CCE-Net: A rib fracture diagnosis network based on contralateral, contextual, and edge enhanced modules

Y. Gao, H. Liu, L. Jiang, C. Yang, X. Yin, Jean-Louis Coatrieux, Y. Chen

CCE-Net: A rib fracture diagnosis network based on contralateral, contextual, and edge enhanced modules

Yuan Gao[1,2], Hongzhi Liu[1,2], Liang Jiang[3*], Chunfeng Yang[1,2], Xindao Yin[3], Jean-Louis Coatrieux[4], Yang Chen[1,2*]

1 Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing 210096, China

2 Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, School of Computer Science and Engineering, Southeast University, Nanjing, 210096, China

3 Department of Radiology, Nanjing First Hospital, Nanjing Medical University, Nanjing 210006, China

4 Centre de Recherche en Information Biomédicale Sino-Francais, Inserm, University of Rennes 1, Rennes 35042, France

First author Yuan Gao's email address: yuangao@seu.edu.cn

Co-first author Hongzhi Liu's email address: liuhongzhi@seu.edu.cn

Corresponding author Yang Chen's email address: chenyang.list@seu.edu.cn

Co-corresponding author Liang Jiang's email address: Jiangliang0402@163.com

ABSTRACT

Rib fracture is a common disease that requires prompt treatment. This study focuses on developing a rib fracture diagnosis deep learning method using contralateral, contextual, and edge enhanced modules and evaluating its detection performance. A novel rib fracture diagnosis method was designed, named CCE-Net. To evaluate the performance of this method, 1639 digital radiography (DR) images were enrolled. Fracture features were extracted for three modules: contralateral, contextual, and edge enhanced modules. These modules can be used to identify fracture features in rib DR images, imitating the experience of broad-certificated radiologists. The contralateral module assists in diagnosing rib fractures by comparing the difference between the detected target region and the contralateral region. The contextual module helps to aid rib fracture detection by extracting contextual features. The edge enhanced module improves the accuracy of fracture detection by enhancing the edge information of the rib bone. The head of this two-stage detection network uses the multi-path fusion mechanism as the main architecture to integrate and utilize the above modules. The qualitative results show that with the ground truth of rib fracture as the evaluation standard, CCE-Net can achieve a better visual effect of fracture detection than other methods. The quantitative results show that CCE-Net can achieve the best performance in various detection indicators include AP50 0.911, AP75 0.794, AP25 0.913, and Recall 0.934. Experimental results show that CCE-Net can acquire the excellent ability of rib fracture diagnosis. We invasion that this approach will be applied to clinical study.

Keywords: rib fracture, deep learning, contralateral module, contextual module, edge enhanced module

1. Introduction

Rib fractures are the most common injury in blunt chest trauma [1–4]. It can cause chest pain and restrict physical mobility [5-11]. People over 45 years old with more than four rib fractures are considered dangerous [12,13]. Rib fractures need to be diagnosed and treated as soon as possible [14,15,16]. There are many reasons why it is necessary to diagnosis rib fractures: it will easily cause respiratory complications such as posttraumatic pneumonia if the diagnosis of rib fractures is not timely [4,17,18]; they are indicators of trauma-related diseases that require immediate treatment, such as pneumothorax and their onset can be delayed for several days [19]; the diagnosis of rib fractures can be used as the basis for further comprehensive treatment strategies [20,21]. DR is usually the preferred method of rib fracture detection [22]. Due to the different shapes of rib fractures, the rate of missed diagnosis and misdiagnosis is relatively high [23,24]. Rib fractures can become a life-threatening disease unless detected and treated appropriately, especially in elderly patients [25,26,27]. With the rapid development of artificial intelligence, it is worth introducing deep learning technology to improve rib fracture diagnosis and recognition accuracy as much as possible.

Previous work mainly focused on rib fracture detection in CT images. Weikert et al. [28] proposed a deep learning-based prototype algorithm detecting rib fractures on trauma CT on a pre-examination level. Zhou et al. [29] built a CNN model combining CT images and clinical information to detect and classify rib fractures automatically. Urbaneja et al. [30] proposed that CT with unfolded cylindrical projection can be used for rib fracture detection and characterization. Jin et al. [31] used a 3D-UNet model to solve the segmentation problem of rib fractures. Meng et al. [32] helped radiologists achieve high performance in diagnosing and classifying rib fractures on CT images with the assistance of deep learning algorithms. Although the above-mentioned CT-based methods have achieved good performance, there are few excellent methods based on DR. Compared with CT, DR has a faster imaging speed and a smaller radiation dose. DR is the first choice for the radiologist to diagnose rib fractures. Lindsey et al. [33] train a deep learning model to detect fractures on radiographs with a diagnostic accuracy similar to senior subspecialized orthopedic surgeons. Yahalomi et al. [34] train a Faster R-CNN, a machine vision neural network for object detection, to identify and locate distal radius fractures in anteroposterior X-ray images. Thian et al. [35] demonstrate the ability of an object detection CNN to detect and localize radius and ulna fractures on wrist radiographs with high sensitivity and specificity. Kim et al. [36] identify the extent to which transfer learning from deep convolutional neural networks (CNNs), pre-trained on non-medical images, can be used for automated fracture detection on plain radiographs. Kitamura et al. [37] use a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation to detect ankle fracture. Although there are many deep learning-based fracture detection studies on DR, there are scant clinically reliable rib fracture detection

