BJR
© 2023 The Authors. Published by the British Institute of Radiology
https://doi.org/10.1259/bjr.20221006

Received: 27 October 2022 | Revised: 14 March 2023 | Accepted: 15 March 2023 | Published online: 17 April 2023

Cite this article as:
Li N, Wu Z, Jiang C, Sun L, Li B, Guo J, et al. An automatic fresh rib fracture detection and positioning system using deep learning. *Br J Radiol* (2023) 10.1259/bjr.20221006.

FULL PAPER

An automatic fresh rib fracture detection and positioning system using deep learning

[1]NING LI, MD, [1]ZHE WU, MD, [1]CHAO JIANG, MD, [1]LULU SUN, MD, [1]BINGYAO LI, MD, [1]JUN GUO, MD, [2]FENG LIU, PhD, [2]ZHEN ZHOU, PhD, [1]HAIBO QIN, MD, [2]WEIXIONG TAN, MS and [1]LUFENG TIAN, MD

[1]Department of Radiology, Fushun Central Hospital of Liaoning Province, Fushun, Liaoning Province, China
[2]Deepwise Artificial Intelligence (AI) Lab, Deepwise Inc., Beijing, China

Address correspondence to:
Weixiong Tan
E-mail: *tanweixiong@deepwise.com*
Lufeng Tian
E-mail: *tianlufeng22@163.com*

The authors Ning Li and Zhe Wu contributed equally to the work.

Objective: To evaluate the performance and robustness of a deep learning-based automatic fresh rib fracture detection and positioning system (FRF-DPS).
Methods: CT scans of 18,172 participants admitted to eight hospitals from June 2009 to March 2019 were retrospectively collected. Patients were divided into development set (14,241), multicenter internal test set (1612), and external test set (2319). In internal test set, sensitivity, false positives (FPs) and specificity were used to assess fresh rib fracture detection performance at the lesion- and examination-levels. In external test set, the performance of detecting fresh rib fractures by radiologist and FRF-DPS were evaluated at lesion, rib, and examination levels. Additionally, the accuracy of FRF-DPS in rib positioning was investigated by the ground-truth labeling.
Results: In multicenter internal test set, FRF-DPS showed excellent performance at the lesion- (sensitivity: 0.933 [95%CI, 0.916–0.949], FPs: 0.50 [95%CI, 0.397–0.583]) and examination-level. In external test set, the sensitivity and FPs at the lesion-level of FRF-DPS (0.909 [95%CI, 0.883–0.926], $p < 0.001$; 0.379 [95%CI, 0.303–0.422], $p = 0.001$) were better than the radiologist (0.789 [95%CI, 0.766–0.807]; 0.496 [95%CI, 0.383–0.571]), so were the rib- and patient-levels. In subgroup analysis of CT parameters, FRF-DPS were robust (0.894–0.927). Finally, FRF-DPS(0.997 [95%CI, 0.992–1.000], $p < 0.001$) is more accurate than radiologist (0.981 [95%CI, 0.969–0.996]) in rib positioning and takes 20 times less time.
Conclusion: FRF-DPS achieved high detection rate of fresh rib fractures with low FP values, and precise positioning of ribs, thus can be used in clinical practice to improve the detection rate and work efficiency.
Advances in knowledge: We developed the FRF-DPS system which can detect fresh rib fractures and rib position, and evaluated by a large amount of multicenter data.

INTRODUCTION

Rib fracture is the most common type of chest injury in patients with blunt trauma, and approximately 40–50% of patients with blunt chest trauma have rib fractures.[1] Reportedly, up to 12% of patients with traumatic rib fractures die from acute or subacute injury, and nearly 50% of patients eventually need intensive care and surgery.[2] Some studies have shown that 81% of patients with traumatic rib fractures have hemothorax, pneumothorax, and other organ injuries that are mainly related to specific rib fractures.[3] Identifying the degree of displacement and the number of rib fractures is the basis for developing treatment plans.[4] As the main imaging method to evaluate chest trauma, CT provides a detailed evaluation of rib fractures and multiple traumas of a patient with chest and abdominal injuries.[5] Influenced by the shape of the rib, CT imaging equipment, and angle, radiologists might misdiagnose or miss the diagnosis of rib fractures in CT. Furthermore, CT scans should be evaluated sequentially when the degree of displacement and the number of rib fractures are calculated. These are meticulous and time-consuming tasks.

Presently, artificial intelligence (AI), the core technique mainly represented by deep neural networks or known as deep learning (DL), has great potential in computer vision and medical images.[6,7] It also has some applications in the

field of rib fracture detection.[8–10] Nonetheless, the majority of previous studies have not been verified by human–machine comparison and large-scale data, and hence the clinical value cannot be confirmed. Zhou et al[11] proposed a model based on DL to automatically detect and classify the rib fractures into three categories (fresh, healing, and old). Yang et al[12] utilized the DL model to diagnose and classify rib fractures, which is useful to doctors in improving diagnosis efficiency and reducing workload. However, both studies used small test data sets to evaluate the performance of the respective models; moreover, the first study does not indicate the anatomical location of each rib fracture.

Thus, this study aimed to develop DL-based automatic fresh rib fracture detection and positioning in an emergency scenario, which was validated using a multicenter data set.

METHODS AND MATERIALS

Ethics approval

This multicenter retrospective study was approved by the ethics committee of each hospital, and the requirement for informed consent was waived due to the retrospective design of this study.

