Medical Image Analysis 102 (2025) 103506

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Preoperative fracture reduction planning for image-guided pelvic trauma surgery: A comprehensive pipeline with learning[☆]

Yanzhen Liu [a,1], Sutuke Yibulayimu [a,1], Yudi Sang [c,*], Gang Zhu [c], Chao Shi [a], Chendi Liang [a], Qiyong Cao [b], Chunpeng Zhao [b], Xinbao Wu [b], Yu Wang [a,*]

[a] 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

[b] Department of Orthopaedics and Traumatology, Beijing Jishuitan Hospital, Capital Medical University, Beijing, 100035, China

[c] Beijing Rossum Robot Technology Co., Ltd., Beijing, 100088, China

ARTICLE INFO

Dataset link: https://zenodo.org/records/10927452, https:// github.com/Sutuk/Clinical-data-on-pelvic-fractures, https://github.com/YzzLiu/FracSegNet

Keywords:
Pelvic fracture
Fracture segmentation
Reduction planning
Deep learning
Morphable models

ABSTRACT

Pelvic fractures are among the most complex challenges in orthopedic trauma, which usually involve hipbone and sacrum fractures, as well as joint dislocations. Traditional preoperative surgical planning relies on the operator's subjective interpretation of CT images, which is both time-consuming and prone to inaccuracies. This study introduces an automated preoperative planning solution for pelvic fracture reduction, addressing the limitations of conventional methods. The proposed solution includes a novel multi-scale distance-weighted neural network for segmenting pelvic fracture fragments from CT scans, and a learning-based approach to restore pelvic structure, combining a morphable model-based method for single-bone fracture reduction and a recursive pose estimation module for joint dislocation reduction. Comprehensive experiments on a clinical dataset of 30 fracture cases demonstrated the efficacy of our methods. Our segmentation network outperformed traditional max-flow segmentation and networks without distance weighting, achieving a Dice similarity coefficient (DSC) of 0.986 $\pm$ 0.055 and a local DSC of 0.940 $\pm$ 0.056 around the fracture sites. The proposed reduction method surpassed mirroring and mean template techniques, and an optimization-based joint matching method, achieving a target reduction error of (3.265 $\pm$ 1.485) mm, rotation errors of (3.476 $\pm$ 1.995)°, and translation errors of (2.773 $\pm$ 1.390) mm. In the proof-of-concept cadaver studies, our method achieved a DSC of 0.988 in segmentation and 3.731 mm error in reduction planning, which senior experts deemed excellent. In conclusion, our automated approach significantly improves traditional preoperative planning, enhancing both efficiency and accuracy in pelvic fracture reduction.

1. Introduction

Pelvic trauma represents one of the most severe orthopedic injuries, with a disability rate of 50%–60% and a mortality rate exceeding 13% (Halawi, 2016; Miller et al., 2003). Among pelvic trauma patients, 74.5% require surgical intervention to restore stability (Hermans et al., 2017; Khurana et al., 2014). Traditional open reduction surgery, while effective under direct visualization, requires large incisions, resulting in significant blood loss and posing risks of vascular and nerve damage, as well as postoperative infections. The deep anatomical position of the pelvis and its proximity to vital organs make the procedure both complex and demanding. With advances in intraoperative imaging and the adoption of minimally invasive techniques, closed reduction under image guidance has emerged as an ideal alternative. This approach offers advantages such as reduced bleeding, lower infection rates, decreased soft tissue damage, faster postoperative recovery, and the potential for robot-assisted surgery (Tosounidis et al., 2012).

Effective preoperative surgical planning, which typically includes segmenting fracture fragments from CT and determining their optimal positions, is crucial for successful image-guided closed reduction. Errors in planning can lead to incomplete anatomical restoration and unstable fixation, resulting in complications such as post-traumatic osteoarthritis and recurrent fractures (Olson et al., 2012). Among patients with insufficient reduction, 63% experience persistent pain, 28% suffer

☆ This work was supported by the National Key Research and Development Program of China (Grant 2022YFC2504304), the Beijing Science and Technology Project (Grant Z221100003522007), and the Natural Science Foundation of Beijing (Grant L222136).

* Corresponding authors.
E-mail addresses: sangyudi@rossumrobot.cn (Y. Sang), wangyu@buaa.edu.cn (Y. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.media.2025.103506
Received 2 August 2024; Accepted 11 February 2025
Available online 21 February 2025

sensory and motor deficits, and 32% cannot return to their previous occupations (Smith et al., 2005).

Conventionally, surgical plans are made manually through interactive medical image processing software. These tools enable operators to delineate fragments in 3D views or modify masks layer by layer to segment bone fragments and then virtually manipulate the 3D bone surface models towards the imagined healthy state of the pelvis (Luque-Luque et al., 2021; Chowdhury et al., 2009). However, for complex fractures, especially those involving multiple bone fragments and collision/compression between fragments, manual planning is very challenging and time-consuming, often taking more than two hours to first separate each fragment and then assemble the puzzle (Suero et al., 2010; Zeng et al., 2024b). Additionally, the manual planning process heavily relies on operator's skill and experience, making it difficult to achieve consistent and high-quality outcomes. An automatic planning solution is in great demand to guarantee the planning accuracy and efficiency, as well as to alleviate the reliance on medical experts and therefore help spread the application of advanced image-guided reduction surgery.

Identifying and isolating each bone fragment in the CT is the first crucial step in the planning process. Geometry-based techniques, such as max-flow, enable semi-automatic or automatic segmentation of non-contact fractures. However, these methods often fail with more complex fracture types, such as contact or compression fractures (Krčah et al., 2011). Learning-based methods are considered promising for addressing these challenges. Nevertheless, existing research predominantly focuses on segmenting overall anatomical structures or detecting fractures, with few studies specifically addressing fracture segmentation (Liu et al., 2021; Zeng et al., 2024a). This task remains challenging for learning-based methods for several reasons: (1) Unlike organ or tumor segmentation tasks where models can implicitly learn the shape priors of objects, it is difficult to learn the shape information of bone fragments due to the significant variations in fracture types and shapes (Kuiper et al., 2022); (2) The fracture surface can exhibit various forms, including large spaces (fragments isolated and moved), small gaps (fragments isolated but not moved), creases (fragments not completely isolated), compression (fragments collided), and their combinations. These variations result in diverse image intensity profiles around the fracture site; (3) The variable number of bone fragments in pelvic fractures complicates the development of a consistent labeling strategy applicable to all types and cases. These factors collectively underscore the complexities of fracture segmentation, highlighting the need for continued research and development in this area.

The other major step in the planning process is to assemble the puzzle of the extracted bone fragments and restore the healthy pelvic anatomy. Some attempts has been made to match the features of complementary fracture surfaces but was challenged by the precise definition of fracture lines. Other methods mainly focus on various template methods, including mean templates, mirror templates, and adaptive templates. Notably, statistical shape models (SSM) have been employed as adaptive templates to warp and align with the fragments (Han et al., 2019). However, these methods mainly focus on single-bone structures and often do not perform as well for composite and complex structures like the pelvis as a whole. Recently, deep network has also been employed to reconstruct the target morphology of fractured bones as template for reduction (Zeng et al., 2024b). However, like other template-based methods, deep learning generated template faces the challenge of balancing the need for structural regularity and the flexibility required to accurately align the unique shape of injured bones.

To address these challenges and streamline the preoperative workflow, we propose a learning-based solution for pelvic fracture segmentation and reduction planning. Our approach consists of a two-stage deep learning-based automatic segmentation method for pelvic fractures and a novel recursive pose estimation module combined with statistical shape modeling to restore healthy pelvic morphology. Major contributions of this study include:

- We propose a comprehensive, fully automated preoperative planning process for pelvic fracture reduction that includes both automatic segmentation and target pose planning, with significant improvements in both accuracy and efficiency.
- We develop a novel pelvic fracture segmentation network, which incorporates a multi-scale distance-weighted loss function in the training process to improve segmentation accuracy near the fracture regions.
- We develop a novel learning-based approach to address the restoration of pelvic structure, combining a morphable model-based method for reduction of single-bone fractures and a simulation-to-reality (sim-to-real) recursive pose estimation module for reduction of joint dislocations, addressing the challenge of limited fractured data.

2. Related works

2.1. Fracture segmentation

Existing research in orthopedic image segmentation primarily focuses on segmenting the overall anatomical structures such as pelvis, ribs, spine, and skull (Yang et al., 2021; Cheng et al., 2021; Liu et al., 2022). In pelvic imaging, Liu et al. (2021) utilized a cascade 3D UNet to segment the hipbone, sacrum, and lumbar vertebrae in CT scans. However, in the context of pelvic fractures, it is crucial to extract the fracture fragments rather than the whole anatomical region. Methods used for fracture segmentation can be categorized into intensity-based methods, geometric methods, and deep learning methods.

Intensity-based methods, such as fixed or adaptive thresholding, watershed algorithms, and region growing, typically rely on grayscale intensity similarity and boundary gradient continuity for segmentation (Tomazevic et al., 2010; Neubauer et al., 2005; Ruikar et al., 2019). These methods often face obstacles due to the intensity difference between cortical and trabecular bones, as well as the overlapping intensity ranges between trabecular bones and soft tissues. Such challenges usually pose great impact on the occurrence of hollows within the segmentation masks, resulting in the inaccurate representation of the bone shape and topology (Paulano et al., 2014).

Geometry-based methods, such as graph cuts and template-based methods, depend on geometric connectivity or morphological similarity. Yuan et al. proposed a semi-automatic graph-cut method based on continuous max flow for segmenting pelvic fractures, which involves pre-determining the number of fragments and selecting seed points on each fragment (Yuan et al., 2014). Similarly, Wang et al. developed an automatic max-flow segmentation method using boundary-enhanced filters (Wang et al., 2019). While graph-cut can effectively segment non-contacting fragments, they face difficulties when fragments are in contact or compressed. Template-based methods involve registering a CT scan to a healthy template and using graph partitioning techniques for label production. Deng et al. used the contralateral bone as a template and implemented a dual-task synergy approach for fragment segmentation and reduction planning using a deep learning-based dense feature matching strategy (Deng et al., 2023). However, this method is effective only when a complete contralateral template is available, and its success heavily depends on the registration accuracy. The complex and varying forms of fractures can lead to unreliable segmentation results (Seim et al., 2008).

Deep learning-based fracture segmentation is still a relatively underexplored area, yet some studies have demonstrated its significant potential. Yang et al. applied a two-stage Mask R-CNN model to locate and segment intertrochanteric fractures in 2D images (Yang et al., 2022). This approach, however, overlooked the 3D spatial context and the overall features, which can result in misinterpretation of fractures. Kim et al. used the DeepLab model to automatically segment bone fragments from tibia and fibula in CT (Kim et al., 2023). Wang et al. employed a V-net-based network, IFFCT, to segment intertrochanteric fractures (Wang et al., 2022). Similarly, Zeng et al. proposed a dual-stream learning network for automatic segmentation of pelvic fractures (Zeng et al., 2024a). Many of these studies have reported decreased performance on small bone fragments. Additionally, they have identified the limited availability of fractured data as a significant challenge that limits segmentation accuracy and robustness. Additionally, due to the variability in shape, location, and the varying number of bone fragments, models need to possess advanced spatial relationship understanding and detailed feature learning capabilities, which current methods struggle to achieve.