algorithms. Due to low texture contrast, large differences in patient anatomy, and overlapping organs, it is very challenging to detect and locate rib fractures in DR images automatically.

This paper proposes a contralateral, contextual, and edge enhanced network (CCE-Net) address the above challenges of detecting rib fractures. CCE-Net mainly adopts a region-based two-stage detector. It combines three novel feature extraction streams: the contralateral module, contextual module, and edge enhancement module, so that accurate feature extraction. Since the left and right parts of the human chest ribs have many similar skeleton structures, radiologists often use the left-right comparison method to help them diagnose rib fractures. To extract the similarity information of the contralateral ribs, the spine line is used as the axis of symmetry to obtain a symmetrical rib patch. The contralateral patch is used as one of the neural network inputs. A fusion module is designed to integrate the features of the disease proposal and its contralateral reference patch. The above experience also applies to the upper and lower ribs, which have structural similarities. The contextual module is integrated into the pipeline. The flatness and smoothness of the bone edges are an important basis for fracture judgment when radiologists diagnose rib fractures. To capture bone edge information, the edges containing the key texture structure of the ribs are extracted. Specifically, the edge information reflecting the bone texture structure is integrated as one of the neural network inputs. Therefore, the above methods can help CCE-Net obtain rich characteristic information and excellent performance. The contributions of the study can be listed as follows.

1. This study aimed to develop a novel deep learning-based model for rib fractures automatically detection. The performance of our model is compared with other models in rib fractures detection.

2. A two-stage detection method is creatively proposed that can effectively integrate multiple feature extraction modules of the contralateral, contextual, edge enhanced, which makes full use of the unique feature information of the rib image. These modules are similar to the experience of radiologists in diagnosing rib fractures.

3. A weight distribution fusion method aims to fuse different image detail features and texture structure features of rib fractures at the decision level. It helps obtain a complete feature representation and enables the model to be trained end-to-end.

2. Material and Methods

2.1. Network architecture overview

This paper aims to detect rib fractures in DR images automatically. The flow chart for these steps is shown in Fig.1. The radiologists often focus on the differences in different regions and the smoothness of the bone edge as the diagnosis basis of rib fractures. Based on these essentials, the CCE-Net is proposed to exploit contralateral, contextual, and edge texture information to enhance the feature representations of rib fractures. The network architecture of CCE-Net is depicted in Fig.2. The contralateral and contextual patches are extracted according to the spine line segmentation algorithm. The edge images are acquired based on the edge extraction algorithm. These three

modules will be used as the neural network inputs to help the two-stage detector achieve better feature representation. The details are illustrated below.