Patients

The inclusion criteria were as follows: (1) patients admitted to the Emergency department because of chest trauma; (2) CT imaging of all ribs was performed bilaterally, and images in digital imaging and communications in medicine (DICOM) format were collected after examination; (3) patients aged >18-years-old. The exclusion criteria were as follows: (1) patients with image artifacts or inability to meet the diagnostic requirements due to poor breath-holding, (2) patients with pathological rib fractures, (3) patients with a fracture history of >3 weeks.

From June 2009 to March 2019, CT scans of 18,903 participants admitted to the different centers (A, B, C, D, E, F, G, and H) in China and who met the inclusion criteria were collected retrospectively. Scans of patients (n = 784) who met the exclusion criteria were removed, finally leaving a data set of 18,172 CT scans. The data of the seven centers (A, B, C, D, E, F, and G) were randomly divided into training set (n = 12,636), validation set (n = 1605) and internal test set (n = 1612) according to 8:1:1 patient ratio. The training and validation sets were collectively referred to as the development set. The data of center H comprised the external test set (n = 2,319). The internal test set was used to evaluate the detection performance of fresh rib fracture detection and positioning system (FRF-DPS) on multicenter data. The external test set was used to evaluate the performance of the FRF-DPS in the detection and recognition of fresh rib fractures and that of the rib positioning algorithm. Figure 1 and Table 1 summarize the data distribution and patient characteristics, respectively.

CT examinations

CT examinations were obtained by helical CT scans using the 64-slice CT (LightSpeed VCT, Sensation Medical Systems, Philips Medical Systems), the 128-slice CT (LightSpeed VCT, Neuviz Medical Systems), and the uCT960+Medical Systems. In addition, CT examinations were captured under voltage 120 kVp, current 250–350 mA, and multiple convolution kernel reconstruction algorithms. The reconstruction thicknesses ranged from 0.625 to 1.25 mm. Table 1 lists the CT parameters.

Ground-truth (GT) labeling

The GT was determined by three senior physicians with experience of at least 10 years (WZ, LL, and JC). Specifically, each CT scan was performed individually by each physician, labeling the bounding boxes of fresh rib fractures, as well as the corresponding rib number and position. The inconsistent annotations were resolved by discussion among the three physicians to reach the final decision. Furthermore, each rib, irrespective of the fresh rib fractures, was labeled by its position.

Figure 1. Flowchart of patient inclusion.

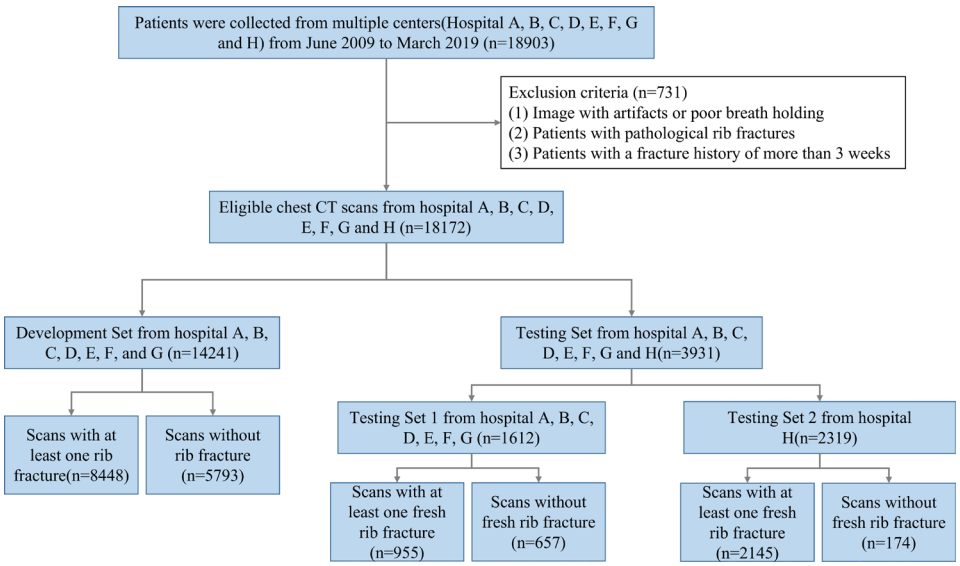
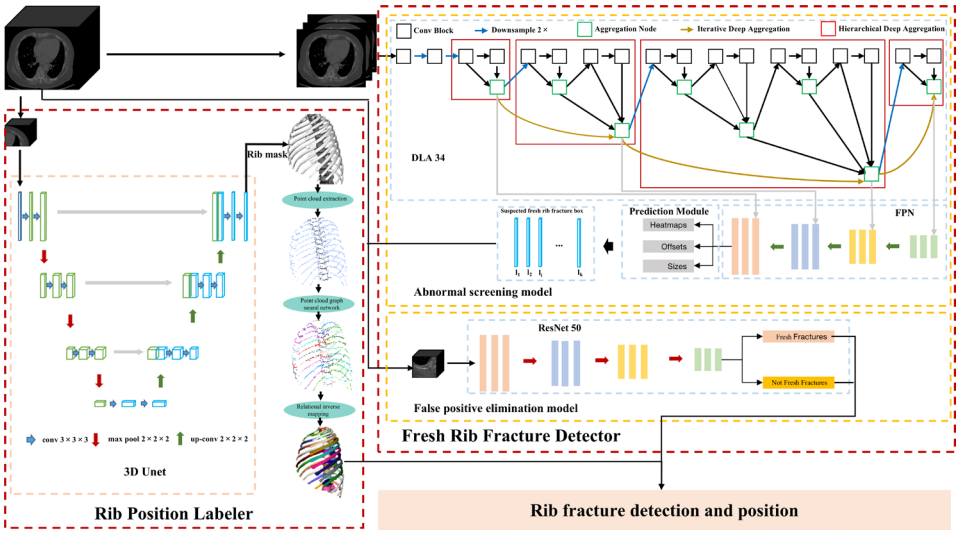