2.2. Fracture reduction planning

Fracture reduction planning, also known as target pose planning or virtual fracture reduction, involves repositioning bone fragments to restore the original anatomical structure. Fracture surface matching is an intuitive method that aligns fragments based on their fracture surface features (Willis et al., 2007; Paulano-Godino and Jiménez-Delgado, 2017; Luque-Luque et al., 2021). However, this method relies heavily on precise identification and segmentation of fracture lines, which is particularly challenging in comminuted fractures with multiple fragments and missing small pieces.

Template-based approaches avoid the fracture line identification process by constructing a target template and aligning the fragments to this template, making them more robust. These methods include using a mean template modeled from pelvic shapes or a template mirrored from the healthy contralateral side (Ead et al., 2021; Zhang et al., 2016; Ead et al., 2020; Zhao et al., 2022). The mean template approach is less adaptable to individual morphology, often resulting in poor robustness, while the contralateral mirroring method, effective only for iliac fractures, is limited by the inherent asymmetry between the two ilia (Gnat et al., 2009). Liu et al. (2024a) proposed using the mirrored contralateral innominate bone for initial registration, followed by local adjustments to address structural asymmetry, but this method faces significant challenges in cases of bilateral injuries where the mirrored part cannot provide accurate references.

Han et al. built SSM as adaptive templates for the entire pelvis. However, due to the complex and composite structure of the pelvis, the results were less than satisfactory (Han et al., 2020). To improve this, they combined single-bone SSM with a statistical pose model (SPM) to create a comprehensive pelvis model as a two-step solution (Han et al., 2021). In this approach, the SPM focused solely on pose parameters, neglecting detailed shape nuances, thereby limiting its accuracy in accommodating minor morphological discrepancies. Lu et al. (2023) constructed separate SSM for male and female pelvises but did not include the sacral part in their models. Alternatively, SSM combined with joint surface detection and matching methods have been proposed to address pelvic reduction planning (Yibulayimu et al., 2023). However, for complex trauma types, especially those involving distorted joint surfaces or comminuted fractures, manually designed matching criteria may not robustly capture the features.

In reconstructive bone surgery, deep learning has been applied to model complex structures such as the skull and jaw using generative models (Kodym et al., 2020; Abdi et al., 2019). Recently, Zeng et al. (2024b) utilized deep learning for pelvic fracture reduction, learning the deformation field between fractured and healthy bones to create the target template. This work separately addressed the iliac and sacral bone fractures, but the restoration of the whole pelvic structure from joint dislocation was not covered in its method.

2.3. Differences from the conference version

This study extends our works in our initial conference papers presented at the 26th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2023) by significantly enhancing performance, depth, and practical utility (Liu et al., 2023; Yibulayimu et al., 2023).

- We have improved the automatic pelvic fracture reduction planning pipeline, demonstrating its robustness and effectiveness through comprehensive experimental analyses.
- We have optimized the fracture segmentation method and expanded the fracture CT dataset to encompass a wider range of fracture types, thereby improving segmentation accuracy and robustness.
- We have refined the fracture reduction planning method by integrating a novel pose estimation network, which enhances the accuracy and practical applicability of the reduction plans.

3. Methods

The workflow for preoperative planning of pelvic fracture reduction consists of an image segmentation step and a fracture reduction step, illustrated in Fig. 1. Initially, an anatomical segmentation network is employed to generate 3D models of the left hipbone, right hipbone, and sacrum. Then, a fracture segmentation network identifies and segments bone fragments within each extracted hipbone and sacral region. Following segmentation, a post-processing step refines the segmentation by isolating and labeling individual components to produce the final segmented result. In cases involving bone fractures, a morphable model-based approach is utilized. This method registers fragments of each bone to an adaptive template, thereby reconstructing the morphology of individual bone anatomy. Finally, a recursive pose estimation module iteratively predicts the target pose of the dislocated bone to restore the overall pelvic morphology to its healthy state.

3.1. Pelvic bone extraction

Our initial step aims to accurately extract pelvic bone regions from CT scans. We employ a cascaded 3D UNet framework designed for predicting anatomical labels from preprocessed CT images. This framework consists of two sequential five-layer UNet models. The first model operates on low-resolution images to establish a broad contextual understanding and segments the hipbones and sacrum in coarse detail. Subsequently, the second model refines this segmentation on full-resolution images by integrating the initial coarse segmentation with the CT volumes.

To ensure reliable performance, the networks are pre-trained on the *CTPelvic1K* dataset, which includes over 1000 high-resolution pelvis scans. They are further fine-tuned using the *Pelvic Bone Fragments with Injuries (PENGWIN)* dataset, a curated collection that encompasses diverse fracture cases, discussed in Section 4.1.

3.2. Pelvic fracture segmentation

A deep network is developed to segment bone fragments within each identified region. Fig. 2 illustrates our approach. A 3D UNet is used as the foundational model (Isensee et al., 2021). The network model learns a non-linear relationships between the masked CT volume (X) and the ground truth fragment labels (Y). To establish a consistent labeling strategy applicable to all fracture types, we categorize each bone into three distinct labels: background, main fragment, and minor fragments. The main fragment refers to the largest central fragment.

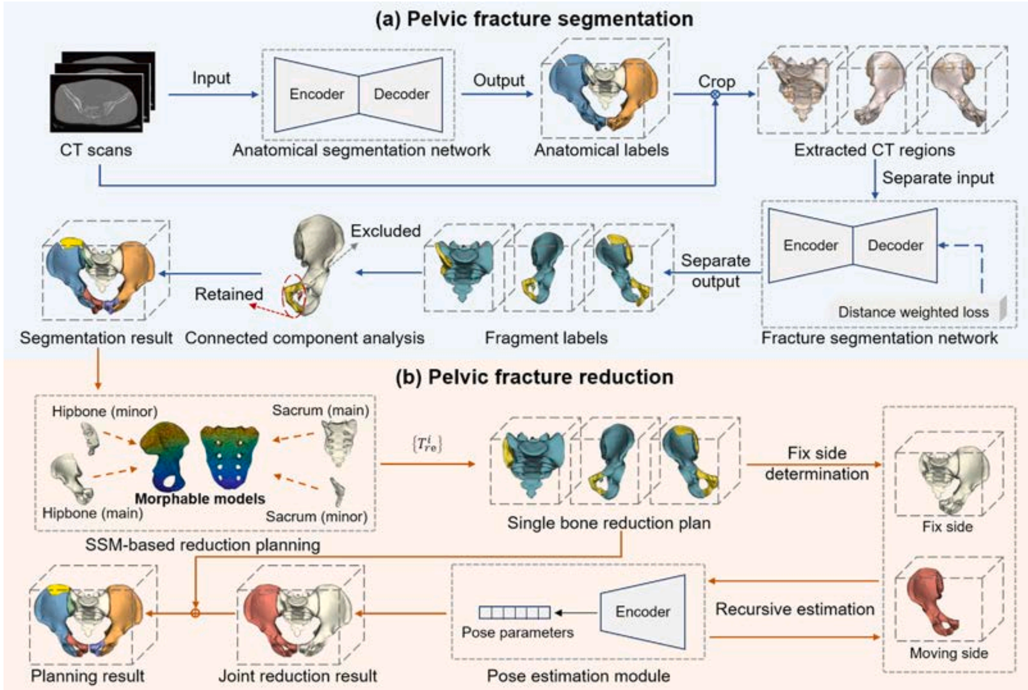

Fig. 1. Workflow of the proposed preoperative planning method for pelvic fracture reduction. (a) An anatomical segmentation network is used to generate 3D masks of the left hipbone, right hipbone, and sacrum. Subsequently, a fracture segmentation network isolates the bone fragments within each extracted hipbone and sacral region. The post-processing step further separates isolated components to yield the final segmentation result. (b) The morphable model-based reduction planning method computes the transformation of each bone fragment to restore the single-bone morphology. Subsequently, the target pose of each dislocated bone relative to the entire pelvic anatomy is estimated using a recursive pose estimation module.

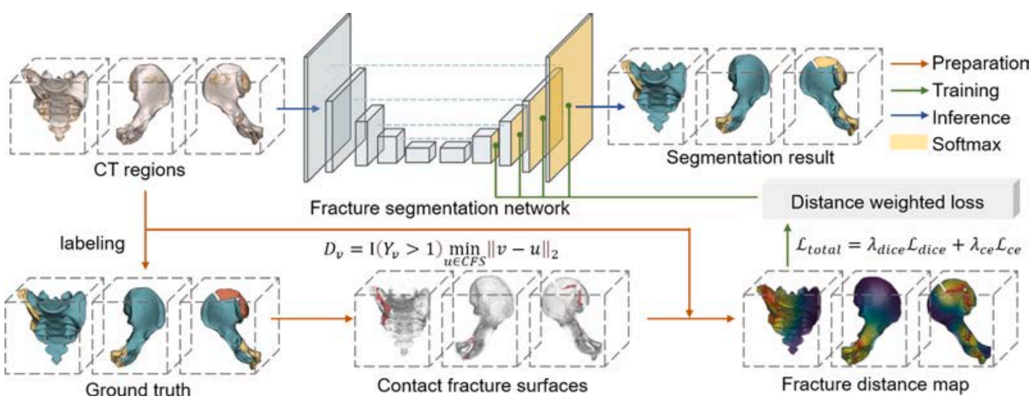

Fig. 2. Training of the fracture segmentation network. The network takes masked bone regions of CT as input and generates main and minor fragment labels as output. Ground truth labels are used to derive the fracture distance map (FDM), which are then used to compute a distance-weighted loss to enhance segmentation accuracy, particularly in contact fracture surface (CFS).

3.2.1. Distance-weighted loss

The contact fracture surface (CFS) represents the area where bones collide and overlap due to compression. This region poses significant challenges for both human operators and network models to accurately delineate. Our primary objective is to improve segmentation accuracy specifically within this complex region. To address this, we incorporate fracture distance map (FDM) into the training process.

The FDM is derived from ground-truth segmentations of each data sample prior to training. This representation conveys boundary, shape, and positional information of the segmented object. Initially, CFS regions are identified by comparing labels within each voxel's neighborhood. Subsequently, we compute the distance of each foreground voxel to the nearest CFS, denoted as D_v, and normalize it:

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

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

where V represents all foreground voxels, v and u are voxel indices, Y is the ground-truth segmentation, $\mathrm{I}(Y_v \geq 1)$ is the foreground indicator function, and $\hat{D}$ is the normalized distance. This distance informs the FDM weight $\hat{W}$:

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

$$\hat{W}_v = \frac{W_v \cdot |V|}{\sum_{v \in V} W_v}, \tag{4}$$

where λ_{back} is the weight for background voxels, and λ_{FDM} adjusts the slope of the activation function. To ensure the equivalence across different samples, weights are normalized by their sum.

The FDM weight $\hat{W}$ is utilized to calculate the weighted Dice loss $\mathcal{L}_{dice}$ and cross-entropy loss $\mathcal{L}_{ce}$, enhancing the importance of CFS in training.

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

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

where L represents the number of classes, and P_v^l and Y_v^l are the predicted output and ground truth for the vth voxel of the lth label, respectively. The overall loss is a weighted sum of the two components:

$$\mathcal{L}_{total} = \lambda_{dice}\mathcal{L}_{dice} + \lambda_{ce}\mathcal{L}_{ce}, \tag{7}$$

where λ_{dice} and λ_{ce} are the balancing weights.

