; SD, surface distance.

* Corresponding author at: Smart Medical Imaging, Learning and Engineering (SMILE) Lab, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen, 518060, China.

** Corresponding author at: Shenzhen Second People's Hospital, 3002 SunGang Road West, Shenzhen 518035, Guangdong Province, China

E-mail addresses: onewang@szu.edu.cn (Y. Wang), xiajun@email.szu.edu.cn (J. Xia).

¹ These two authors have contributed to the work equally and should be regarded as co-first authors.

² These two authors have contributed to the work equally and should be regarded as co-corresponding author.

<https://doi.org/10.1016/j.ejrad.2024.111605>

Received 11 September 2023; Received in revised form 19 May 2024; Accepted 5 July 2024

Available online 7 July 2024

0720-048X/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

research findings, and incorrect treatment [14]. The original Schatzker classification relies on X-ray images of the knee [12], which can result in misdiagnosis, especially among inexperienced physicians. Accurate assessment of fracture type and depth of collapse requires the use of computed tomography (CT) imaging and three-dimensional (3D) reconstruction [11,15–18]. CT imaging has higher accuracy than standard X-ray imaging and can improve surgical planning [13,18–20]. Therefore, there is an urgent need for a method that allows for rapid segmentation and accurate reconstruction of knee CT images of TPF, aiding junior doctors with Schatzker classification.

Deep learning, a subset of artificial intelligence (AI), has been widely used in medical image analysis [21–26], including automatic segmentation of knee X-ray [27], CT [28], and magnetic resonance imaging (MRI) [29]. However, existing studies on automatic knee segmentation are based on normal tibial plateau images, and to the best of our knowledge, no studies have applied deep learning to the automatic segmentation and accurate reconstruction of knee CT images of TPF. In this study, we proposed an automatic method based on 3D U-net, along with a specifically designed training strategy that does not involve any additional elaborate modules, image preprocessing, or post-processing; the knee segmentation workflow is illustrated in Fig. 1. We also examined the usefulness of the automatic segmentation results in assisting junior doctors with correctly applying the Schatzker classification.

2. Material and methods

2.1. Data selection

We retrospectively collected CT image data from 200 patients for model training and internal validation in our hospital from January 2020 to March 2023. An additional 34 patients were later collected for external validation. All patients had to meet the following inclusion criteria: (1) age between 18–80 years, regardless of sex; (2) skeletal maturity and absence of severe deformities; (3) previous spiral CT examinations including the knee and tibial plateau fracture with a maximum scan layer thickness of 1.5 mm. Patients with pre-existing abnormal bone morphology due to conditions such as concomitant bone tumours or old fractures were excluded. The CT scanning parameters were 120 kV, 200 mAs, 512×512 matrix, and 1.5 mm reconstructed slice thickness.

2.2. Data annotation and Schatzker classification

The ground truth used in the training was manually obtained by a chief radiologist and a chief orthopaedic surgeon through ITK-SNAP software (version 3.8.0). To define a gold standard, a panel of two experienced orthopaedic surgeons and a senior radiologist classified 74 knee CT images (40 and 34 cases from the internal and external validation sets, respectively) based on the Schatzker classification until a consensus was reached. All of them had extensive experience in the field of knee injury and 10 to 25 years of experience in clinical practice. One radiology resident and one orthopaedic resident representing junior physicians were enlisted to assess the practicality of the model's results for Schatzker classification. Both physicians were in their second year of general residency training. Prior to the assessment, the residents received no specific research-oriented pre-training and were kept blind to the gold standard results. During the initial evaluation, the radiology resident solely reviewed all images of the original CT images, while the orthopaedic resident first examined all images of the original CT images and then examined all images of the model-generated images. In the second evaluation, the images from the first assessment were switched between the two physicians (Fig. 2). The assessments occurred with at least 1 month between them to prevent learning bias.

2.3. Pre-processing

In recent years, several studies have demonstrated the performance of U-shaped structures in medical image segmentation. Adding excessively complicated modules to the network may lead to overfitting to a specific dataset [30]. Consequently, we aimed to determine the optimal pre-processing method for our datasets. First, we calculated the spacing of each volume and selected the median value as the target spacing. Next, all the volumes were resampled to the spacing of $0.3 \times 0.3 \times 0.7$ mm, where 0.7 mm was the z-axis spacing. Second, we coarsely set the voxels with CT values below -200 HU or above 800 HU to 0 HU, followed by windowing, centring operations, and bilateral filtering to normalize and enhance the volumes. Gradually, we resized the enhanced and unified spacing volumes to 66.7% , and the spacing was set to 1.5 because 3D convolution operations consume a large amount of computing resources and we trained our 3D U-net in patch form, which means that the input of the network is a fixed patch size. The resizing operation aimed to obtain a larger view of the input volume patch.