Table 1. Patients' demographic and CT parameters

Characteristic	Development (n = 14241)	Test 1 (n = 1612)	Test 2 (n = 2319)
Clinical information			
Age in years, mean ± SD	57.94 ± 14.36	58.33 ± 14.07	58.15 ± 14.12
SEX			
Male	7929 (55.7%)	902 (55.9%)	1459 (62.9%)
Female	6313 (44.3%)	710 (44.1%)	860 (37.1%)
CT parameters			
Manufacturer			
GE	2779	288	0
SIEMENS	6907	791	1,724
Philips	3313	410	502
Other	1242	123	93
Kernel			
B70f	5110	573	1,281
B60f'	4742	519	781
Other	3936	520	257
Slice thickness			
0.625 mm	4540	502	1,282
1.00 mm	9578	1,032	1,037
Other	123	78	0
CTDI	41.89 ± 13.93	43.02 ± 14.24	42.17 ± 14.13

CTDI, CT dose index.; SD, standard deviation.
Mean data are ±standard deviation; data in parentheses are percentages;

Figure 2. Pipeline of the FRF-DPS. The FRF-DPS consists of Rib Fracture Detector, which detects rib fractures and distinguishes between fresh and old rib fractures, and Rib Position Labeler, which segments the position of each rib. Rib Fracture Detector consists of two models: the abnormal screening model and false-positive elimination model. The abnormal screening model takes centernet as the framework and DLA as the feature extraction model. 3d, three-dimensional; FRF-DPS, fresh rib fracture detection and positioning system.



FRF-DPS

The FRF-DPS consisted of two modules: a fresh rib fracture detector and a rib position labeler; the pipeline is illustrated in Figure 2. The fresh rib fracture detector detects fresh rib fractures from CT scan images and adopts a cascade-based design,[13,14] wherein an initial object detector first detects the axial images. The detection on adjacent slices is grouped as fracture candidates, followed by a three-dimensional (3D) convolutional neural network[15] to distinguish between the positive (fresh rib fracture) and negative (not fresh rib fracture) candidates. The rib position labeler provides text annotation of the rib number of each rib on every slice. It performs rib segmentation using a customized U-Net model.[16] Each connected component on the segmentation mask is treated as a position candidate for rib number annotation. Then, the position candidate is encoded to a 3D point by concatenating its center co-ordinate and slice number, and all position candidates in the CT form a point cloud. Finally, a graph neural network is used to learn the geometry of different ribs, and then the rib number predictions of each point are presented.[17] For detailed on the fresh rib fracture detector and rib position labeler, see Supplementary Material 1.

The diagnostic process of a radiologist

In order to evaluate the diagnostic performance of the FRF-DPS, a radiologist (R1) with 10 years of experience in thoracic CT evaluated the fresh rib fractures independently. The reader used the labeling platform to read each patient's CT scans independently and framed the lesions of fresh rib fractures slice-by-slice on the platform according to their judgment, as well as the corresponding rib number and rib position. Finally, the lesion box was exported from the platform as fresh rib fractures in each CT image, according to the readers' judgment. Also, exports all the rib numbers and the rib position.

Consistency evaluation

To evaluate a predictor's agreement, GT was divided into lesion-, rib-, and patient-level, the definitions of consistency see Supplementary Material 1. In addition, the "completely accurate" performance metric used in evaluating the degree and number of different fresh rib fractures was defined as follows: when a patient has no FP and FN results for all ribs, the patient is predicted to be correct, otherwise incorrect. Finally, "completely accurate" performance metric was calculated as the number of patients predicted to be correct divided by the total number of patients. The fracture degree was divided into "non-displaced", "slightly displaced" and "severely displaced" categories.[18,19]

Statistical analysis

SPSS 25.0 software (IBM Corp., Armonk, NY) was used to perform statistical analysis. For internal test set, 9874 fresh rib fractures were included in the 1612 chest CT scans. In internal test set, the GT labeling included 13,524 fresh rib fractures and all rib positions in the 2319 chest CT scans. Sensitivity, precision, F1-score, and FP values were calculated for the detection of fresh rib fractures at the lesion level. At the rib- and patient-level, the sensitivity, specificity, precision, and F1-scores were calculated. To verify the robustness of the model, the χ^2 test was used to compare the differences in the detection performance of AI under different CT parameters (such as slice thickness, model, and reconstruction convolution kernel) of fresh rib fractures. Furthermore, in different degrees and numbers of rib fractures, the χ^2 test was used to compare the differences in the detection performance between AI and radiologists to verify the value of the FRF-DPS in clinical practice. $p < 0.05$ indicated a statistically significant difference.

RESULTS

Performance of FRF-DPS in multicenter evaluation

Table 2 lists the statistical results of the FRF-DPS on internal test set. The statistical analysis on fresh rib fractures was performed at the lesion- and examination-levels, respectively. For the lesion-level, the sensitivity was 0.933 [95%CI, 0.916–0.949], and the FPs was 0.5 [95%CI, 0.397–0.583]. The sensitivity was 0.993 [95%CI, 0.984–1.000], and the specificity was 0.934 [95%CI, 0.916–0.949] at the examination-level.