3.2.2. Training strategies

We employ a multi-scale deep supervision strategy during model training to enhance feature learning at different scales (Wang et al., 2023). The deeper layers primarily capture global features, including shape and structural information, while the shallower layers focus on local features that help delineate fracture surfaces. Auxiliary losses are integrated into the decoder at all resolution levels, except the lowest. These losses are computed using the corresponding down-sampled FDM $\hat{W}_v^n$ and down-sampled ground truth Y_v^n. For the nth level, the loss $\mathcal{L}^n$ is calculated with a specific λ_{FDM} as shown in Eq. (3), where λ_{FDM} for each layer decreases by a factor of 2 with increasing depth:

$$\lambda_{FDM}^n = \lambda_{FDM}\cdot\frac{1}{2^{n-1}}. \tag{8}$$

This approach ensures that local CFS information receives more focus in the shallow layers, while the weights become more uniform in the deeper layers.

To stabilize network training, we implement a smooth transition strategy, which maintains the model's focus on global features in the early stages and gradually shifts attention towards the fracture site as training progresses (Qamar et al., 2020). This transition dynamically adjusts the proportion of the FDM in the overall weight matrix based on the training iterations. The dynamic weight is computed as follows:

$$W_{st} = \begin{cases}\mathcal{J}, & \text{if } t < \tau_{begin},\\ \frac{1}{1+\delta}\mathcal{J} + \frac{\delta}{1+\delta}\hat{W}, & \text{if } \tau_{begin} \le t \le \tau_{begin} + \tau_{smooth},\\ \hat{W}, & \text{if } t > \tau_{begin} + \tau_{smooth},\end{cases} \tag{9}$$

$$\delta = -\ln\left(1 - \frac{t - \tau_{begin}}{\tau_{smooth} + \epsilon}\right), \tag{10}$$

where $\mathcal{J}$ is an all-ones matrix with the same size as the input volume, t is the current iteration number, τ_{begin} marks the start of the transition, τ_{smooth} is the duration of the smooth transition phase, and ϵ is a small positive constant. The dynamic weight W_{st} is adjusted by modulating the relative proportions of $\mathcal{J}$ and $\hat{W}$.

3.2.3. Post-processing

Connected component analysis (CCA) is widely recognized in segmentation tasks (Ghimire et al., 2022). However, directly applying CCA to fracture segmentation can be challenging due to the collision between fragments. To address this, we first remove the main central fragment, which simplifies the isolation of minor fragments. Each individual fragment undergoes CCA, which assigns it a distinct label. Fragments measuring less than 1 cm^3 are excluded from further analysis, as their impact on surgical outcomes is typically negligible.

3.3. Restoration of single-bone anatomies

To virtually restore the structure of fractured hipbone and sacrum, we utilize Gaussian process morphable models (GPMMs) (Luthi et al., 2018) to establish adaptive bone templates for the bones, followed by sequentially registering each fragment to the morphable template.

3.3.1. Gaussian process morphable models

To model the hipbone and sacrum, we first align the bone segmentations to a reference shape Γ_R. A bone shape Γ_B is then obtained by warping Γ_R using a deformation field $u(x)$ through a non-rigid registration algorithm (Audenaert et al., 2019).

GPMMs represent the deformation field as a Gaussian process $GP(\mu, k)$, where μ is the mean function and k is the covariance function. The empirical mean $\mu(x)$ and the covariance function $k(x, y)$ are defined as:

$$\mu(x) = \frac{1}{n}\sum_{i=1}^{n} u_i(x), \tag{11}$$

$$k(x, y) = \frac{1}{n-1}\sum_{i=1}^{n}\left(u_i(x) - \mu_{SM}(x)\right)\left(u_i(y) - \mu_{SM}(y)\right)^T, \tag{12}$$

where n is the number of training surfaces and $u_i(x)$ represents the deformation at point x on the ith sample surface.

GPMMs are parameterized using principal component analysis (PCA), allowing the bone shape to be expressed as:

$$\Gamma_B = \{x + \mu(x) + Pb \mid x \in \Gamma_R\}, \tag{13}$$

where P and b are the principal components and the weight vector, respectively. GPMMS modeled with the empirical kernel here is equivalent to a statistical shape model (SSM) in a finite domain.

A set of CT images with intact pelvises from the Cancer Imaging Archive is utilized to create a comprehensive pelvis atlas (Clark et al., 2013). The dataset includes a total of 53 cases, with 28 male and 25 female. This dataset serves as the foundation for establishing the GPMMs used in our model and is publicly available on GitHub at https://github.com/Sutuk/Clinical-data-on-pelvic-fractures.

3.3.2. Sequential fragment registration algorithm

We introduce the sequential fragment registration algorithm (SFRA) for the incremental alignment of bone fragments to the morphable template. In this approach, each fragment is registered to the GPMM template one at a time using the iterative closest point (ICP) algorithm. After aligning a fragment, the corresponding section is removed from the template before proceeding to the next fragment. Once all fragments are registered, the GPMM is non-rigidly deformed to fit the combined point cloud of all fragments.

During this process, model parameters b are optimized based on bi-directional vertex correspondences. For each source point x, a forward displacement $\delta_{\rightarrow}(x) = v_{\rightarrow}(x) - x$ and one or more backward displacements $\delta_{\leftarrow}(x) = v_{\leftarrow}(x) - x$ are computed. The corresponding point displacement $\delta(x)$ is defined as the maximum of the forward and backward displacements if they are within a given threshold ε:

$$\delta(x) = \begin{cases}\max\left(\left|\delta_{\rightarrow}(x)\right|, \left|\delta_{\leftarrow}(x)\right|\right), & \text{if } |\delta(x)| < \varepsilon,\\ \mathbf{O}, & \text{if } |\delta(x)| \ge \varepsilon.\end{cases} \tag{14}$$

where $\mathbf{O}$ denotes the zero vector.

The optimal parameter $\hat{b}$ is obtained by minimizing the discrepancy function:

$$\hat{b} = \operatorname{argmin}\left\{\delta(x) - \mu(x) - Pb \mid x \in \Gamma_R\right\}. \tag{15}$$

This minimization is constrained within three standard deviations of the eigenvalue λ_i, ensuring physiologically feasible bone shapes (Albrecht et al., 2013).

Fragment registration and model deformation are alternately iterated until the model parameters converge. The adaptive shape of the GPMM is regarded as the target morphology of a bone, and the transformation of registered fragments forms the output of the single bone reduction planning.

3.4. Pose estimation for joint reduction

In this approach, pelvic joint reduction planning is formulated as a problem of estimating transformation parameters. The estimated transformation with a degree of freedom (DoF) of six is used to describe the target pose of the dislocated hipbone.

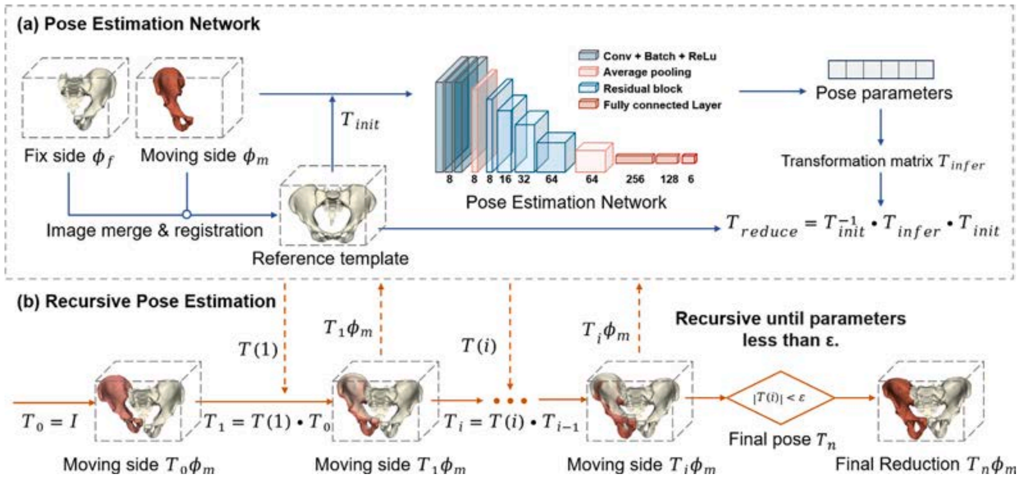

Fig. 3. Recursive pose estimation process for joint reduction planning. (a) The pose estimation network takes two-channel input volumes from the fixed and moving sides and outputs the reduction pose parameters. (b) The network recursively refines the pose estimate until convergence.

3.4.1. Pose estimation network

As shown in Fig. 3, the pose estimation network takes the binary pelvic bone segmentation volumes as input and outputs pose parameters. Specifically, binary masks of the fixed side ϕ_F (sacrum and hipbone) and the moving side ϕ_M (dislocated hipbone) are resampled to a spacing of 2 mm and spatial dimensions of $(180, 128, 96)$. Binary masks are used instead of the original CT volumes mainly for efficient conversion between surface models and volumetric data, as well as a faster data interpolation process. Before being fed into the network, the entire pelvis is rigidly aligned with a predefined template to standardize its initial pose, and is mirror-flipped to ensure data consistency when the right side is the dislocated part to be moved. The network output consists of six parameters T defining the rigid transformation for the reduction of ϕ_M, including three for translation and three for Euler angles. The rotational center is defined as the volume center.

The model incorporates ResNet-18, chosen for its light-weight architecture and widespread validation in image processing tasks (He et al., 2016). Network hyper-parameters, especially the number of convolutional filters, are carefully chosen based on performance on the validation set. The mean squared error (MSE) between the prediction and the ground-truth pose parameters is used as the training loss.

3.4.2. Recursive pose estimation

The trained pose estimation network is applied in a recursive cascade fashion, with each iteration gradually refining the pose estimate towards the final reduction state. Initially, the fixed and moving side binary masks ϕ_F and ϕ_{M0} are fed to the network, which predicts an initial reduction transformation T_1. Then, T_1 is applied to the moving side to form ϕ_{M1}. Subsequently, ϕ_F and ϕ_{M1} are concatenated to form the input for the next iteration. This process continues until the norm of the transformation parameters for T_n is less than ϵ, or until reaching the maximum iteration. Upon convergence on the nth iteration, the overall pose estimate for the dislocated hipbone is calculated by combining the transformation matrix from each iteration: $T_{all} = T_n \cdot T_{n-1} \cdot \cdots \cdot T_1$.

To ensure data fidelity and prevent error accumulation from iterative interpolation, the estimated pose parameters are applied to the original full-resolution images and then downsampled to the network input resolution for subsequent iterations. While the network, with low-resolution inputs, is kept lightweight for fast inference during application.

3.4.3. Training data simulation

One major challenge in training the pose estimation network is data availability. First, while large-scale pelvic CT datasets are publicly accessible, there remains a notable shortfall of fractured pelvic data necessary to underpin the training of robust models (Liu et al., 2021). Second, generating the ground truth on these clinical data requires accurate segmentation and manipulation of the bone fragments, a process that necessitates specialized expertise and is labor-intensive. Finally, as there is no well-established standard for the planning process, manually manipulating the fragments can be subjective and prone to significant intra- and inter-observer variations.

In light of this challenge, we propose a sim-to-real approach that employs existing data of healthy pelvis to represent the ideal morphology of the pelvic structure and introduce artificial dislocations to the data for model training. We utilized the publicly available *COLONOG* dataset as the source of pelvic CT data, with the pelvic bone segmentation labels provided by *CTPelvic1K* (Johnson et al., 2008; Liu et al., 2021). After excluding scans that featured incomplete pelvic bones, we ultimately selected a cohort of 450 data samples for our study.

In order to enhance the robustness of the network and prevent over-fitting, the dataset is enriched by the following augmentation techniques:

- Random scaling of the pelvis in three dimensions to introduce variation in pelvic size and appearance.
- Random removal of specific regions within the pelvis to account for the potential absence of minor fragments in real situations, thereby eliminating the influence of less relevant morphological features.
- Random translation and rotation of the moving side, with the parameters drawn from a centered normal distribution with a standard deviation of 2 mm and 2° to introduce variation in the initial location and pose of the displaced hipbone.