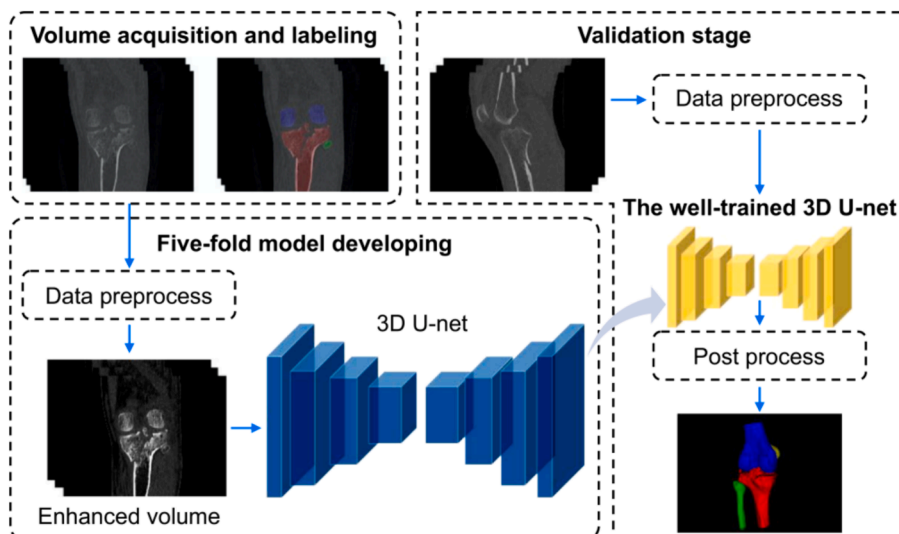


Fig. 1. Flowchart of knee segmentation.

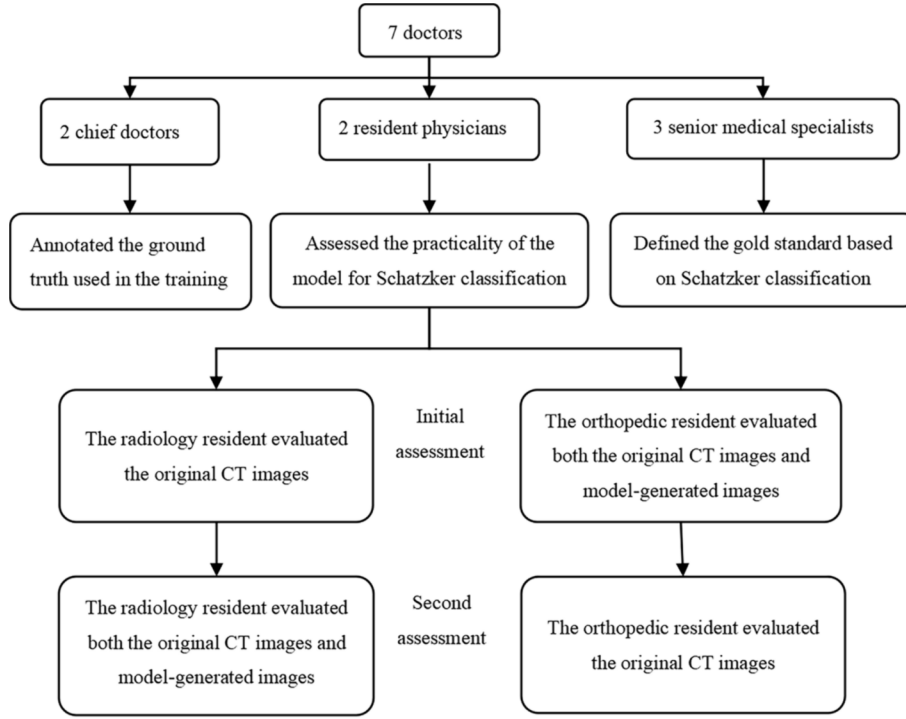


Fig. 2. Flowchart of the study design.

2.4. Proposed deep learning network

The deep learning segmentation network was based on the 3D U-net proposed by Çiçek et al. [31], which is an extension of 2D U-net [32] (usually named as U-net) that retains the conventional “encoder-decoder” structure [32] but expands the convolution operations to three dimensions, requiring three-dimensional image data network as input. Details of our segmentation network, comprising four encoder blocks and three decoder blocks in total, are shown in Fig. 3. Each encoder block comprised two continuous $3 \times 3 \times 3$ convolutions and $2 \times 2 \times 2$ maximum pooling with strides of two in each dimension. Each convolution was followed by batch normalization and a rectified linear unit. Conversely, the decoder block replaced the maximum pooling with a $3 \times 3 \times 3$ transposed convolution with a stride of two and a padding of one in each dimension. Through skip connection, the feature maps extracted from the encoders can be fused with the same-resolution feature maps

generated by the decoders, enabling the network to learn semantic and detailed information. Finally, a $1 \times 1 \times 1$ convolution was applied to obtain the segmented predictions.

2.5. Training strategy and loss function

Maximum number of training epochs was set to 200. Adam optimizer was selected and the initial learning rate was 0.01, while beta1 and beta2 were 0.5 and 0.999, respectively. Batch size was set to 4, and fixed patch size was set to $96 \times 128 \times 128$ in the z, y, and x orders owing to GPU memory limitations. Learning rate changed according to the cosine-annealing learning scheduler. Loss function in our study was the cross-entropy loss, defined as follows:

$$\text{CrossEntropy Loss} = -\frac{1}{N} \sum_i^N \sum_{c=1}^C y_{ic} \log(p_{ic})$$

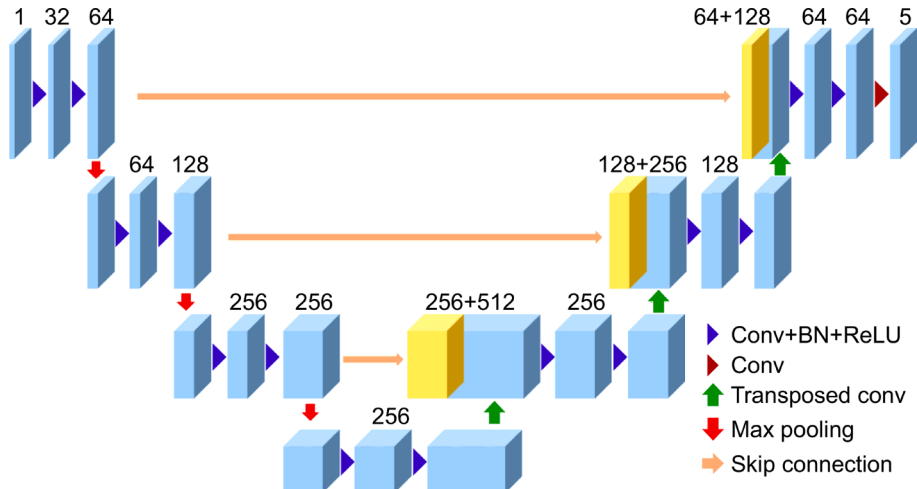


Fig. 3. Details of the segmentation network.