Performance comparison between FRF-DPS and radiologists

For the external test set, the FRF-DPS outperformed the radiologists to detect fresh rib fractures at the lesion-, rib- and examination-level (Table 3). As shown in Figure 3, among the fresh rib fractures detected by FRF-DPS, 1968 cases were not detected by the R1at the lesion-level, 1190 cases were not detected at the rib-level, and 80 cases were not detected at the examination-level. However, among the fresh rib fractures detected by the R1, only 346,235, and 2 were not detected by the FRF-DPS. In the 13,524 fresh rib lesions, the FRF-DPS system detected 1968 new lesions out of lesions that had already been identified by radiologists, with only 177 FPs reported. Of the 10,312 ribs with fresh rib fractures, 1190 new ribs were identified by the FRF-DPS system, with only 79 FPs. Among 2147 examinations with fresh rib fractures, 80 new TP were detected by the FRF-DPS system, with only 6 additional FPs. Figure 4 shows some cases that were detected by both the FRF-DPS and the R1.

Table 2. FRF-DPS evaluates performance in fresh rib fracture detection on the internal test set

	Sensitivity [95% CI]	Specificity [95% CI]	FPs [95% CI]
Lesion-level	0.933 [0.916, 0.949]	–	0.500 [0.397,0.583]
Examination-level	0.993 [0.984, 1.000]	0.934 [0.916, 0.949]	–

CI, confidence interval; FN, false-negative; FPs, false-positive per scan; FRF-DPS, fresh rib fracture detection and positioning system; TN, true-negative; TP, true-positive.
Sensitivity = TP/(TP+FN); Specificity = TN/(TN+FP);

Table 3. Comparison of the detection performance of the FRF-DPS and the R1 model at three levels on the external test set

Lesion-level				
Reader	Sensitivity [95% CI]	Precision [95% CI]	F1-score [95% CI]	FPs [95% CI]
AI	0.909 [0.883, 0.926]	0.937 [0.913, 0.956]	0.923 [0.902, 0.941]	0.379 [0.303, 0.422]
R1	0.789 [0.766, 0.807]	0.902 [0.877, 0.921]	0.842 [0.821, 0.860]	0.496 [0.383, 0.571]
p	<0.001***	0.003**	0.002**	0.001**
Rib-level				
Reader	Sensitivity [95% CI]	Precision [95% CI]	F1-score [95% CI]	Specificity [95% CI]
AI	0.922 [0.904, 0.939]	0.950 [0.932, 0.966]	0.936 [0.911, 0.955]	0.989 [0.981, 1.000]
R1	0.831 [0.814, 0.848]	0.917 [0.901, 0.939]	0.872 [0.854, 0.889]	0.983 [0.977, 0.996]
p	<0.001***	<0.001***	0.002**	0.019*
Examination-level				
Reader	Sensitivity [95% CI]	Precision [95% CI]	F1-score [95% CI]	Specificity [95% CI]
AI	0.989 [0.982, 1.000]	0.986 [0.979, 0.998]	0.987 [0.980, 1.000]	0.822 [0.801, 0.843]
R1	0.953 [0.936, 0.969]	0.972 [0.955, 0.989]	0.962 [0.943, 0.978]	0.661 [0.639, 0.689]
p	0.005**	0.029*	0.016*	0.004**

AI, artificial intelligence; CI, confidence interval; FN, false-negative; FPs, false-positive perscan; FRS-DPS, fresh rib fracture detection and positioning system; R1, Radiologist; TN, true-negative; TP, true-positive.
Sensitivity = TP/(TP+FN); Precision = TP/(TP+FP); F1-score = 2*Sensitivity*Precision/(Sensitivity+Precision); Specificity = TN/(TN+FP); *P value < 0.05.

The detection performance of the R1 and the FRF-DPS in different degrees and numbers of rib fractures was analyzed further. As shown in Table 4, for patients with 1–2 fresh rib fractures, the "completely accurate" performance metric values of the FRF-DPS for non-, slight-, and severe-displaced fractures were 11.2–14.2% higher than those of the R1 (all $p < 0.05$). Similarly, for patients with 3–5 fresh rib fractures and those with more than 5 fresh rib fractures, the FRF-DPS outperformed ($p < 0.05$). Specifically, for 769 patients with severe displacement, the sensitivity value of the FRF-DPS was about 94.3%, *i.e.* 114 cases more were detected than R1.

Performance of FRF-DPS for different CT parameters

Next, we evaluated the generalizability of FRF-DPS for the detection of fresh rib fractures by CT kernel, CT dose index (CTDI), gender, age, CT manufacturer, and slice thickness on the external test set, as FRF-DPS could be applicable in diverse clinical fields. FRF-DPS exhibited a robust performance, with sensitivity values 0.894–0.927 and *p*-values >0.05 at the lesion-level (Figure 5). This performance was not affected by medications, imaging settings, or participants' demographic characteristics. The detailed performance metrics for the different subgroups are presented in Supplementary Material 1.