4. Experiment and result

4.1. Data

We developed PENGWIN, a dataset comprising 150 CT scans of various pelvic trauma cases: (a) pelvic dislocation, (b) unilateral iliac bone fracture, (c) bilateral iliac bone fracture, (d) iliac bone fracture, and (e) iliac and sacral bone fracture. These scans were retrospectively collected from patients who underwent pelvic reduction surgery at Beijing Jishuitan Hospital between 2017 and 2023.

Our study adhered to the guidelines of the Declaration of Helsinki and received approval from the Ethics Committee of Beijing Jishuitan Hospital (202009-04). The CT scans were obtained using different scanner types, including Toshiba Aquilion, United Imaging uCT 550 and uCT 780, and Philips Brilliance models, providing a range of imaging

Table 1
Quantitative results on fracture segmentation. Results (mean ± std) are presented as segmentation accuracy for the hipbone main fragment (H-main), hipbone minor fragment (H-minor), sacral main fragment (S-main), sacral minor fragment (S-minor), and the overall performance. Distance metrics are reported in millimeters.

Method	Metric	Hip-main	Hip-minor	Sacral-main	Sacral-minor	All
Max-Flow Wang et al. (2019)	DSC	0.946 ± 0.053*	0.272 ± 0.396*	0.929 ± 0.150*	0.235 ± 0.320*	0.765 ± 0.251*
	LDSC	0.765 ± 0.087*	0.145 ± 0.298*	0.723 ± 0.110*	0.166 ± 0.256*	0.453 ± 0.180*
	HD_{95}	16.782 ± 20.792*	126.800 ± 100.140*	5.807 ± 9.317*	71.083 ± 59.986*	34.606 ± 50.730*
	ASSD	1.372 ± 1.898*	27.049 ± 21.791*	0.691 ± 1.156*	14.081 ± 10.993*	7.132 ± 10.091*
Swin-UNETR Hatamizadeh et al. (2022)	DSC	0.944 ± 0.016*	0.812 ± 0.193*	0.939 ± 0.023*	0.854 ± 0.097*	0.911 ± 0.077*
	LDSC	0.463 ± 0.156*	0.565 ± 0.194*	0.543 ± 0.096*	0.725 ± 0.069*	0.541 ± 0.132*
	HD_{95}	4.156 ± 8.274*	18.850 ± 20.617*	1.699 ± 1.863*	10.581 ± 13.652	3.195 ± 6.632*
	ASSD	0.557 ± 0.445*	2.880 ± 4.403*	0.392 ± 0.192*	1.482 ± 2.106*	1.035 ± 1.681*
UNet Isensee et al. (2021)	DSC	0.994 ± 0.019	0.944 ± 0.151*	0.990 ± 0.031	0.940 ± 0.094	0.981 ± 0.062*
	LDSC	0.946 ± 0.052*	0.917 ± 0.088*	0.925 ± 0.034*	0.918 ± 0.040	0.929 ± 0.063*
	HD_{95}	2.484 ± 11.517	7.275 ± 23.304*	1.209 ± 4.494	7.367 ± 13.929	1.887 ± 7.776*
	ASSD	0.176 ± 0.798	0.959 ± 2.657	0.093 ± 0.312	0.929 ± 2.053	0.331 ± 1.159*
FDM-UNet	DSC	**0.997 ± 0.010**	0.955 ± 0.133	0.990 ± 0.034	0.938 ± 0.100	0.984 ± 0.054
	LDSC	0.957 ± 0.043	0.930 ± 0.083	0.921 ± 0.035*	0.913 ± 0.043	0.938 ± 0.058
	HD_{95}	1.091 ± 5.277	5.722 ± 15.094	0.986 ± 3.489	7.415 ± 13.786	**1.154 ± 5.481**
	ASSD	0.077 ± 0.305	1.139 ± 3.503	0.080 ± 0.259	1.132 ± 2.206*	0.346 ± 1.290*
FDMSS-UNet	DSC	0.996 ± 0.014	**0.957 ± 0.146**	**0.993 ± 0.014**	**0.955 ± 0.056**	**0.986 ± 0.055**
	LDSC	**0.958 ± 0.042**	**0.932 ± 0.077**	**0.929 ± 0.036**	**0.920 ± 0.050**	**0.940 ± 0.056**
	HD_{95}	**1.202 ± 5.883**	**3.595 ± 12.403**	**0.548 ± 0.898**	**6.325 ± 12.351**	1.262 ± 5.811
	ASSD	**0.082 ± 0.366**	**0.719 ± 2.445**	**0.054 ± 0.101**	**0.764 ± 1.759**	**0.234 ± 0.981**

* Statistically significant differences $p < 0.05$ compared to the last method.

characteristics. The mean voxel spacing is 0.82 × 0.82 × 0.94 mm, and the mean image dimensions are 480 × 397 × 310. The dataset includes patients aged 16 to 94 years, with a gender distribution of 63 females and 87 males, representing a typical demographic affected by pelvic fractures.

We divided the dataset into training and test sets using stratified sampling in a 4:1 ratio. All experiments were conducted on the 30 cases in the test set. Three experts with over five years of experience in pelvic surgery planning performed the segmentation and reduction annotations, which were then cross-checked for accuracy. This dataset is publicly available at https://zenodo.org/records/10927452.

4.2. Implementation

Anatomical segmentation network. All images were resampled to (0.83, 0.83, 0.89) mm voxel spacing using B-spline interpolation and then standardized with Z-score normalization. To increase dataset variability, each training sample was augmented four times. Augmentation involved random elastic distortions (80% to 120%), translations (−20 to 20 mm), and rotations (−30 to 30 degrees) on each axis.

Fracture segmentation network. The resampled bone volumes were cropped with bounding boxes and normalized using z-scores. Each training sample was augmented to create eight new images through mirror flipping along three axes, and random distortions, translations, and rotations. We employed the ADAM optimizer with a learning rate of 0.0001 and a batch size of 2. The parameters λ_{back}, λ_{dice}, λ_{ce}, and λ_{FDM} were set to 0.2, 1, 1, and 16, respectively. A five-fold cross-validation was conducted, with each model trained for 2000 epochs.

Morphable models construction. We constructed separate GPMMs for hipbone and sacrum. The hipbone model used 7520 point cloud samples, while the sacrum model used 6549 point cloud samples. Both models had a parameter dimension of 52. In the model-based fragment registration method, the bidirectional correspondence threshold ϵ was set to 5.

Pose estimation network training. The simulated dataset was partitioned into a training set of 400 samples and a validation set of 50 samples before augmentation. The network was trained for 40 epochs using ADAM optimizer with a learning rate of 10^{-3} and batch size of 8. The stopping threshold ϵ for recursive prediction was set at 0.5.

Platform. The experiments were performed on a system with an Intel i9-12900KS CPU and an NVIDIA RTX 3080Ti GPU. All learning-based methods was implemented using PyTorch 1.12 and SimpleITK. The source code is available at https://github.com/YzzLiu/FracSegNet.

4.3. Experiments on fracture segmentation

The proposed method for fracture segmentation was tested on CT scans to segment each pelvic fracture fragment. We assessed the accuracy of the segmentation using four key metrics: the Dice similarity coefficient (DSC), the 95th percentile of Hausdorff distance (HD_{95}), the average symmetric surface distance (ASSD), and the local Dice similarity coefficient (LDSC), where LDSC was calculated by measuring the Dice coefficient within a 10 mm range near the CFS.

To evaluate our method, we compared the proposed FDMSS-UNet against a graph-cut based max-flow method and the state-of-the-art deep learning model, Swin-UNETR. These methods were implemented as described in their original publications. Additionally, ablation experiments were conducted to compare the performance of FDMSS-UNet with a version without smooth transition and deep supervision (FDM-UNet) and another without distance weighting (UNet).

The quantitative results are shown in Table 1 and Fig. 4. The results indicate that deep learning methods, particularly FDMSS-UNet, were more effective in identifying fracture fragments compared to the max-flow approach, especially for minor fragments. FDMSS-UNet achieved the highest performance across most metrics, with DSC and LDSC values of 0.986 ± 0.055 and 0.940 ± 0.056, respectively, and an ASSD of (0.234 ± 0.981) mm, significantly outperforming the other methods. Main fragments showed better results than minor fragments, partially due to their larger volume.

The ablation experiments highlighted the impact of FDM and training strategies on improving prediction accuracy in the CFS region. While FDM-UNet performed well in terms of DSC and LDSC in certain areas and achieved the best HD95 results, it compromised the global performance of ASSD compared to FDMSS-UNet. The inclusion of deep supervision and smooth transition strategies in FDMSS-UNet contributed to more stable training, yielding balanced local and global metrics. All the improvements were statistically significant, as confirmed by paired t-tests with $p < 0.05$.

Figs. 5 and 6 shows qualitative comparisons in 3D and 2D views. The max-flow method generated reasonable segmentation only when the CFS was clear and mostly non-contact. Although Swin-UNETR and

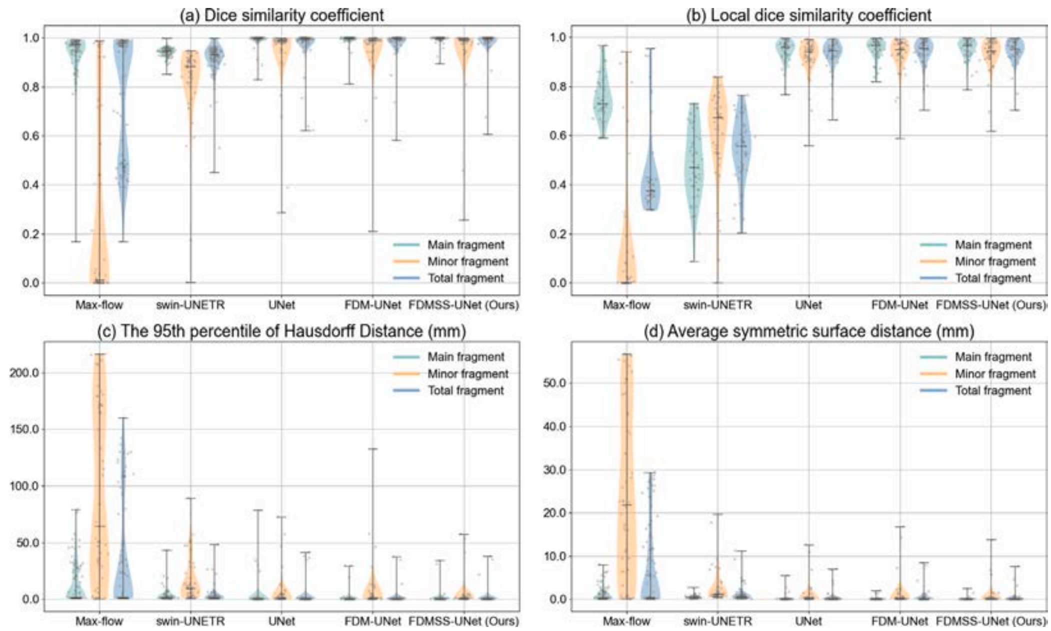

Fig. 4. Distribution of segmentation errors with various methods in main fragments, minor fragments and all fragments of each fractured bone.