where C is the number of foreground classes, N is the number of voxels, p_{ic} is the probability of the i^{th} voxel belonging to class c , and y_{ic} is 1 if the i^{th} voxel belongs to class c else 0.

Four different classes of bone were labelled 1–4; however, the total voxel number of each class was not equivalent. The number of voxels classified as the tibia and femur were usually more than those classified as the fibula and patella. Cross-entropy loss cannot resolve the class imbalance problem; therefore, we designed a weighted sampling rule: for the entire volume, an increasing proportion of the class voxels lowers the probability of being chosen as the centre of a volume patch. The weights of the tibia, fibula, femur, patella, and background were set to 15, 30, 15, 30, and 10 %, respectively, according to our previous experiments.

2.6. Post-processing

To accommodate the patch-based training of the segmentation network, the validated volume was partitioned into multiple patch volumes, segmented individually, and patched sequentially. Considering that the voxels at the edge of the patch usually have low confidence, we used a Gaussian distribution rule to decrease their weight and increase the centre voxels of the patch. Finally, the segmented predictions of the validation set underwent 10 rounds of hole-filling and were resampled to the original CT volume spacing for better evaluation against the ground truth.

2.7. Evaluation metrics

Dice and Jaccard coefficients are used to evaluate segmentation results of a model. Larger values indicate higher sample similarity. Mean surface distance (ASD) and Hausdorff distance (HD) are distance-related parameters used to evaluate the segmentation effect of medical images and can better reflect the segmentation effect of image edges. Smaller values indicate better neural network segmentation effects. Accuracy, sensitivity, and specificity were used to assess the residents' classification diagnostic performance before and after the model-generated images assistance. Additionally, we used Cohen's kappa to determine the agreement between the physicians and the gold standard. To interpret k values, Landis and Koch [33] proposed the following division: a k value between 0.01 and 0.20 reflects slight agreement, between 0.21 and 0.40 reflects fair agreement, between 0.41 and 0.60 reflects moderate agreement, between 0.61 and 0.80 reflects substantial agreement, and greater than 0.81 reflects almost-perfect agreement.

2.8. Statistical analysis

All calculations were performed using SPSS version 26.0. Values with normal distribution (Shapiro–Wilk test) were recorded as mean \pm standard deviation. Paired t -test was used to compare the diagnostic times of the two residents before and after model assistance. Statistical significance was set at $P < 0.05$.

3. Results

3.1. Segmentation in internal and external validation sets

We trained the model with the highest mean Dice score using five-fold cross-validation. The final knee segmentation results obtained in the internal and external validation sets are shown in Table 1. All four bones of the knee, including the fractured tibia, yielded satisfactory results, with mean Dice coefficients > 0.950 . Femoral segmentation errors were minimal, with a Dice coefficient of 0.995 ± 0.001 in the internal validation set. Time required to automatically segment each knee CT image was less than 40 s, significantly shorter than the manual segmentation time of 2 h. Fig. 4 shows the 3D models of six objects with

Table 1
Results of automatic segmentation using 3D U-Net on the four component bones of the knee [mean \pm standard deviation].

| Objective | Validation set | Dice | Jaccard | ASD | 95HD |
|-------------------------|----------------|----------------------|----------------------|----------------------|-----------------------|
| Tibial plateau fracture | internal | 0.986 ± 0.011 | 0.963 \pm 0.026 | 0.239 \pm 0.183 | 2.175 \pm 1.857 |
| | external | 0.958 \pm 0.027 | 0.920 \pm 0.049 | 0.477 \pm 0.286 | 3.833 \pm 2.602 |
| Fibula | internal | 0.978 \pm 0.011 | 0.943 \pm 0.034 | 0.142 \pm 0.122 | 1.067 \pm 1.342 |
| | external | 0.950 \pm 0.037 | 0.907 \pm 0.062 | 0.200 \pm 0.194 | 1.003 \pm 1.580 |
| Femur | internal | 0.995 \pm 0.001 | 0.984 \pm 0.009 | 0.074 \pm 0.053 | 0.521 \pm 0.611 |
| | external | 0.978 \pm 0.032 | 0.958 \pm 0.053 | 0.256 \pm 0.414 | 0.934 \pm 1.707 |
| Patella | internal | 0.987 \pm 0.006 | 0.960 \pm 0.015 | 0.175 \pm 0.435 | 2.329 \pm 11.339 |
| | external | 0.966 \pm 0.018 | 0.934 \pm 0.032 | 0.144 \pm 0.094 | 0.633 \pm 0.674 |