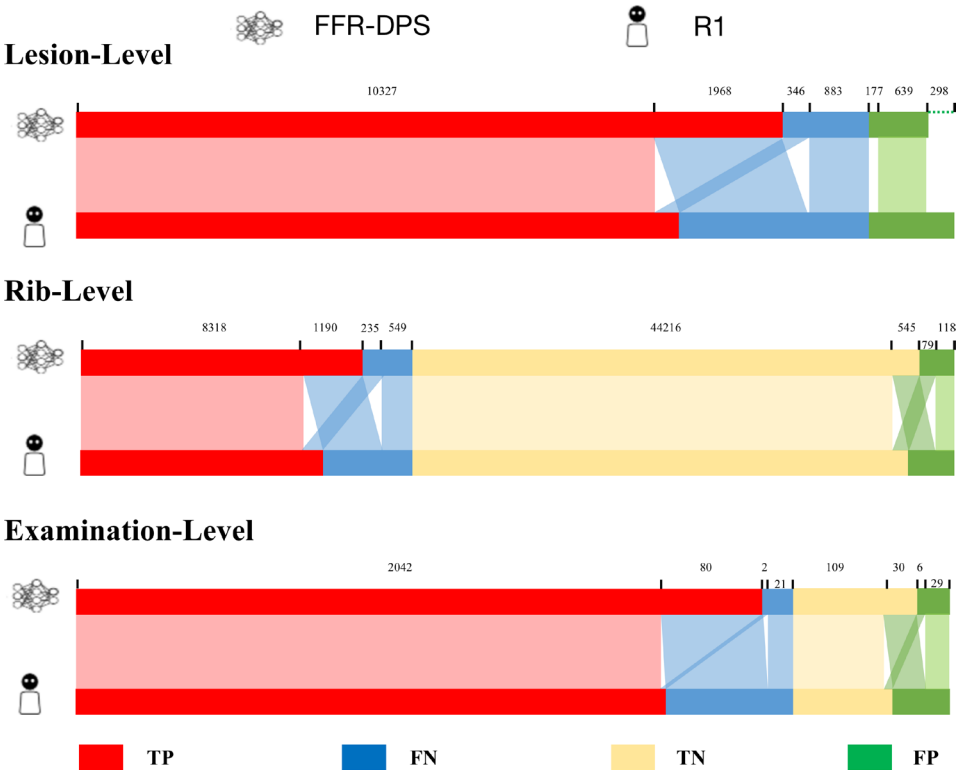
Comparison of rib position performance between FRF-DPS and radiologists

Herein, the rib position performance of the FRF-DPS and R1 was assessed on the external test set (Table 5). Of the 13,524 fresh rib fractures (GT), the FRF-DPS had only 38 fresh rib fractures (accuracy: 0.997 [95%CI, 0.992–1.000]), wherein the rib position was incorrectly labeled, while R1 had 247 fractures. In terms of time consumption, the average duration from the detection of each fresh rib fracture to the determination of its rib position by FRF-DPS was <3 s, while that of R1 was about 60 s. In addition, we also calculated the performance of FRF-DPS and R1 detection of FP in the rib position. Among the 816 FP detected by FRF-DPS, except for 10 FP not detected on the rib, the FRF-DPS only had 3 FP with wrongly labeled rib positions, while R1 had 33 such FP. Among the 1114 FP detected by R1, except for 12 FP not detected on the rib, the FRF-DPS only had 2 FP with wrongly labeled rib positions, while R1 had 17 such FP. The detailed performance metrics for the different subgroups are presented in Supplementary Material 1.

DISCUSSION

Chest CT identifies many injuries that chest X-rays may miss, including lung contusions, hemothorax or pneumothorax, and rib fractures. Therefore, CT examination is often the first choice for patients with thoracic trauma. Rib fractures were considered indicators of severe trauma.[20] Traditional detection methods require careful evaluation of the entire data set of the CT scans in sequence, which is time-consuming and error-prone. Moreover, the number of rib fractures and the degree of dislocation were related to the plan of follow-up treatment, especially the choice of drugs and surgery planning.[21,22] Thus, a definite diagnosis of a fresh rib fracture, the detection of the extent of the fracture, and positioning of the ribs were essential. Especially in emergency scenarios, it was crucial to promptly detect and locate fresh rib fractures, while detecting old fractures was not necessary. Therefore, we developed a deep learning-based FRF-DPS for fresh rib fractures, using 14,241 CT scans for model development, and evaluation of 3931 CT scans. The results showed that FRF-DPS efficiently and quickly detected fresh rib fractures and determined the position of the rib harboring

Figure 3. Comparison of TP, FP, TN, and FN between the FRF-DPS and R1 models at the lesion-, rib-, and patient-level. Among these, in the middle area on each level map, the same color represents cases with FRF-DPS and R1 model. For example, in the lesion-level, in the middle reddish area, 10,327 identical fresh rib fractures were correctly detected by AI and R1. In three light blue areas: 346 identical fresh ribs were correctly detected by R1 but missed by the FRF-DPS; 1968 identical fresh ribs were correctly detected by FRF-DPS but missed by R1; 883 identical fresh ribs were missed. In the light green area, 639 identical sites were falsely detected as fractures. The remaining 177 and 298 are the number of FRF-DPS and R1 independent false detections, respectively. FRF-DPS, fresh rib fracture detection and positioning system; FN, false-negative; FP, false-positive; TN, true-negative; TP, true-positive.



the fracture. In addition, we compared the performance of FRF-DPS with that of radiologists, and the results showed that FRF-DPS outperformed the radiologists.

According to previous reports, some models based on DL model have been used to diagnose, classify, and locate rib fractures.[11,12,23] However, these studies were difficult to verify the performance

Figure 4. Cases detected by FRF-DPS and the R1 model. (a) Detection by both the RF-DP FRF-DPS and R1 model; (b) Detection by only the FRF-DPS; (c) Detection by only the R1 model; (d) No detection. FRF-DPS, fresh rib fracture detection and positioning system.