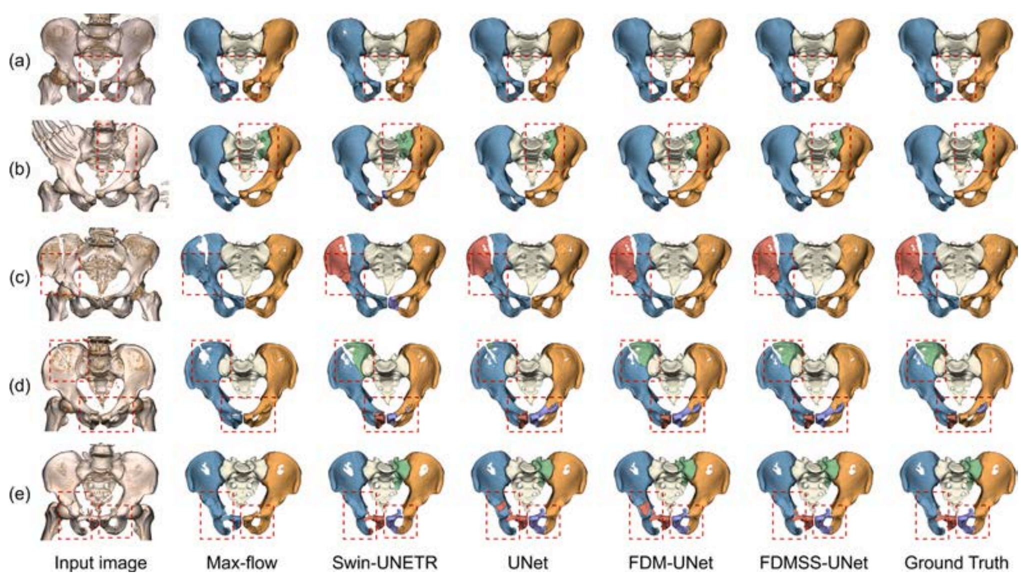

Fig. 5. Example fracture segmentation results in 3D view. (a) Pelvic ring dislocation. (b) Unilateral hip fracture. (c) Bilateral hip fractures. (d) Sacral fracture. (e) Combined sacral and hip fractures. Red boxes indicate the primary fracture areas.

UNet could identify fragments, they often struggled with the complex CFS region, resulting in inaccurate fracture surface predictions. The incorporation of FDM enhanced segmentation accuracy in the CFS region but made the network overly sensitive to fracture features, sometimes neglecting overall morphological characteristics. FDMSS-UNet, by integrating smooth transition and deep supervision strategies, improved accuracy in the CFS region while preserving overall morphology, making it the most effective model in our tests.

4.4. Experiments on fracture reduction planning

The proposed reduction planning method was applied to estimate the transformations of bone fragments to achieve proper reduction in each case. The accuracy of the reduction planning was evaluated using three error metrics: target reduction error (TRE), rotation error, and translation error. TRE is defined as the root-mean-square error (RMSE) of point-wise distances between the planned location and the ground truth, calculated across all surface points. The rotation error is derived from the axis-angle difference between the planned and ground truth pose matrices for each fragment, specifically the rotation angle between the two pose matrices. The translation error is determined by the Euclidean distance between the planned and ground-truth translation vectors for each fragment. The origin of coordinates for reporting the translation and rotation errors is defined as the geometric center of each fragment.

We implemented and compared the following benchmark methods: (1) Mean template (Ead et al., 2021), (2) contralateral mirroring template (Zhang et al., 2016), and (3) optimization-based pelvic articular surface matching algorithm (AS Matching) (Yibulayimu et al., 2023). These methods were implemented in accordance with the specifications and procedures outlined in the original publications. Our proposed fracture reduction planning method is abbreviated as RPEM. Additionally,

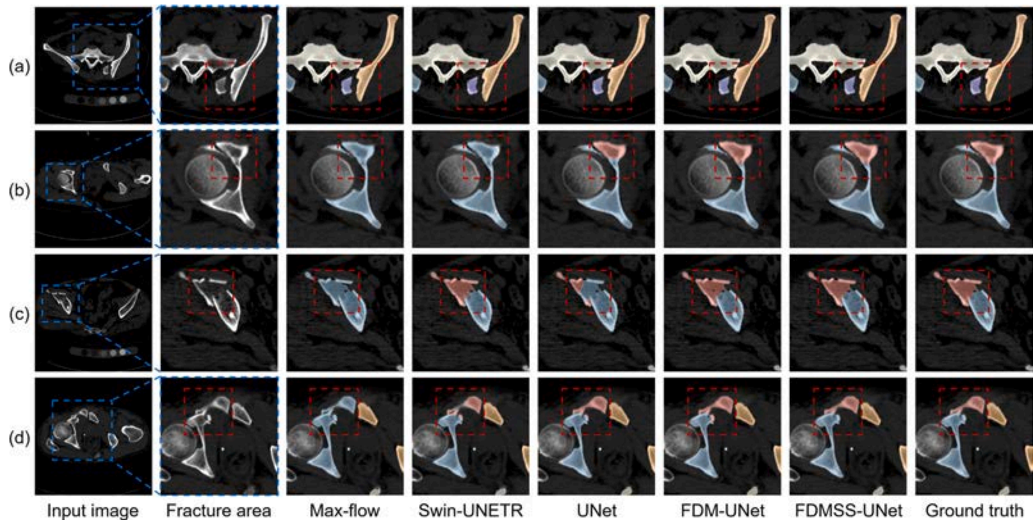

Fig. 6. Example fracture segmentation results displayed on 2D axial slices. (a) Isolated and displaced fragments. (b) Isolated but stable fragments. (c) Compressed and colliding fragments. (d) Partially separated fragments. Red boxes indicate the primary fracture areas.

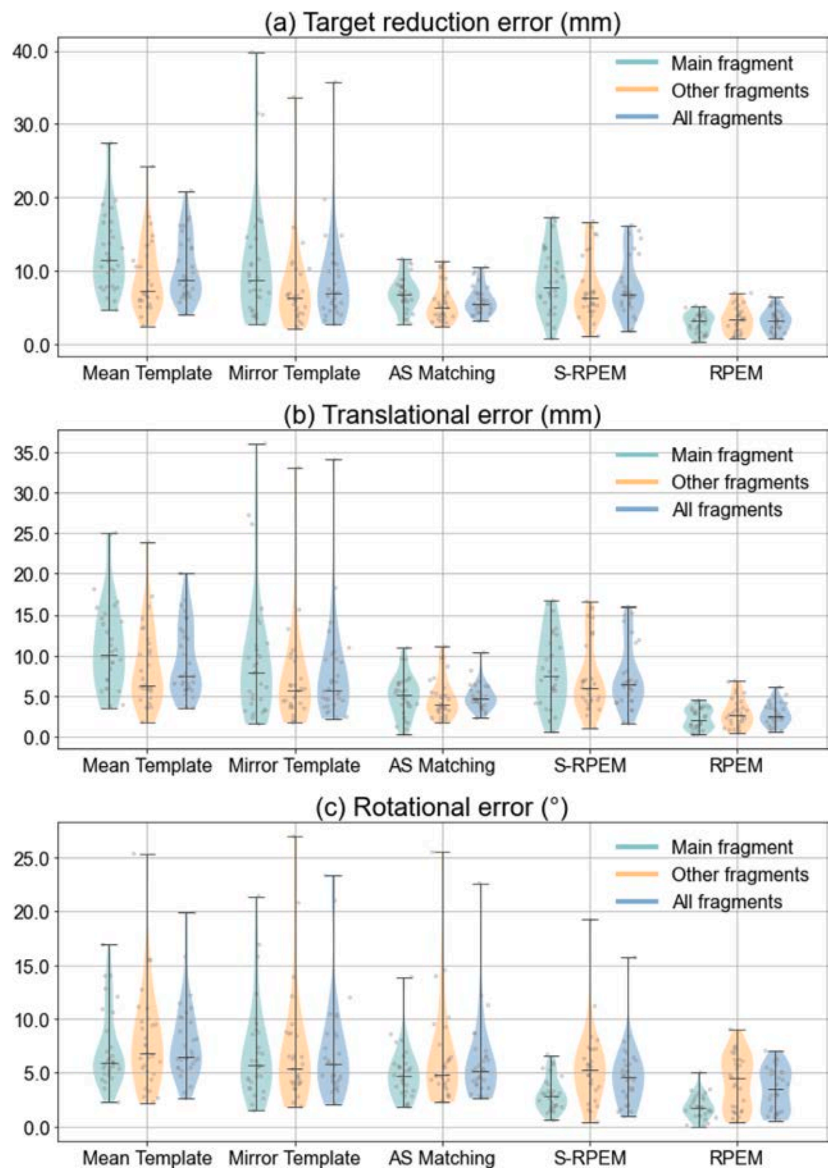

Fig. 7. Distribution of the reduction planning errors for different methods across main, other, and all fragments.

Table 2
Quantitative comparisons of fracture reduction accuracy. Results are represented by target reduction error (mm), translational error (mm), and rotational error (°) to the ground truth. The best values are shown in bold.

Fragment type	Metric	Mean template (Ead et al., 2021)	Mirror template (Zhang et al., 2016)	AS Matching	S-RPEM	RPEM
Main fragment	TRE	12.725 ± 5.754*	11.253 ± 8.892*	6.860 ± 2.265*	8.593 ± 4.353*	**2.900 ± 1.362**
	Rotation	7.266 ± 3.582*	6.947 ± 4.587*	4.946 ± 2.475*	3.013 ± 1.565*	**1.833 ± 1.114**
	Translation	11.186 ± 5.533*	9.540 ± 7.964*	5.013 ± 2.585*	8.072 ± 4.498*	**2.236 ± 1.273**
Other fragments	TRE	9.257 ± 4.877*	7.599 ± 6.039*	5.434 ± 2.468*	7.561 ± 4.237*	**3.356 ± 1.838**
	Rotation	8.068 ± 4.837*	7.113 ± 5.502*	6.485 ± 4.791*	5.496 ± 3.568*	**4.126 ± 2.518**
	Translation	8.652 ± 5.013*	7.069 ± 6.054*	4.844 ± 2.467*	7.187 ± 4.294*	**2.967 ± 1.819**
All fragments	TRE	10.304 ± 4.310*	8.765 ± 6.431*	5.984 ± 1.882*	7.778 ± 3.957*	**3.265 ± 1.485**
	Rotation	7.855 ± 3.695*	7.092 ± 4.796*	6.008 ± 3.925*	4.747 ± 2.880*	**3.476 ± 1.995**
	Translation	9.342 ± 4.426*	7.818 ± 6.220*	4.950 ± 1.653*	7.338 ± 4.021*	**2.773 ± 1.390**

* Statistically significant differences $p < 0.05$ compared to the last method.