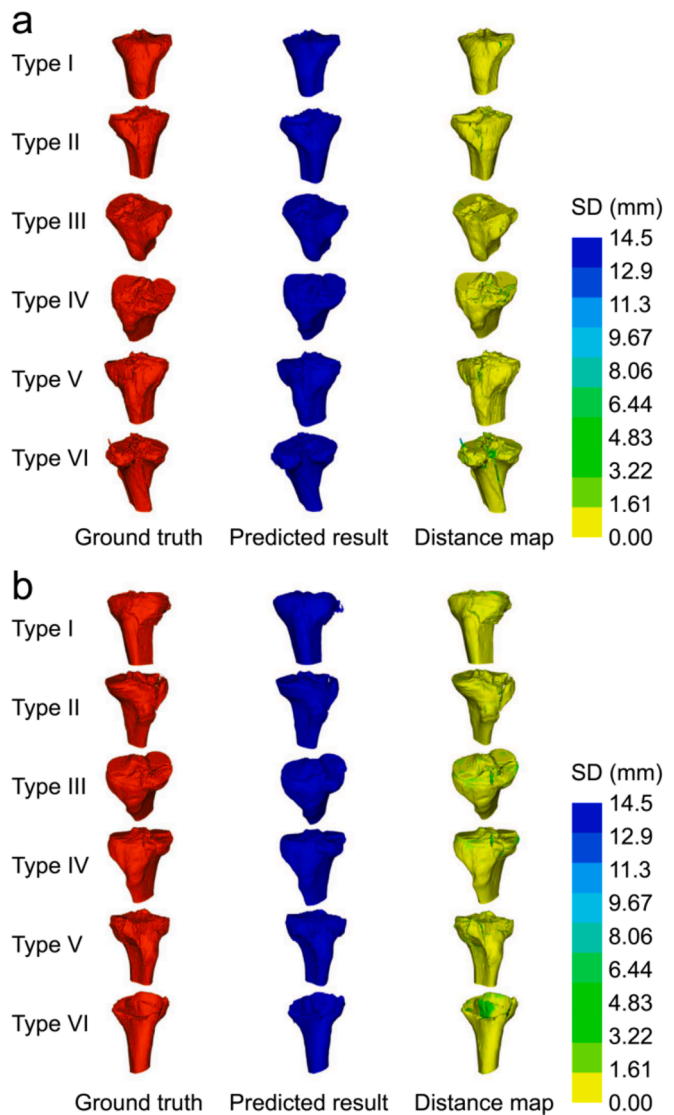


Fig. 4. 3D visualization of the tibia with tibial plateau fracture and the surface distance (SD) map between the ground truth and the automatic segmentation. a: internal validation set; b: external validation set.

the highest Dice coefficient from each of the six Schatzker types. Surface distance mapping between model segmentation results and ground truth was plotted. For each surface voxel in the segmentation results, the Euclidean distance to the nearest contour voxel in the ground truth was calculated as the surface distance (SD) error. Fig. 4 visualizes the participants' SD errors in colour.

3.2. Comparison with Graph-Cut

Graph-Cut [34] is a traditional image segmentation method. We compared the segmentation results of Graph-Cut with those of our model. With 74 validation set examples, the Graph-Cut image segmentation results showed 35 cases where inference failed (the four bones of the knee could not be successfully segmented automatically) and 11 cases where the fibula was not successfully segmented, but the femur, patella, and fractured tibia were segmented automatically. We selected the CT image of the knee with the highest Dice coefficient of the tibial plateau fracture in the Graph-Cut segmentation results and selected four representative axial images from this CT image for comparison with the corresponding segmentation results of our model. The segmentation results of our model were very close to the ground truth (Fig. 5).

3.3. AI-assisted Schatzker classification task

The diagnostic efficacies of the radiology and orthopaedic residents before and after the model images assistance in terms of accuracy, sensitivity, and specificity are presented in Table 2. In all six subtypes,

results of these assessment metrics after the model images assistance were greater than or equal to the pre-assisted results, and the diagnosis time was shorter for both physicians (Table 3). After the model results assistance, >0.912 accuracy and >0.889 specificity were achieved for all subtypes, and based on the sensitivity metric, three important observations were made. First, type IV was the best-performing subtype, with a sensitivity of 1.000 for both physicians' diagnoses in both datasets. Second, type III in the external validation set had a sensitivity of only 0.500 for both physicians' diagnoses, most likely because of its very low percentage [5.9 % (2/34)]. Third, for type VI in the external validation set, the difference in diagnostic opinions between the two physicians was larger than that for the other subtypes, which showed a different level of diagnosis between the two physicians. Both physicians showed significant improvement in most kappa values after model assistance (Table 4 and 5), especially for the orthopaedic resident in external validation set type V ($k_1: 0.675, k_2: 1.000$). However, with such excellent diagnostic results after model assistance, one type III in the external validation set was an exception, correctly typed by the orthopaedic resident in the first round, but incorrectly typed with model assistance.

4. Discussion

We proposed a method based on a 3D U-net for automatic segmentation of knee CT images of TPF. To the best of our knowledge, no studies have applied deep learning for automatic segmentation and accurate reconstruction of knee CT images of TPF. Our model has a very high