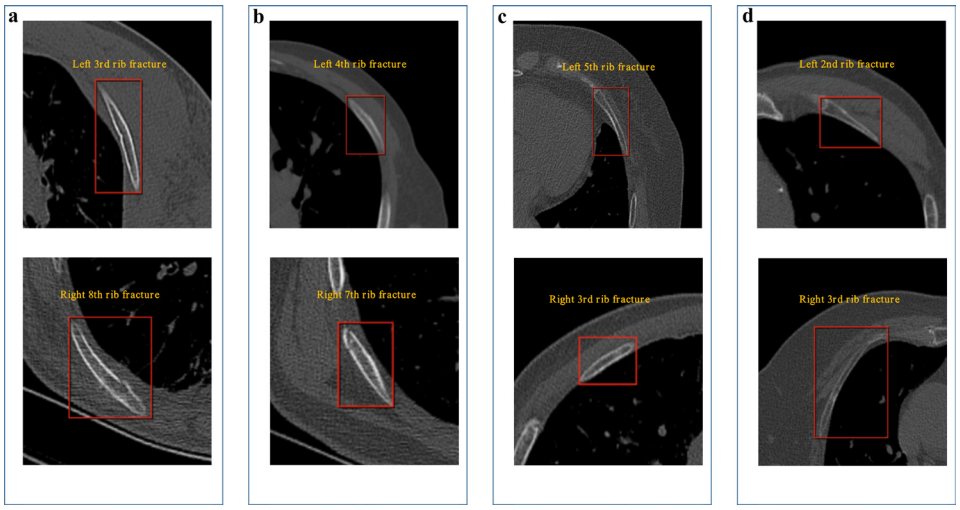


Table 4. Completely accurate statistics of the FRF-DPS in Number of fresh rib fractures per CT and degree of displacement

Number of fresh rib fractures per CT	Degree of displacement	Completely accurate		*p*
		AI [95% CI]	R1 [95% CI]	
1–2	Non-displaced	0.831 (304/366)[0.811, 0.852]	0.689 (252/366) [0.661, 0.725]	0.002**
	Slightly displaced	0.934 (184/197)[0.906, 0.961]	0.822 (162/197) [0.791, 0.850]	0.005**
	Severely displaced	0.960 (72/75)[0.929, 0.992]	0.840 (63/75) [0.807, 0.874]	0.002**
3–5	Non-displaced	0.825 (213/258)[0.802, 0.848]	0.663 (171/258) [0.639, 0.688]	0.001**
	Slightly displaced	0.923 (265/287)[0.902, 0.944]	0.812 (233/287) [0.791, 0.833]	0.004**
	Severely displaced	0.945 (239/253)[0.922, 0.967]	0.837 (212/253) [0.820, 0.856]	0.003**
6–24	Non-displaced	0.807 (67/83)[0.771, 0.846]	0.578 (48/83) [0.539, 0.619]	<0.001***
	Slightly displaced	0.917 (176/192)[0.901, 0.938]	0.682 (131/192) [0.660, 0.701]	<0.001***
	Severely displaced	0.938 (414/441)[0.922, 0.953]	0.762 (336/441) [0.749, 0.770]	0.001**

AI, artificial intelligence; CI, confidence interval; FRS-DPS, fresh rib fracture detection and positioning system; R1, Radiologist.
Data in parentheses are percentages; *p-value < 0.05.

and generalizability of the models due to the small amount of data. In addition, they did not use the data set to objectively assess the performance of rib positioning. Wu et al[24] validated the performance of the DL-based rib detection model using a large amount of data and evaluated the performance of rib segmentation, but the rib location was not clearly defined, and the rib positioning was not evaluated. Besides, the rib fractures detected by this study include fresh fractures and old fractures. Unlike previous studies, our system aimed to quickly detect fresh rib fractures and locate its rib positioning. In the actual emergency clinic, it was necessary to quickly detect fresh rib fracture and locate its rib position, but the old fracture did not be required to be detected because it did not need treatment. Therefore, our study was more in line with the actual clinical needs. We used a large amount of multicenter data to verify the efficiency of our study, fresh rib fractures and positions. From the analysis of the results, it can be concluded that the

Figure 5. Histogram distribution of the sensitivity of FRF-DPS among different subgroups. FRF-DPS, fresh rib fracture detection and positioning system.

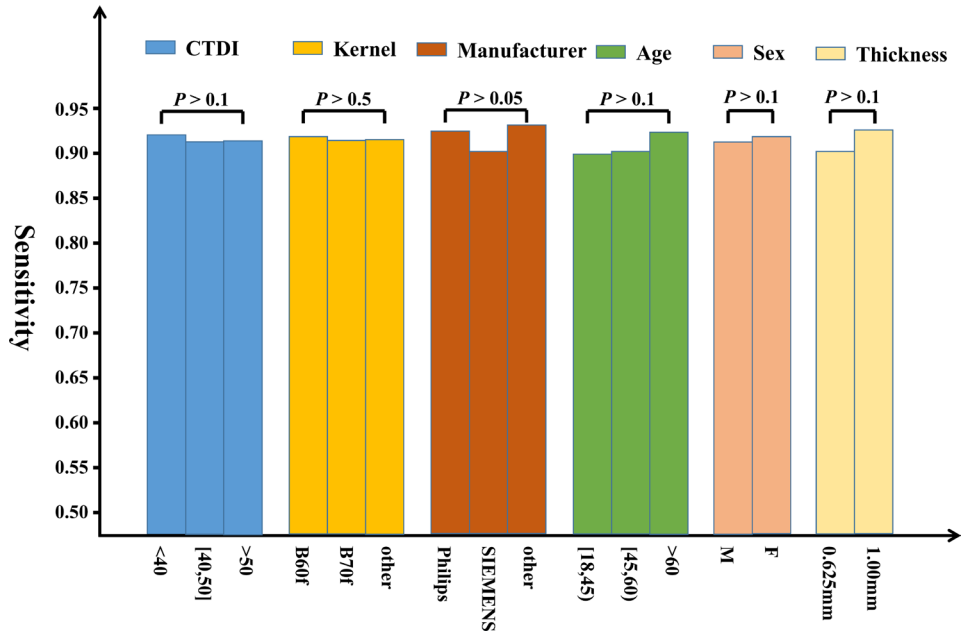


Table 5. Comparison of rib position performance between FRF-DPS and radiologists

	Accuracy			Time(/s) of per lesion	
	AI [95% CI]	R1 [95% CI]	*p*	AI	R1
GT	0.997 [0.992, 1.000]	0.981 [0.969, 0.996]	<0.001*	2.97	60.12
RRF-DPS FP	0.996 [0.990, 1.000]	0.959 [0.944, 0.973]	<0.001*	3.01	61.37
AIFP	0.998 [0.995, 1.000]	0.984 [0.973, 0.998]	<0.001*	2.94	59.84