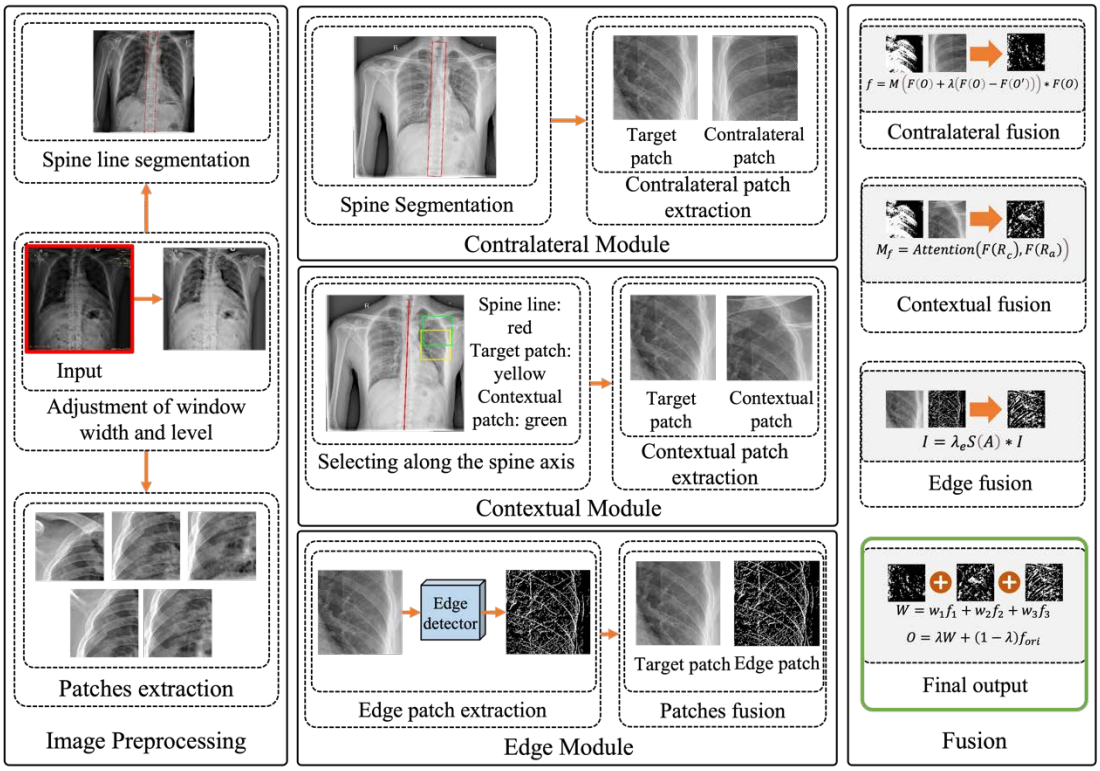


Fig.1 The workflow of this study. Image preprocessing can extract image patches and the spine line from original images. Three novel modules (contralateral, contextual, and edge enhanced) can obtain more features of rib fractures. The feature fusion of the three modules can assist the basic detection network in achieving performance improvement. It should be noted that the DR image with the red outline on the left is used as the input of this method, and the fusion result of the green outline on the right is used as the output of this method.

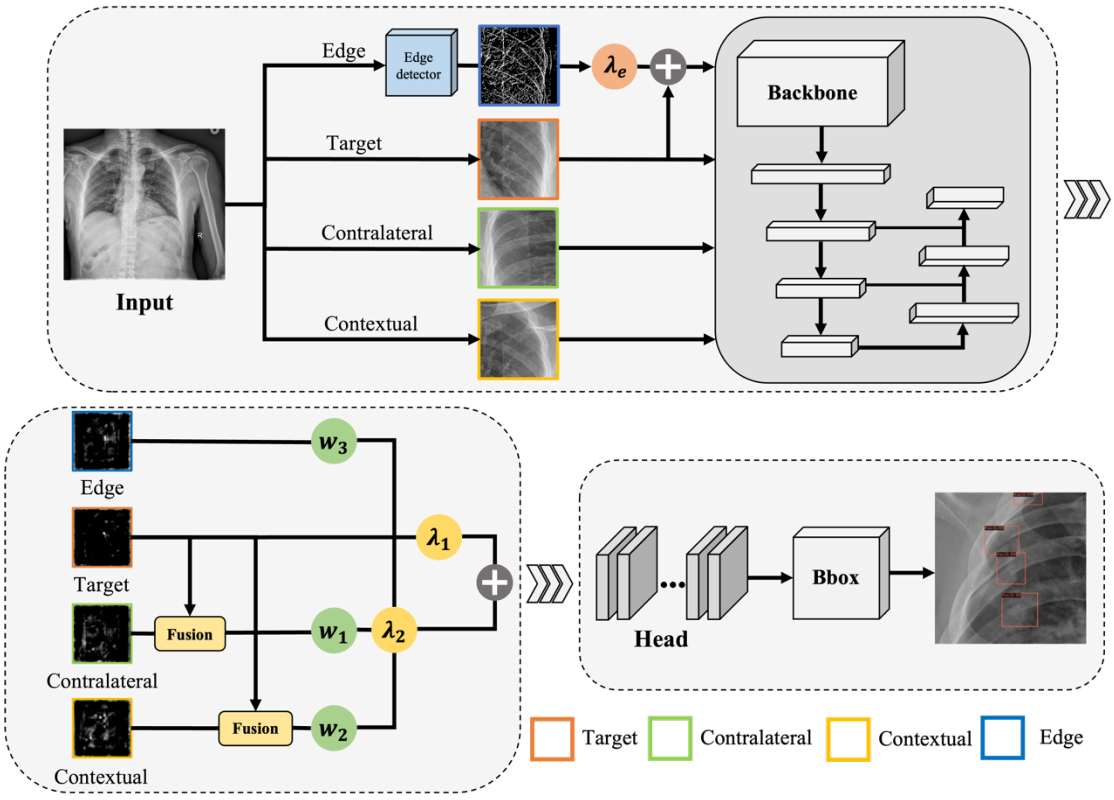


Fig.2. The network architecture of CCE-Net. Our proposed method is based on the two-stage network. It extracts valid information on rib fractures, integrating four branches marked in orange, green, yellow and blue: target, contralateral, contextual, and edge. Four branches perform the same feature extraction on the backbone network. The results of four branches can be effectively fused in the neck stage to provide more reliable feature information in the subsequent stages.