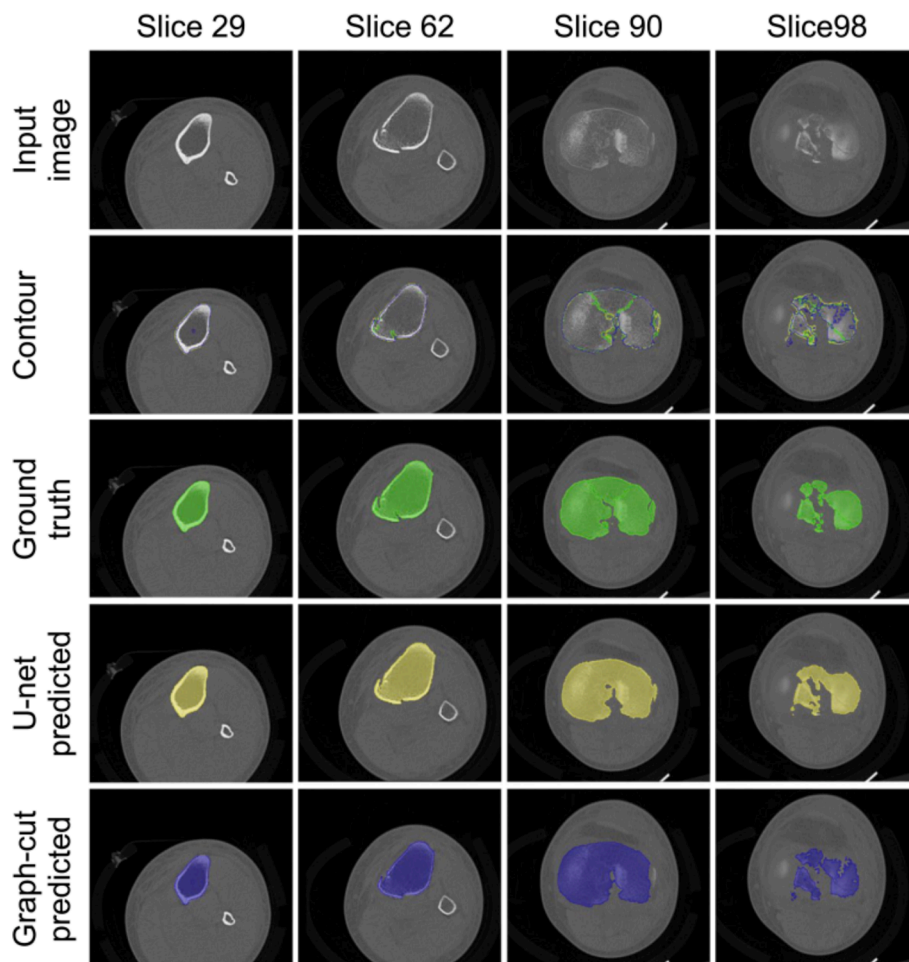


Fig. 5. Tibial contour segmentation. The green, yellow, and blue contours represent the results segmented with the ground truth, 3D U-net, and Graph-Cut, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Performance of a radiology and an orthopaedic resident on Schatzker classification tasks.

| | | Internal validation set | | | External validation set | | |
|-----|-----|-------------------------|-------------|-------------|-------------------------|-------------|-------------|
| | | Accuracy | Sensitivity | Specificity | Accuracy | Sensitivity | Specificity |
| I | RR1 | 0.925 | 0.625 | 1.000 | 0.941 | 0.750 | 0.967 |
| | RR2 | 1.000 | 1.000 | 1.000 | 0.971 | 1.000 | 0.967 |
| | OR1 | 0.975 | 0.875 | 1.000 | 0.971 | 0.750 | 1.000 |
| | OR2 | 0.975 | 0.875 | 1.000 | 1.000 | 1.000 | 1.000 |
| II | RR1 | 0.950 | 0.833 | 1.000 | 0.941 | 0.875 | 0.962 |
| | RR2 | 0.975 | 0.917 | 1.000 | 1.000 | 1.000 | 1.000 |
| | OR1 | 0.950 | 0.833 | 1.000 | 0.971 | 0.875 | 1.000 |
| | OR2 | 1.000 | 1.000 | 1.000 | 0.971 | 1.000 | 0.962 |
| III | RR1 | 0.875 | 1.000 | 0.853 | 0.941 | 0.500 | 0.969 |
| | RR2 | 0.975 | 1.000 | 0.971 | 0.971 | 0.500 | 1.000 |
| | OR1 | 0.925 | 1.000 | 0.912 | 0.941 | 1.000 | 0.882 |
| | OR2 | 0.975 | 1.000 | 0.971 | 0.971 | 0.500 | 1.000 |
| IV | RR1 | 0.950 | 0.750 | 0.972 | 1.000 | 1.000 | 1.000 |
| | RR2 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | OR1 | 0.975 | 1.000 | 0.972 | 0.941 | 0.750 | 0.967 |
| | OR2 | 0.975 | 1.000 | 0.972 | 1.000 | 1.000 | 1.000 |
| V | RR1 | 0.950 | 0.857 | 0.970 | 0.912 | 1.000 | 0.889 |
| | RR2 | 1.000 | 1.000 | 1.000 | 0.912 | 1.000 | 0.889 |
| | OR1 | 0.975 | 0.857 | 1.000 | 0.882 | 0.857 | 0.889 |
| | OR2 | 0.975 | 0.857 | 1.000 | 1.000 | 1.000 | 1.000 |
| VI | RR1 | 1.000 | 1.000 | 1.000 | 0.912 | 0.667 | 1.000 |
| | RR2 | 1.000 | 1.000 | 1.000 | 0.912 | 0.667 | 1.000 |
| | OR1 | 1.000 | 1.000 | 1.000 | 0.941 | 0.778 | 1.000 |
| | OR2 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

RR1 and OR1 represent the radiology resident's and orthopaedic resident's results with the original CT images, respectively. RR2 and OR2 represent the radiology resident's and orthopaedic resident's results with the addition of model images, respectively.