AI, artificial intelligence; CI, confidence interval; FP, false-positive; FRS-DPS, fresh rib fracture detection and positioning system; GT, ground-truth.

FRF-DPS can assist physicians in improving the rib fracture detection rate and speed without misleading them to add FP cases. In addition, the FRF-DPS can greatly improve the speed and accuracy of rib position positioning.

To date, although several studies have used DL to detect rib fractures with promising sensitivity,[25,26] precise positioning of the ribs was rarely performed; thus, in clinical practice, positioning of the ribs was essential. The FRF-DPS provided 99.7% rib localization while precisely detecting rib fractures. In the current study, the detection performance of the FRF-DPS was further analyzed at different degrees and numbers of rib fractures. Importantly, FRF-DPS achieved a "completely accurate" performance of >91% for different numbers of rib fractures in both mildly displaced and severely dislocated patients. Although the "completely accurate" performance was lower on well-aligned fractures, it was >80% and it was at least 15% more than that of the R1. Specifically, in patients with severe dislocation of more than three rib fractures, the performance of the FRF-DPS could reach 94.1%. We also observed that the FRF-DPS efficiently detects the degree of displacement and the number of rib fractures and could be advantageous for the formulation of treatment plans, which has great significance in clinical practice.

Nevertheless, the present study has several limitations. One of the main improvements is that the performance verification of radiologists with FRF-DPS in detecting rib fractures. Although a senior radiologist was recruited to compare the radiologist's findings with those achieved by the FRF-DPS, no comparison of the detection performance between the FRF-DPS+physician and the FRF-DPS was made, and the number of physicians also needs to be increased. Therefore, the application of the FRF-DPS in clinical practice requires multicenter data and multiple physicians' feedback.

CONCLUSION

In conclusion, FRF-DPS showed promising performance and robustness in the automatic detection of fresh rib fractures and rib localization, indicating its potential in the clinical practice.

AUTHOR CONTRIBUTION:

Conceptualization: Ning Li, Zhe Wu,Weixion Tan, Lufeng Tian
Data curation: Ning Li, Zhe Wu, Lufeng Tian, Chao Jiang
Formal analysis: Weixion Tan, Zhen Zhou
Funding acquisition: Lufeng Tian
Investigation: Zhe Wu, Lulu Sun
Methodology: Weixion Tan, Feng Liu
Project administration: Zhe Wu, Lufeng Tian, Bingyao Li
Resources: Ning Li, Jun Guo
Software: Weixion Tan, Feng Liu, Zhen Zhou
Supervision: Zhe Wu, Lufeng Tian
Validation: Feng Liu
Visualization: Weixion Tan
Writing-original draft: Haibo Qin, Zhen Zhou
Writing-review & editing: Ning Li, Zhe Wu

CONFLICTS OF INTEREST

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

FUNDING

This work was supported in part by the Beijing Municipal Science and Technology Planning Project (Grant Nos. Z201100005620008) and the Natural Science Foundation of Liaoning Province in China (Grant No. 2023-MS-352).

DATA AVAILABILITY STATEMENT

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

REFERENCES

1. Hamilton C, Barnett L, Trop A, Leininger B, Olson A, Brooks A, et al. Emergency department management of patients with rib fracture based on a clinical practice guideline. *Trauma Surg Acute Care Open* 2017; **2**(1): e000133. https://doi.org/10.1136/tsaco-2017-000133
2. Dunham CM, Hileman BM, Ransom KJ, Malik RJ. Trauma patient adverse outcomes are independently associated with rib cage fracture burden and severity of lung, head, and abdominal injuries. *Int J Burns Trauma* 2015; **5**: 46–55.
3. Ivey KM, White C C, Wallum T E. Thoracic injuries in US combat casualties: A 10-year review of operation enduring freedom and iraqi freedom[J]. *J Trauma Acute Care Surg* 2012; 73. https://doi.org/10.1097/TA.0b013e3182754654
4. Ingoe HM, Eardley W, McDaid C, Rangan A, Lawrence T, Hewitt C. Epidemiology of adult rib fracture and factors associated with surgical fixation: analysis of a chest wall