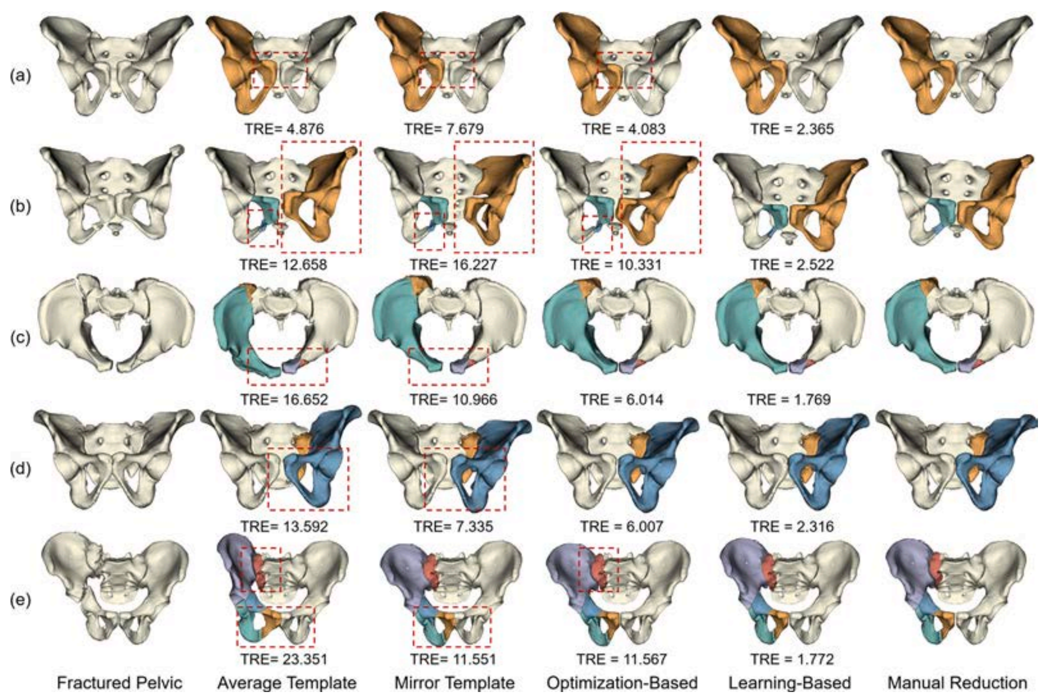

Fig. 8. Example reduction planning results. Each row represents a different type of fracture. TRE values are reported in millimeters. Misalignments are indicted with red boxes.

ablation experiments were performed to test the single pose estimation network without the iterative component (S-RPEM).

Table 2 and Fig. 7 show the quantitative reduction performance in clinical cases, categorizing bone fragments into major and other fragments. The major fragment refers to the primary fragment that is targeted for reduction during fracture surgery, typically the largest fragment of the hipbone. Our method achieved the smallest reduction error across all metrics and fragment types, showing statistically significant improvement over other methods (all paired t-tests showed $p < 0.05$). The TRE for all fragments was (3.265 ± 1.485) mm, with rotation and translation errors of (3.476 ± 1.995)° and (2.773 ± 1.390) mm, respectively, which meets the clinical requirements (Smith et al., 2005; Shillito et al., 2014). The ablation results for S-RPEM demonstrate the advantages of the recursive prediction strategy.

A qualitative comparison in Fig. 8 highlights the limitations of other methods. Due to individual variations, the mean reference often failed to address most fracture reduction problems, leading to uneven fragments and distorted bone posture. Relying solely on sacrum symmetry, the mirror template method frequently resulted in bone overlap or tilted reduction poses, particularly ineffective for sacral and bilateral fractures. In straightforward cases without bone fracture or deformation, results from RPEM and AS Matching were both excellent. However, for complex pelvic trauma, other methods' accuracy declined significantly. In contrast, our method consistently performed satisfactorily, even in scenarios beyond the training set, such as segmentation errors or slightly distorted fixed sides.

4.5. Impact of segmentation on reduction planning

We evaluated the impact of the proposed segmentation method on the reduction planning task by comparing the TRE from three segmentation methods: FDMSS-UNet, FDMSS-UNet with manual correction, and manual segmentation. Table 3 summarizes the results. The success rate is defined as the proportion of cases where segmentation results can be directly used for automatic reduction planning without further manual modification.

The results show that using FDMSS-UNet alone achieved a success rate of 86.67% (26/30) in reduction planning. For the remaining 13.34% (4/30) of cases that did not meet accuracy requirements, manual correction was necessary. The correction step took 1.5 min on average and significantly improved the planning results, with the average TRE improved from 5.057 mm to 3.234 mm. Fig. 9 shows two typical cases that requires manual modifications. Although planning results can be obtained without correcting these fragment labels, inaccuracy in segmentation can lead to incorrect matches during planning. This can even affect the subsequent joint alignment step, as shown in

Table 3
Comparison of different segmentation methods in terms of success rate, planning error, and time.

Segmentation method	Success rate (%)	Planning error (mm)		Time (s)
		Not success	Total	
Algorithm segmentation	86.67	5.057 ± 2.306*	3.557 ± 1.654	32.576 ± 12.563*
Algorithm segmentation + Manual correction	100	3.234 ± 1.273	3.231 ± 1.305	80.443 ± 69.655*
Manual segmentation	100	3.283 ± 1.466	3.265 ± 1.485	1929.931 ± 1106.468

* Statistically significant differences $p < 0.05$ compared to the last method.

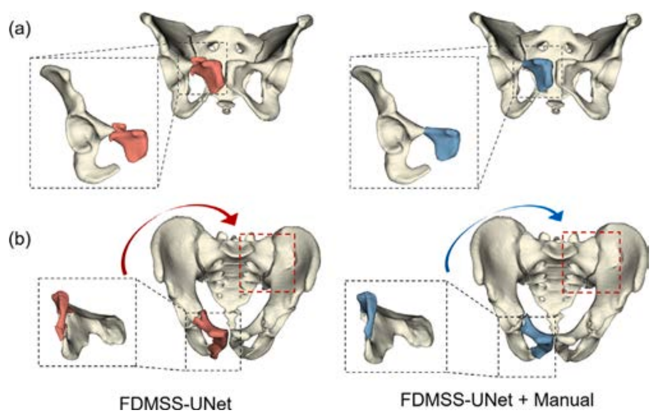

Fig. 9. Impact of segmentation on reduction planning accuracy. (a) and (b) show two typical cases where segmentation leads to increased planning error. Red regions indicate results from FDMSS-UNet, while blue regions indicate results from FDMSS-UNet with manual adjustments.

Fig. 9(b).

While the accuracy of FDMSS-UNet with manual correction matches that of manual segmentation, the latter is significantly more time-consuming. Therefore, combining FDMSS-UNet with manual correction not only maintains high accuracy but also significantly improves processing efficiency.

4.6. Impact of fragment size

We evaluated the impact of fragment size on the accuracy of our segmentation and planning methods. The results are shown in Fig. 10. The clinical fracture dataset includes fragments with size ranging from 1 cm^3 to 500 cm^3. The experimental results indicate no significant correlation between fragment size and overall segmentation accuracy. Specifically, there is a slight positive correlation between DSC and fragment size, while HD_{95} and ASSD exhibit a slight negative correlation with fragment size.

For reduction planning, the TRE, translation error, and rotation error all show a decreasing trend with increasing fragment size, indicating more robust performance for larger fragments. Among these metrics, the rotation error is more pronounced in smaller fragments compared to the translation error.

Fig. 11 shows two cases where the algorithm fails to manage small fragments precisely. In the repositioning of fracture fragments, errors in dense correspondences significantly influence the planning accuracy. Extracting reliable geometric features from small fragments is more challenging, making it difficult to find correct template matching points. Comparing case (a) and case (b), a similar phenomenon is also observed in manual planning, that small, indistinct fragments are typically hard to plan manually. Nevertheless, in the context of pelvic fracture reduction surgery, such small fragments are generally not important for the overall reduction of the pelvic ring and are usually ignored, thus making these errors tolerable.

4.7. Cadaver experiment

We examined the proposed preoperative surgical planning pipeline in a proof-of-concept cadaver experiment, where the ground truth target morphology of the pelvis is available form the CT scan acquired before injuries are induced.

As shown in Fig. 12, a Tile C1 pelvic ring fracture was imparted in the left hipbone of the cadaver, resulting in three fragments. One of the fragments was isolated and dislocated, and was not connected to the SI joint or PS. The proposed fracture segmentation network achieved highly accurate results, with DSC of 0.988, LDSC of 0.964, HD_{95} of 0.794 mm, and ASSD of 0.164 mm, which aligns with the results in the previous experiments. The fracture reduction planning result is shown in Fig. 12(d). Reduction planning accuracy was calculated with respect to the ground-truth positions in the CT scan acquired prior to the injury. The proposed reduction planning method achieved TRE of 3.731 mm, and translation and rotation errors of 2.532 mm and 4.244°, respectively. These difference are deemed acceptable as the result was within the target accuracy of 5 mm for pelvic ring fractures (Smith et al., 2005; Shillito et al., 2014).

Furthermore, our method significantly outperforms manual operations in efficiency. For a typical fracture case, the total time for segmentation and planning using our method is (267.53 ± 168.29) seconds, compared to approximate 120 min in manual planning.

5. Discussion and conclusion

In this study, we have introduced an innovative automated preoperative planning solution for pelvic fracture reduction surgery. Our approach integrates a two-stage deep learning-based automatic segmentation method for pelvic fractures, and a novel recursive pose estimation module combined with statistical shape modeling to restore a healthy pelvic morphology. This comprehensive pipeline effectively tackles critical challenges in pelvic fracture reduction planning. By integrating fracture distance weighted loss in to the fracture segmentation network, the accuracy of segmentation, particularly near fracture regions in significantly enhanced, and the planning process is streamlined. The recursive pose estimation module, combined with the morphable model-based template method, effectively restored the healthy pelvic structure, addressing the challenge of limited fractured data and providing a reliable solution for complex fracture types.

Our approach significantly improves both the accuracy and efficiency of preoperative surgical planning process. Experiments on retrospective clinical data across all types of pelvic fractures demonstrated the method's versatility and robustness. The proof-of-concept cadaver experiments further confirmed the system's superior efficiency and precision.

Despite the promising results, our study has limitations. The automated segmentation method, while highly accurate, may still require some manual corrections in cases with severe deformities or low image quality. Specifically, as shown in Fig. 13, the segmentation algorithm struggled with anterior ring that were significantly distorted but only slightly separated, which could lead to suboptimal alignment in reduction planning. However, simple adjustments, such as separate the fragment into two distinct parts, significantly improved the reduction outcome, highlighting the remaining need for clinical decision when

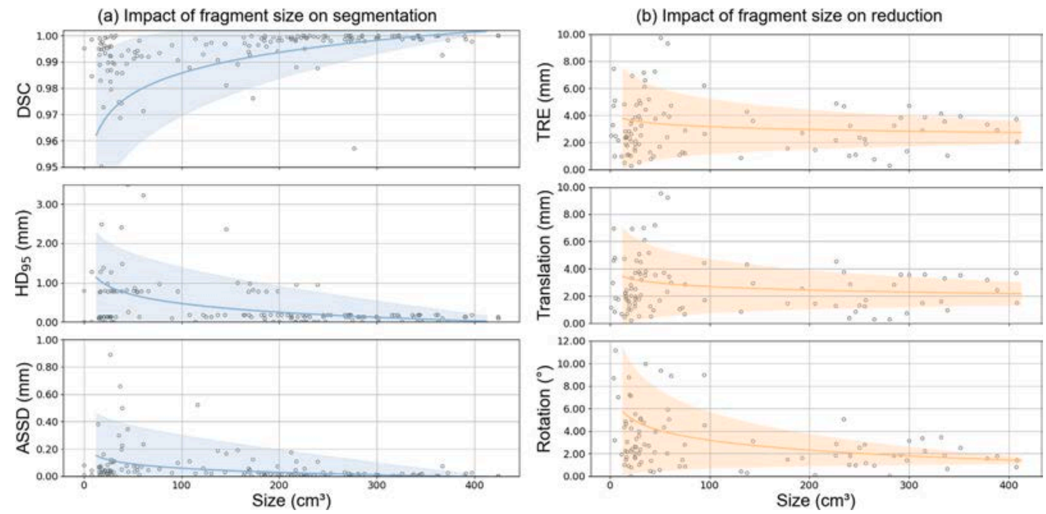

Fig. 10. Impact of fragment size to segmentation and reduction. (a) DSC, HD_{95}, and ASSD on segmentation result. (b) TRE, translational error, and rotational error on reduction planning.