Table 3
Time consumption per diagnosis by orthopaedic and radiology residents [mean \pm standard deviation] (min).

| Validation set | RR1 | RR2 | P_1 | OR1 | OR2 | P_2 |
|----------------|---------------|---------------|--------|---------------|---------------|--------|
| Internal | 1.8 \pm 0.6 | 1.1 \pm 0.4 | <0.001 | 2.2 \pm 0.4 | 1.6 \pm 0.4 | <0.001 |
| | 1.9 \pm 0.7 | 1.2 \pm 0.8 | <0.001 | 2.1 \pm 0.8 | 1.2 \pm 0.7 | <0.001 |

RR1 and OR1 represent the radiology resident's and orthopaedic resident's results on the original CT images, respectively. RR2 and OR2 represent the radiology resident's and orthopaedic resident's results with the addition of model images, respectively.

Table 4
The diagnostic results agreement between the residents and the gold standard in the internal validation set.

| — | RR and Gold standard | | — | OR and Gold standard | |
|-----|----------------------|-------|---|----------------------|-------|
| | k_1 | k_2 | | k_1 | k_2 |
| I | 0.727 | 1.000 | | 0.918 | 0.918 |
| II | 0.875 | 0.939 | | 0.875 | 1.000 |
| III | 0.635 | 0.908 | | 0.756 | 0.908 |
| IV | 0.722 | 1.000 | | 0.875 | 0.875 |
| V | 0.827 | 1.000 | | 0.908 | 0.908 |
| VI | 1.000 | 1.000 | | 1.000 | 1.000 |

k_1 and k_2 represent the kappa values of the original CT image only and the added model image, respectively. RR: radiology resident; OR: orthopaedic resident.

segmentation accuracy and can obtain a 3D map of the knee in less than 40 s. Our model's 3D maps are comparable in accuracy to the ground truth, which indicates that the automatically generated 3D maps can help surgeons to visualize fracture line alignment, bone fragment spatial distribution, and fracture end separation or angulation in 3D, facilitating understanding of fracture mechanisms and clinical assessment of fracture stability. Further, compared to Graph-Cut, our model excelled in the validation sets by accurately segmenting all four knee bones, with

Table 5
The diagnostic results agreement between the residents and the gold standard in the external validation set.

| — | RR and Gold standard | | — | OR and Gold standard | |
|-----|----------------------|-------|---|----------------------|-------|
| | k_1 | k_2 | | k_1 | k_2 |
| I | 0.717 | 0.872 | | 0.841 | 1.000 |
| II | 0.837 | 1.000 | | 0.915 | 0.922 |
| III | 0.469 | 0.653 | | 0.638 | 0.653 |
| IV | 1.000 | 1.000 | | 0.717 | 1.000 |
| V | 0.767 | 0.767 | | 0.675 | 1.000 |
| VI | 0.746 | 0.746 | | 0.837 | 1.000 |

k_1 and k_2 represent the kappa values of the original CT image only and the added model image, respectively. RR: radiology resident; OR: orthopaedic resident.

average Dice coefficients > 0.950 . Conversely, Graph-Cut failed to infer 46 cases of 74. The segmentation performance of Graph-Cut was significantly worse than that of our model.

Two resident physicians assessed the practicality of the model's automated segmentation results for Schatzker classification. Using the model's results, both residents improved the accuracy, sensitivity, and specificity of their classification diagnoses while spending less time. From the results of evaluating the original CT images alone, the two residents frequently misclassified type I and II as type III. However, with the aid of the model's results, they were able to accurately reclassify most cases that were initially misclassified. Schatzker Type V and VI tibial plateau fractures are usually caused by high-energy external forces, and these are complex injuries that usually require open reduction and internal fixation procedures with a high risk of infection [35]; thus, early and correct diagnosis of the fracture type is necessary. In this study, the results indicated that 3D maps for types V and VI significantly improved classification efficiency and accuracy; this is because 3D maps allow for multi-directional and multi-angle views of the distribution of fracture lines. However, owing to the poor segmentation of very fine fracture lines by our model, this resulted in one type II misdiagnoses as type III, even with the use of the model results. In summary, our findings suggest that the model can segment knee CT images of TPF with high accuracy and may assist clinical junior doctors in assessing patients with

TPF and performing Schatzker classification.

Our study had some limitations. First, we excluded diseases such as combined bone tumours, old fractures, and other conditions that lead to abnormalities in pre-injury bone morphology, so our model may not be able to achieve satisfactory results in this population. Second, our study was single-centred, and the dataset included only adults, which may limit the generalizability of our model to other hospitals and paediatric populations. Subsequently, we will conduct multicentre studies involving all age groups.

5. Conclusion

In conclusion, we proposed a 3D U-net-based deep neural network using a specifically designed strategy for the automatic segmentation of CT knee images of TPF. Our model can effectively segment the four bones of the knee and rapidly generate an accurate 3D map within 40 s. Additionally, the excellent results of the automatic segmentation can assist junior doctors with the Schatzker classification of TPF, which can help to improve diagnostic efficiency and reduce the workload of junior doctors in clinical practice.

Ethics approval

This study was conducted in accordance with the recommendations of our hospital's Ethics Committee.

Funding

This work was supported by the National Natural Science Foundation of China [grant numbers 82171913, 62071305]; Shenzhen Fundamental Research Program [grant numbers JCYJ20190806164409040, JCYJ20220818101816036]; Basic and Applied Basic Research Foundation of Guangdong [grant numbers 2022A1515011241, 2019A1515010847]; and Guangdong-Hong Kong Joint Funding for Technology and Innovation [grant number 2023A0505010021].