injury dataset from england and wales. *Injury* 2020; **51**: 218–23. https://doi.org/10.1016/j.injury.2019.10.030
5. Huber-Wagner S, Lefering R, Qvick L-M, Körner M, Kay MV, Pfeifer K-J, et al. Effect of whole-body CT during trauma resuscitation on survival: A retrospective, multicentre study. *Lancet* 2009; **373**: 1455–61. https://doi.org/10.1016/S0140-6736(09)60232-4
6. Wang J, Zhu H, Wang S-H, Zhang Y-D. A review of deep learning on medical image analysis. *Mobile Netw Appl* 2021; **26**: 351–80. https://doi.org/10.1007/s11036-020-01672-7
7. Wang M, Wei Z, Jia M, Chen L, Ji H. Deep learning model for multi-classification of infectious diseases from unstructured electronic medical records. *BMC Med Inform Decis Mak* 2022; **22**: 41. https://doi.org/10.1186/s12911-022-01776-y
8. Choi J, Mavrommati K, Li NY, Patil A, Chen K, Hindin DI, et al. Scalable deep learning algorithm to compute percent pulmonary contusion among patients with rib fractures. *J Trauma Acute Care Surg* 2022; **93**: 461–66. https://doi.org/10.1097/TA.0000000000003619
9. Zhang B, Jia C, Wu R, Lv B, Li B, Li F, et al. Improving rib fracture detection accuracy and reading efficiency with deep learning-based detection software: a clinical evaluation. *Br J Radiol* 2021; **94**(1118): 20200870. https://doi.org/10.1259/bjr.20200870
10. Niiya A, Murakami K, Kobayash R, et al. Development of an artificial intelligence-assisted computed tomography diagnosis technology for rib fracture and evaluation of its clinical usefulness. *Sci Rep* 2022; **12** (8363). https://doi.org/10.1038/s41598-022-12453-5
11. Zhou QQ, Wang J, Tang W, Hu ZC, Xia ZY, Li XS, et al. Automatic detection and classification of rib fractures on thoracic CT using convolutional neural network: Accuracy and feasibility. *Korean J Radiol* 2020; **21**: 869–79. https://doi.org/10.3348/kjr.2019.0651
12. Yang C, Wang J, Xu J, Huang C, Liu F, Sun W, et al. Development and assessment of deep learning system for the location and classification of rib fractures via computed tomography. *Eur J Radiol* 2022; **154**: S0720-048X(22)00284-4. https://doi.org/10.1016/j.ejrad.2022.110434
13. Duan K, Bai S, Xie L, Qi H, Huang Q, Tian Q. CenterNet: Keypoint Triplets for Object Detection. 2019 IEEE/CVF International Conference on Computer Vision (ICCV); Seoul, Korea (South). Vis; 2019. pp. 6569–78. https://doi.org/10.1109/ICCV.2019.00667
14. Yu F, Wang D, Shelhamer E, Darrell T. Deep Layer Aggregation. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Salt Lake City, UT. ; 2018. pp. 2403–12. https://doi.org/10.1109/CVPR.2018.00255
15. Qiu Z, Yao T, Mei T. Learning Spatio-Temporal Representation with Pseudo-3D Residual Networks. 2017 IEEE International Conference on Computer Vision (ICCV); Venice. ; 2017. pp. 5533–41. https://doi.org/10.1109/ICCV.2017.590
16. BarkauRL, JohnsonMC, JacksonMG. *UNET: A model of unsteady flow through a full network of open channels*. ASCE; 2015.
17. Liu Y , Zhang M , Ma C , et al. Graph neural network[J]. 2020.
18. Talbot BS, Gange CP Jr, Chaturvedi A, Klionsky N, Hobbs SK, Chaturvedi A. Traumatic rib injury: Patterns, imaging pitfalls, complications, and treatment. *Radiographics* 2017; **37**: 628–51. https://doi.org/10.1148/rg.2017160100
19. Gupta A, Jamshidi M, Rubin JR. Traumatic first rib fracture: Is angiography necessary? A review of 730 cases. *Cardiovasc Surg* 1997; **5**: 48–53. https://doi.org/10.1016/s0967-2109(97)00060-4
20. Murphy CE, Raja AS, Baumann BM, Medak AJ, Langdorf MI, Nishijima DK, et al. Rib fracture diagnosis in the panscan era. *Ann Emerg Med* 2017; **70**: 904–9. https://doi.org/10.1016/j.annemergmed.2017.04.011
21. Sirmali M, Türüt H, Topçu S, Gülhan E, Yazici U, Kaya S, et al. A comprehensive analysis of traumatic rib fractures: Morbidity, mortality and management. *Eur J Cardiothorac Surg* 2003; **24**: 133–38. https://doi.org/10.1016/s1010-7940(03)00256-2
22. Bugaev N, Breeze JL, Alhazmi M, Anbari HS, Arabian SS, Holewinski S, et al. Magnitude of rib fracture displacement predicts opioid requirements. *J Trauma Acute Care Surg* 2016; **81**: 699–704. https://doi.org/10.1097/TA.0000000000001169
23. Zhou Q-Q, Hu Z-C, Tang W, Xia Z-Y, Wang J, Zhang R, et al. Precise anatomical localization and classification of rib fractures on CT using a convolutional neural network. *Clin Imaging* 2022; **81**: 24–32. https://doi.org/10.1016/j.clinimag.2021.09.010
24. Wu M, Chai Z, Qian G, Lin H, Wang Q, Wang L, et al. Development and evaluation of a deep learning algorithm for rib segmentation and fracture detection from multicenter chest CT images. *Radiol Artif Intell* 2021; **3**(5): e200248. https://doi.org/10.1148/ryai.2021200248
25. Ibanez V, Gunz S, Erne S, Rawdon EJ, Ampanozi G, Franckenberg S, et al. RiFNet: Automated rib fracture detection in postmortem computed tomography. *Forensic Sci Med Pathol* 2022; **18**: 20–29. https://doi.org/10.1007/s12024-021-00431-8
26. Tan H, Xu H, Yu N, Yu Y, Duan H, Fan Q, et al. The value of deep learning-based computer aided diagnostic system in improving diagnostic performance of rib fractures in acute blunt trauma. *In Review* 2021. https://doi.org/10.21203/rs.3.rs-413049/v1