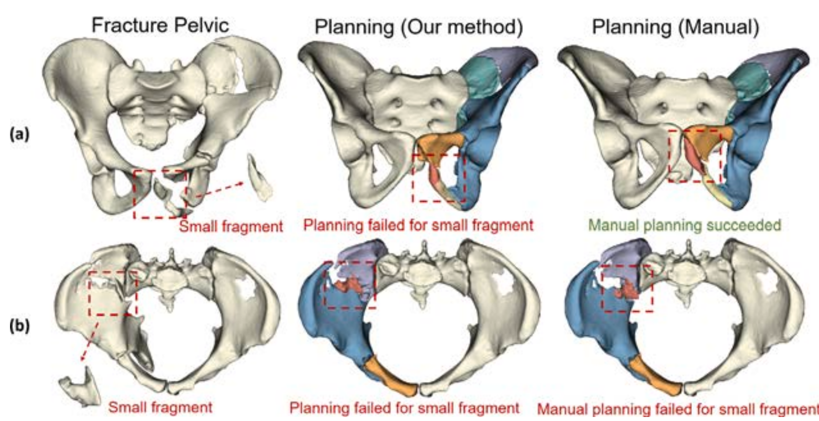

Fig. 11. Reduction planning results involving small fragments. (a) Reduction of small fragments where the algorithm fails but manual correction is possible. (b) Fragments lacking geometric features where both the algorithm and manual planning fail.

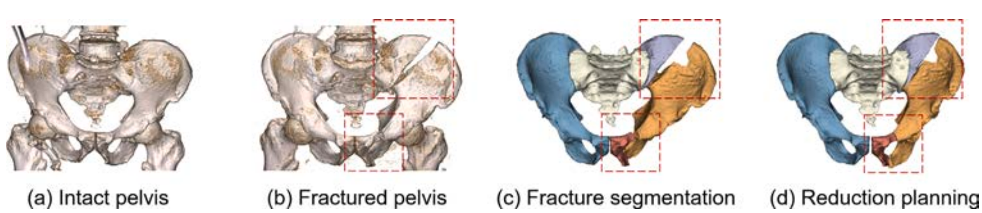

Fig. 12. Experimental results of preoperative fracture reduction planning on cadaveric pelvic fractures. (a) Initial 3D reconstruction of the cadaveric pelvis showing pre-injury morphology; (b) Induced Tile C1 pelvic ring fracture; (c) Automated segmentation using FDMSS-UNet; (d) Automated reduction planning using RPEM.

planning on such cases. Additionally, the recursive pose estimation module faced challenges in managing extremely small or irregular fragments, implying the need for further refinement of the module to solve point-wise correspondence and better handle such complexities.

In future work, we plan to focus on broader validation of our method across a more extensive spectrum of clinical cases, including diverse cohorts and fracture types, involving larger sample sizes and various imaging modalities to ensure the approach's generalizability and robustness. Furthermore, we aim to explore an instance segmentation framework capable of handling a variable number of fragment labels. This approach could provide a more detailed and precise representation of fracture cases. In order to further enhance the computational efficiency and ensure seamless adoption and utilization in surgical settings, we plan to develop a deep learning-based

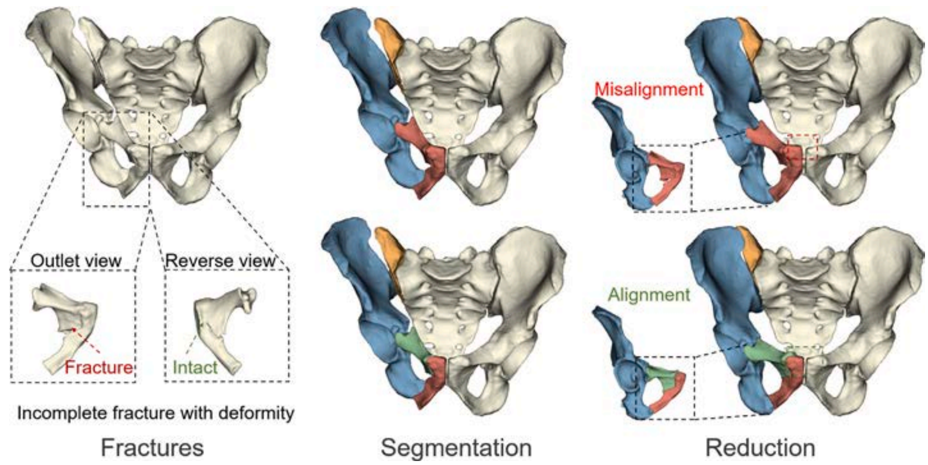

Fig. 13. Challenges faced by the segmentation algorithm in separating fragments from specific fractures that are not completely separated but have undergone significant distortion.

single-bone fracture reduction module to expedite the planning process, currently constrained by traditional statistical shape modeling methods. In addition, we plan to integrate the proposed planning solution with surgical robot systems to achieve automatic closed reduction of pelvic fractures (Ge et al., 2022; Liu et al., 2024b).

CRediT authorship contribution statement

Yanzhen Liu: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Conceptualization, Validation, Visualization. **Sutuke Yibulayimu:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft. **Yudi Sang:** Conceptualization, Data curation, Methodology, Project administration, Software, Supervision, Writing – review & editing, Formal analysis, Validation. **Gang Zhu:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Validation. **Chao Shi:** Formal analysis, Investigation, Software, Validation, Visualization, Writing – review & editing. **Chendi Liang:** Conceptualization, Formal analysis, Investigation, Software, Visualization, Writing – review & editing. **Qiyong Cao:** Data curation, Formal analysis, Funding acquisition, Resources, Validation. **Chunpeng Zhao:** Data curation, Formal analysis, Investigation, Resources, Validation. **Xinbao Wu:** Data curation, Funding acquisition, Investigation, Resources, Supervision, Validation. **Yu Wang:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

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.

Data availability

The data are available on https://zenodo.org/records/10927452 and https://github.com/Sutuk/Clinical-data-on-pelvic-fractures. And the code is available at https://github.com/YzzLiu/FracSegNet.

References

Abdi, A.H., Pesteie, M., Prisman, E., Abolmaesumi, P., Fels, S., 2019. Variational shape completion for virtual planning of jaw reconstructive surgery. In: Medical Image Computing and Computer Assisted Intervention: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part V 22. Springer, pp. 227–235. http://dx.doi.org/10.1007/978-3-030-32254-0_26.

Albrecht, T., Lüthi, M., Gerig, T., Vetter, T., 2013. Posterior shape models. Med. Image Anal. 17 (8), 959–973. http://dx.doi.org/10.1016/j.media.2013.05.010.

Audenaert, E.A., Van Houcke, J., Almeida, D.F., Paelinck, L., Peiffer, M., Steenackers, G., Vandermeulen, D., 2019. Cascaded statistical shape model based segmentation of the full lower limb in CT. Comput. Methods Biomech. Biomed. Eng. 22 (6), 644–657. http://dx.doi.org/10.1080/10255842.2019.1577828.

Cheng, P., Yang, Y., Yu, H., He, Y., 2021. Automatic vertebrae localization and segmentation in CT with a two-stage dense-U-Net. Sci. Rep. 11 (1), 22156. http://dx.doi.org/10.1038/s41598-021-01296-1.

Chowdhury, A.S., Bhandarkar, S.M., Robinson, R.W., Jack, C.Y., 2009. Virtual multi-fracture craniofacial reconstruction using computer vision and graph matching. Comput. Med. Imaging Graph. 33 (5), 333–342. http://dx.doi.org/10.1016/j.compmedimag.2009.01.006.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., et al., 2013. The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging 26, 1045–1057. http://dx.doi.org/10.1007/s10278-013-9622-7.

Deng, Z., Jiang, J., Huang, R., Zhang, W., Chen, Z., He, K., Yao, Q., 2023. Synergistically segmenting and reducing fracture bones via whole-to-whole deep dense matching. Comput. Graph..

Ead, M.S., Palizi, M., Jaremko, J.L., Westover, L., Duke, K.K., 2021. Development and application of the average pelvic shape in virtual pelvic fracture reconstruction. Int. J. Med. Robot. 17 (2), e2199. http://dx.doi.org/10.1002/rcs.2199.

Ead, M.S., Westover, L., Polege, S., McClelland, S., Jaremko, J.L., Duke, K.K., 2020. Virtual reconstruction of unilateral pelvic fractures by using pelvic symmetry. Int. J. Comput. Assist. Radiol. Surg. 15, 1267–1277. http://dx.doi.org/10.1007/s11548-020-02140-z.

Ge, Y., Zhao, C., Wang, Y., Wu, X., 2022. Robot-assisted autonomous reduction of a displaced pelvic fracture: a case report and brief literature review. J. Clin. Med. 11 (6), 1598. http://dx.doi.org/10.3390/jcm11061598.

Ghimire, K., Chen, Q., Feng, X., 2022. Head and neck tumor segmentation with deeply-supervised 3D unet and progression-free survival prediction with linear model. In: Head and Neck Tumor Segmentation and Outcome Prediction: Second Challenge, HECKTOR 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings. Springer, pp. 141–149. http://dx.doi.org/10.1007/978-3-030-98253-9_13.

Gnat, R., Saulicz, E., Biały, M., Kłaptocz, P., 2009. Does pelvic asymmetry always mean pathology? Analysis of mechanical factors leading to the asymmetry. J. Hum. Kinet. 21 (2009), 23–32. http://dx.doi.org/10.2478/v10078-09-0003-8.

Halawi, M.J., 2016. Pelvic ring injuries: Surgical management and long-term outcomes. J. Clin. Orthop. Trauma 7 (1), 1–6. http://dx.doi.org/10.1016/j.jcot.2015.08.001.

Han, R., Uneri, A., Ketcha, M., Vijayan, R., Sheth, N., Wu, P., Vagdargi, P., Vogt, S., Kleinszig, G., Osgood, G.M., Siewerdsen, J.H., 2020. Multi-body 3D-2D registration for image-guided reduction of pelvic dislocation in orthopaedic trauma surgery. Phys. Med. Biol. 65 (13), 135009. http://dx.doi.org/10.1088/1361-6560/ab843c.

Han, R., Uneri, A., Silva, T.D., Ketcha, M., Siewerdsen, J.H., 2019. Atlas-based automatic planning and 3D-2D fluoroscopic guidance in pelvic trauma surgery. Phys. Med. Biol. 64 (9), http://dx.doi.org/10.1088/1361-6560/ab1456.

Han, R., Uneri, A., Vijayan, R.C., Wu, P., Vagdargi, P., Sheth, N., Vogt, S., Kleinszig, G., Osgood, G.M., Siewerdsen, J.H., 2021. Fracture reduction planning and guidance in orthopaedic trauma surgery via multi-body image registration. Med. Image Anal. 68, 101917. http://dx.doi.org/10.1016/j.media.2020.101917.

Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D., 2022. Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. In: Crimi, A., Bakas, S. (Eds.), Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Springer International Publishing, Cham, pp. 272–284. http://dx.doi.org/10.1007/978-3-031-08999-2_22.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 770–778. http://dx.doi.org/10.1109/CVPR.2016.90.

Hermans, E., Biert, J., Edwards, M.J.R., 2017. Epidemiology of pelvic ring fractures in a level 1 trauma center in the netherlands. Hip Pelvis 29 (4), 253–261. http://dx.doi.org/10.5371/hp.2017.29.4.253.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 18 (2), 203–211. http://dx.doi.org/10.1038/s41592-020-01008-z.