CRedit authorship contribution statement

Die Cai: Writing – original draft. **Yu Zhou:** Writing – original draft. **Wenjie He:** Formal analysis, Data curation. **Jichun Yuan:** Data curation. **Chenyuan Liu:** Formal analysis. **Rui Li:** Formal analysis. **Yi Wang:** Writing – review & editing, Conceptualization. **Jun Xia:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I would like to extend my gratitude to all study participants

References

- [1] B. Rudran, C. Little, A. Wiik, K. Logishetty, Tibial plateau fracture: anatomy, diagnosis and management, *Br. J. Hosp. Med.*, Lond. 81 (2020) 1–9, <https://doi.org/10.12968/hmed.2020.0339>.
- [2] C.T. Born, J.A. Gil, J.P. Johnson, Periprosthetic tibial fractures, *J. Am. Acad. Orthop. Surg.* 26 (2018) e167–e172, <https://doi.org/10.5435/JAAOS-D-16-00387>.
- [3] M.K. Júnior, F. Fogagnolo, R.C. Bitar, R.L. Freitas, R. Salim, C.A. Jansen Paccolla, Tibial plateau fractures, *Rev. Bras. Ortop.* 44 (2009) 468–474, [https://doi.org/10.1016/S2255-4971\(15\)30142-7](https://doi.org/10.1016/S2255-4971(15)30142-7).
- [4] T. Vendeuvre, L.E. Gayet, Percutaneous treatment of tibial plateau fractures, *Orthop. Traumatol. Surg. Res.* 107 (2021) 102753, <https://doi.org/10.1016/j.otsr.2020.102753>.
- [5] K. Li, S. Zhang, X. Qiu, H. Huang, H. Sheng, Y. Zhang, J. Chang, J. Kuang, J. Yang, Optimal surgical timing and approach for tibial plateau fracture, *Technol. Health Care.* 30 (2022) 545–551, <https://doi.org/10.3233/THC-228050>.
- [6] D. Wasserstein, P. Henry, J.M. Paterson, H.J. Kreder, R. Jenkinson, Risk of total knee arthroplasty after operatively treated tibial plateau fracture: a matched-population-based cohort study, *J. Bone Joint Surg. Am.* 96 (2014) 144–150, <https://doi.org/10.2106/JBJS.L.01691>.
- [7] A. Gamulin, A. Lübbecke, P. Belinga, P. Hoffmeyer, T.V. Perneger, M. Zingg, G. Cunningham, Clinical and radiographic predictors of acute compartment syndrome in the treatment of tibial plateau fractures: a retrospective cohort study, *BMC Musculoskelet. Disord.* 18 (2017) 307, <https://doi.org/10.1186/s12891-017-1680-4>.
- [8] J. Mthethwa, A. Chikate, A review of the management of tibial plateau fractures, *Musculoskelet. Surg.* 102 (2018) 119–127, <https://doi.org/10.1007/s12306-017-0514-8>.
- [9] A. Kumar, A.P. Whittle, Treatment of complex (Schatzker Type VI) fractures of the tibial plateau with circular wire external fixation: retrospective case review, *J. Orthop. Trauma.* 14 (2000) 339–344, <https://doi.org/10.1097/00005131-200006000-00006>.
- [10] P.J. Papagelopoulos, A.A. Partsinevelos, G.S. Themistocleous, A.F. Mavrogenis, D. S. Korres, P.N. Souacos, Complications after tibia plateau fracture surgery, *Injury.* 37 (2006) 475–484, <https://doi.org/10.1016/j.injury.2005.06.035>.
- [11] H. Tschern, P. Lobenhoffer, Tibial plateau fractures. Management and expected results, *Clin. Orthop. Relat. Res.* 292 (1993) 87–100, <https://doi.org/10.1097/00003086-199307000-00011>.
- [12] J. Schatzker, Compression in the surgical treatment of fractures of the tibia, *Clin. Orthop. Relat. Res.* 105 (1974) 220–239, <https://doi.org/10.1097/00003086-197411000-00015>.
- [13] B.K. Markhardt, J.M. Gross, J.U. Monu, Schatzker classification of tibial plateau fractures: use of CT and MR imaging improves assessment, *RadioGraphics.* 29 (2009) 585–597, <https://doi.org/10.1148/rq.292085078>.
- [14] J.M. Huitema, N. van der Gaast, L. Brouwers, R.L. Jaarsma, J.N. Doornberg, M.J. R. Edwards, E. Hermans, Traumatoplatform 3D Consortium, Are 3D-printed models of tibial plateau fractures a useful addition to understanding fractures for junior surgeons? *Clin. Orthop. Relat. Res.* 480 (2022) 1170–1177, <https://doi.org/10.1097/CORR.0000000000002137>.
- [15] K. Atesok, J. Finkelstein, A. Khoury, A. Peyser, Y. Weil, M. Liebergall, R. Mosheiff, The use of intraoperative three-dimensional imaging (ISO-C-3D) in fixation of intraarticular fractures, *Injury.* 38 (2007) 1163–1169, <https://doi.org/10.1016/j.injury.2007.06.014>.
- [16] T. Gössling, K. Klingler, J. Geerling, H. Shin, M. Fehr, C. Krettek, T. Hüfner, Improved intra-operative reduction control using a three-dimensional mobile image intensifier – a proximal tibia cadaver study, *Knee.* 16 (2009) 58–63, <https://doi.org/10.1016/j.knee.2008.07.012>.
- [17] W. Hackl, J. Riedl, M. Reichkendler, K.P. Benedetto, M. Freund, R. Bale, Preoperative computerized tomography diagnosis of fractures of the tibial plateau, *Unfallchirurg.* 104 (2001) 519–523, <https://doi.org/10.1007/s001130170115>. [In German, English Abstract].
- [18] S. Wicky, P.F. Blaser, C.H. Blanc, P.F. Leyvraz, P. Schnyder, R.A. Meuli, Comparison between standard radiography and spiral CT with 3D reconstruction in the evaluation, classification and management of tibial plateau fractures, *Eur. Radiol.* 10 (2000) 1227–1232, <https://doi.org/10.1007/s003300000326>.
- [19] L. Macarini, M. Murrone, S. Marini, R. Calbi, M. Solarino, B. Moretti, Tibial plateau fractures: evaluation with multidetector-CT, *Radiol. Med.* 108 (2004) 503–514 [English, Italian].
- [20] S.V. Yacoubian, R.T. Nevins, J.G. Sallis, H.G. Potter, D.G. Lorch, Impact of MRI on treatment plan and fracture classification of tibial plateau fractures, *J. Orthop. Trauma.* 16 (2002) 632–637, <https://doi.org/10.1097/00005131-200210000-00004>.
- [21] M. Adams, W. Chen, D. Holcdorf, M.W. McCusker, P.D. Howe, F. Gaillard, Computer vs human: deep learning versus perceptual training for the detection of neck of femur fractures, *J. Med. Imaging Radiat. Oncol.* 63 (2019) 27–32, <https://doi.org/10.1111/1754-9485.12828>.
- [22] A. Brett, C.G. Miller, C.W. Hayes, J. Krasnow, T. Ozanian, K. Abrams, J.E. Block, C. van Kuijk, Development of a clinical workflow tool to enhance the detection of vertebral fractures: accuracy and precision evaluation, *Spine (Phila Pa 1976).* 34 (2009) 2437–2443. Doi: 10.1097/BRS.0b013e3181b2eb69.
- [23] C. Chen, B. Liu, K. Zhou, W. He, F. Yan, Z. Wang, R. Xiao, CSR-Net: cross-scale residual network for multi-objective scaphoid fracture segmentation, *Comput. Biol. Med.* 137 (2021) 104776, <https://doi.org/10.1016/j.combiomed.2021.104776>.
- [24] S.W. Chung, S.S. Han, J.W. Lee, K.S. Oh, N.R. Kim, J.P. Yoon, J.Y. Kim, S.H. Moon, J. Kwon, H.J. Lee, Y.M. Noh, Y. Kim, Automated detection and classification of the proximal humerus fracture by using deep learning algorithm, *Acta Orthop.* 89 (2018) 468–473, <https://doi.org/10.1080/17453674.2018.1453714>.
- [25] Y. Deng, L. Wang, C. Zhao, S. Tang, X. Cheng, H.W. Deng, W. Zhou, A deep learning-based approach to automatic proximal femur segmentation in quantitative CT images, *Med. Biol. Eng. Comput.* 60 (2022) 1417–1429, <https://doi.org/10.1007/s11517-022-02529-9>.
- [26] D.H. Kim, T. MacKinnon, Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks, *Clin. Radiol.* 73 (2018) 439–445, <https://doi.org/10.1016/j.crad.2017.11.015>.
- [27] Y.J. Kim, S.R. Lee, J.Y. Choi, K.G. Kim, Using convolutional neural network with Taguchi parametric optimization for knee segmentation from X-ray images, *BioMed Res. Int.* 2021 (2021) 5521009, <https://doi.org/10.1155/2021/5521009>.
- [28] P. Song, Z. Fan, X. Zhi, Z. Cao, S. Min, X. Liu, Y. Zhang, X. Kong, W. Chai, Study on the accuracy of automatic segmentation of knee CT images based on deep learning,

- Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi. 36 (2022) 534–539, <https://doi.org/10.7507/1002-1892.202201072> [In Chinese, English Abstract].
- [29] Z. Zhou, G. Zhao, R. Kijowski, F. Liu, Deep convolutional neural network for segmentation of knee joint anatomy, *Magn. Reson. Med.* 80 (2018) 2759–2770, <https://doi.org/10.1002/mrm.27229>.
- [30] F. Isensee, P.F. Jaeger, S.A.A. Kohl, J. Petersen, K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, *Nat. Methods.* 18 (2021) 203–211, <https://doi.org/10.1038/s41592-020-01008-z>.
- [31] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger (Eds.), U-net: learning dense volumetric segmentation from sparse annotation. *Med image comput comput assist Interv MICCAI, Proceedings, Part II 19: 19th International Conference; October the 17–21 2016; Athens, Greece, Springer International Publishing, 2016, p. 3D.*
- [32] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation. *Med image comput comput assist Interv MICCAI, Proceedings, Part III 18: 18th International Conference; October the 5–9 2015; Munich, Germany, Springer International Publishing, 2015.*
- [33] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, *Biometrics.* 33 (1977) 159–174, <https://doi.org/10.2307/2529310>.
- [34] M. Krčah, G. Székely, R. Blanc, Fully Automatic and Fast Segmentation of the Femur Bone from 3D-CT, in: *Images with No Shape Prior IEEE International Symposium on Biomedical Imaging: from Nano to Macro 2011, IEEE Publications, 2011, pp. 2087–2090.*
- [35] T.B. Pun, V.P. Krishnamoorthy, P.M. Poonnoose, A.T. Oommen, R.J. Korula, Outcome of Schatzker type V and VI tibial plateau fractures, *Indian J. Orthop.* 48 (2014) 35–41. Doi: 10.4103/0019-5413.125490, <http://www.ncbi.nlm.nih.gov/pubmed/24600061>.