Johnson, C.D., Chen, M.-H., Toledano, A.Y., Heiken, J.P., Dachman, A., Kuo, M.D., Menias, C.O., Siewert, B., Cheema, J.I., Obregon, R.G., et al., 2008. Accuracy of CT colonography for detection of large adenomas and cancers. N. Engl. J. Med. 359 (12), 1207–1217. http://dx.doi.org/10.1056/NEJMoa0800996.

Khurana, B., Sheehan, S.E., Sodickson, A.D., Weaver, M.J., 2014. Pelvic ring fractures: what the orthopedic surgeon wants to know. Radiogr. 34 (5), 1317–1333. http://dx.doi.org/10.1148/rg.345135113.

Kim, H., Jeon, Y.D., Park, K.B., Cha, H., Kim, M.-S., You, J., Lee, S.-W., Shin, S.-H., Chung, Y.-G., Kang, S.B., et al., 2023. Automatic segmentation of inconstant fractured fragments for tibia/fibula from CT images using deep learning. Sci. Rep. 13 (1), 20431. http://dx.doi.org/10.1038/s41598-023-47706-4.

Kodym, O., Španěl, M., Herout, A., 2020. Skull shape reconstruction using cascaded convolutional networks. Comput. Biol. Med. 123, 103886. http://dx.doi.org/10.1016/j.compbiomed.2020.103886.

Krčah, M., Székely, G., Blanc, R., 2011. 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. IEEE, pp. 2087–2090. http://dx.doi.org/10.1109/ISBI.2011.5872823.

Kuiper, R.J.A., Sakkers, R.J.B., van Stralen, M., Arbabi, V., Viergever, M.A., Weinans, H., Seevinck, P.R., 2022. Efficient cascaded V-net optimization for lower extremity CT segmentation validated using bone morphology assessment. J. Orthop. Res. 40 (12), 2894–2907. http://dx.doi.org/10.1002/jor.25314.

Liu, P., Han, H., Du, Y., Zhu, H., Li, Y., Gu, F., Xiao, H., Li, J., Zhao, C., Xiao, L., et al., 2021. Deep learning to segment pelvic bones: large-scale CT datasets and baseline models. Int. J. Comput. Assist. Radiol. Surg. 16, 749–756. http://dx.doi.org/10.1007/s11548-021-02363-8.

Liu, J., Li, H., Zeng, B., Wang, H., Kikinis, R., Joskowicz, L., Chen, X., 2024a. An end-to-end geometry-based pipeline for automatic preoperative surgical planning of pelvic fracture reduction and fixation. IEEE Trans. Med. Imaging http://dx.doi.org/10.1109/TMI.2024.3429403.

Liu, Y., Wu, X., Sang, Y., Zhao, C., Wang, Y., Shi, B., Fan, Y., 2024b. Evolution of surgical robot systems enhanced by artificial intelligence: A review. Adv. Intell. Syst. 6 (5), 2300268. http://dx.doi.org/10.1002/aisy.202300268.

Liu, J., Xing, F., Shaikh, A., Linguraru, M.G., Porras, A.R., 2022. Learning with context encoding for single-stage cranial bone labeling and landmark localization. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII. Springer, pp. 286–296. http://dx.doi.org/10.1007/978-3-031-16452-1_28.

Liu, Y., Yibulayimu, S., Sang, Y., Zhu, G., Wang, Y., Zhao, C., Wu, X., 2023. Pelvic fracture segmentation using a multi-scale distance-weighted neural network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 312–321. http://dx.doi.org/10.1007/978-3-031-43996-4_30.

Lu, S., Yang, Y., Li, S., Zhang, L., Shi, B., Zhang, D., Li, B., Hu, Y., 2023. Preoperative virtual reduction planning algorithm of fractured pelvis based on adaptive templates. IEEE Trans. Biomed. Eng. 70 (10), 2943–2954. http://dx.doi.org/10.1109/TBME.2023.3272007.

Luque-Luque, A., Pérez-Cano, F.D., Jiménez-Delgado, J.J., 2021. Complex fracture reduction by exact identification of the fracture zone. Med. Image Anal. 72, 102120. http://dx.doi.org/10.1016/j.media.2021.102120.

Luthi, M., Gerig, T., Jud, C., Vetter, T., 2018. Gaussian process morphable models. IEEE Trans. Pattern Anal. Mach. Intell. 40 (8), 1860–1873. http://dx.doi.org/10.1109/tpami.2017.2739743.

Miller, P.R., Moore, P.S., Mansell, E., Meredith, J.W., Chang, M.C., 2003. External fixation or arteriogram in bleeding pelvic fracture: initial therapy guided by markers of arterial hemorrhage. J. Trauma Acute Care Surg. 54 (3), 437–443. http://dx.doi.org/10.1097/01.TA.0000053397.33827.DD.

Neubauer, A., Bühler, K., Wegenkittl, R., Rauchberger, A., Rieger, M., 2005. Advanced virtual corrective osteotomy. In: International Congress Series. Vol. 1281, Elsevier, pp. 684–689. http://dx.doi.org/10.1016/j.ics.2005.03.254.

Olson, S.A., Furman, B., Guilak, F., 2012. Joint injury and post-traumatic arthritis. http://dx.doi.org/10.1007/s11420-011-9247-7.

Paulano, F., Jiménez, J.J., Pulido, R., 2014. 3D segmentation and labeling of fractured bone from CT images. Vis. Comput. 30, 939–948. http://dx.doi.org/10.1007/s00371-014-0963-0.

Paulano-Godino, F., Jiménez-Delgado, J.J., 2017. Identification of fracture zones and its application in automatic bone fracture reduction. Comput. Methods Programs Biomed. 141, 93–104. http://dx.doi.org/10.1016/j.cmpb.2016.12.014.

Qamar, S., Jin, H., Zheng, R., Ahmad, P., Usama, M., 2020. A variant form of 3D-UNet for infant brain segmentation. Future Gener. Comput. Syst. 108, 613–623. http://dx.doi.org/10.1016/j.future.2019.11.021.

Ruikar, D.D., Santosh, K., Hegadi, R.S., 2019. Segmentation and analysis of CT images for bone fracture detection and labeling. In: Medical Imaging. CRC Press, pp. 130–154. http://dx.doi.org/10.1201/9780429029417-7.

Seim, H., Kainmueller, D., Heller, M., Lamecker, H., Zachow, S., Hege, H.-C., 2008. Automatic segmentation of the pelvic bones from CT data based on a statistical shape model. VCBM 8, 93–100. http://dx.doi.org/10.2312/VCBM/VCBM08/093-100.

Shillito, M., Linn, M., Girard, P., Schwartz, A., 2014. Anterior sacroiliac dislocation: a case report and review of the literature. JBJS Case Connect. 4 (3), e78. http://dx.doi.org/10.2106/JBJS.CC.M.00269.

Smith, W., Shurnas, P., Morgan, S., Agudelo, J., Luszko, G., Knox, E.C., Georgopoulos, G., 2005. Clinical outcomes of unstable pelvic fractures in skeletally immature patients. JBJS 87, http://dx.doi.org/10.2106/JBJS.C.01244v.

Suero, E.M., Hüfner, T., Stübig, T., Krettek, C., Citak, M., 2010. Use of a virtual 3D software for planning of tibial plateau fracture reconstruction. Inj. 41 (6), 589–591. http://dx.doi.org/10.1016/j.injury.2009.10.053.

Tomazevic, M., Kreuh, D., Kristan, A., Puketa, V., Cimerman, M., 2010. Preoperative planning program tool in treatment of articular fractures: process of segmentation procedure. In: XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010: May 27–30, 2010 Chalkidiki, Greece. Springer, pp. 430–433. http://dx.doi.org/10.1007/978-3-642-13039-7_108.

Tosounidis, T., Kanakaris, N., Nikolaou, V., Tan, B., Giannoudis, P.V., 2012. Assessment of lateral compression type 1 pelvic ring injuries by intraoperative manipulation: which fracture pattern is unstable? Int. Orthop. 36, 2553–2558. http://dx.doi.org/10.1007/s00264-012-1685-4.

Wang, D., Wu, Z., Fan, G., Liu, H., Liao, X., Chen, Y., Zhang, H., 2022. Accuracy and reliability analysis of a machine learning based segmentation tool for intertrochanteric femoral fracture CT. Front. Surg. 9, 913385. http://dx.doi.org/10.3389/fsurg.2022.913385.

Wang, D., Yu, K., Feng, C., Zhao, D., Min, X., Li, W., 2019. 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. http://dx.doi.org/10.1145/3364836.3364837.

Wang, J., Zhang, X., Guo, L., Shi, C., Tamura, S., 2023. Multi-scale attention and deep supervision-based 3D UNet for automatic liver segmentation from CT. Math. Biosci. Eng.: MBE 20 (1), 1297–1316. http://dx.doi.org/10.3934/mbe.2023059.

Willis, A.R., Anderson, D.D., Thomas, T.P., Brown, T.D., Marsh, J.L., 2007. 3D reconstruction of highly fragmented bone fractures. In: SPIE Medical Imaging. http://dx.doi.org/10.1117/12.708683.

Yang, L., Gao, S., Li, P., Shi, J., Zhou, F., 2022. Recognition and segmentation of individual bone fragments with a deep learning approach in ct scans of complex intertrochanteric fractures: A retrospective study. J. Digit. Imaging 35 (6), 1681–1689. http://dx.doi.org/10.1007/s10278-022-00669-w.

Yang, J., Gu, S., Wei, D., Pfister, H., Ni, B., 2021. Ribseg dataset and strong point cloud baselines for rib segmentation from ct scans. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. Springer, pp. 611–621. http://dx.doi.org/10.1007/978-3-030-87193-2_58.

Yibulayimu, S., Liu, Y., Sang, Y., Zhu, G., Wang, Y., Liu, J., Shi, C., Zhao, C., Wu, X., 2023. Pelvic fracture reduction planning based on morphable models and structural constraints. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 322–332. http://dx.doi.org/10.1007/978-3-031-43996-4_31.

Yuan, J., Bae, E., Tai, X.-C., Boykov, Y., 2014. A spatially continuous max-flow and min-cut framework for binary labeling problems. Numer. Math. 126, 559–587. http://dx.doi.org/10.1007/s00211-013-0569-x.

Zeng, B., Wang, H., Joskowicz, L., Chen, X., 2024a. Fragment distance-guided dual-stream learning for automatic pelvic fracture segmentation. Comput. Med. Imaging Graph. 116, 102412. http://dx.doi.org/10.1016/j.compmedimag.2024.102412.

Zeng, B., Wang, H., Tao, X., Shi, H., Joskowicz, L., Chen, X., 2024b. A bidirectional framework for fracture simulation and deformation-based restoration prediction in pelvic fracture surgical planning. Med. Image Anal. 97, 103267. http://dx.doi.org/10.1016/j.media.2024.103267.

Zhang, L.-h., Zhao, J.-x., Zhao, Z., Su, X.-y., Zhang, L.-c., Zhao, Y.-p., Tang, P.-f., 2016. Computer-aided pelvic reduction frame for anatomical closed reduction of unstable pelvic fractures. J. Orthop. Res. 34 (1), 81–87. http://dx.doi.org/10.1002/jor.22987.

Zhao, C., Guan, M., Shi, C., Zhu, G., Gao, X., Zhao, X., Wang, Y., Wu, X., 2022. Automatic reduction planning of pelvic fracture based on symmetry. Comput. Methods Biomech. Biomed. Eng.: Imaging Vis. 10 (6), 577–584. http://dx.doi.org/10.1080/21681163.2021.2012830.