2.2. Contralateral module

The contrast of the contralateral patch is useful for radiologists to reference. Especially when fractures are contained in images, the contralateral information helps differentiate the abnormal and the normal ribs and better highlights the difference of fracture regions. Considering that the ribs of the human body are symmetrical, and the spine is situated in the relatively middle position of the chest. The symmetrical central axis of the ribs on both sides is the same as the line of the human spine. We first roughly segment the spine line and extract the contralateral patch based on the left-right symmetry of the ribs.

2.2.1 Spine line roughly extraction

Examining symmetrical regions on both sides of the spine can help radiologists determine rib fractures. The human spine is a skeletal organ usually located in the center and has a regular shape in the DR image. Therefore, the spine line is only needed to roughly segment to examine the contralateral patch of the rib.

Specifically, the spine line can be roughly segmented by a CNN model. The UNet [38] is selected as the spine line segmentation model. We label the spine region on the

DR data set and use this data set to train the model. Because the symmetrical central axis of the ribs on both sides is not a precise line, there is no need for high segmentation accuracy requirements for the spine segmentation model.

The minimum circumscribed quadrilateral enclosing the spine mask can be obtained. The spine line bridged by the centers of two short edges is regarded as the symmetric axis. The spine line can be expressed as $Ax + By + C = 0$. The target patch is denoted as $P_{target}(x_a, y_a)$, and the contralateral patch can be denoted as $P_{contra}(x_b, y_b)$. T The following formula can express the contralateral patch:

$$x_b = \frac{x_a - (2A * (Ax_a + By_a + C))}{A^2 + B^2}$$
$$y_b = \frac{y_a - (2B * (Ax_a + By_a + C))}{A^2 + B^2} \quad \text{Eq (1)}$$

After the central axis of symmetry is determined, the region symmetrical with the target region according to the spine symmetry axis can be determined as the contralateral patch.

2.2.2. Contralateral images fusion

After obtaining the contralateral patch, the characteristics of the highly similar structure on both sides of the human ribs are used to pair each target patch with its contralateral patch. An attention mechanism uses the difference between the target and contralateral patch in high-level semantic features to guide the potential location of the fracture. The contrast-induced attention [39] is used to fuse features of each target patch and its contralateral patch.

The feature fusion process of the target area and the opposite patch feature is shown in Fig.3. We define the target patch as O, and the contralateral patch obtained as O'. They perform the same convolution operation as F in the backbone network. The feature maps are effectively fused to realize the contralateral information extraction. The fusion operation adopts a pixel-by-pixel method. The subtraction operation between $F(O)$ and $F(O')$ helps provide contrast information and suppress the response of attributes unrelated to fracture recognition and location. On the other hand, adding the above subtraction results from the target patch is helpful to identify the same structural information of the ribs with larger responses in the patches on both sides. The attention module M encodes $F(O)$ and $F(O')$ into the attention maps. The attention maps are multiplied pixel by pixel to weight the original target feature map to obtain f. The operation is shown below.

$$f = M\left(F(O) + \lambda\big(F(O) - F(O')\big)\right) * F(O) \quad \text{Eq (2)}$$

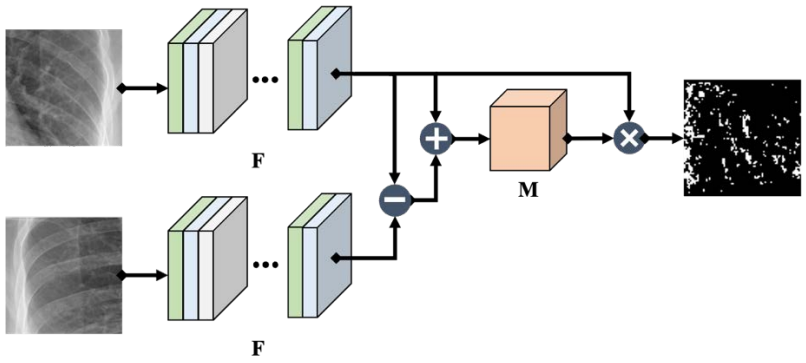


Fig.3. An overview of the fusion flow. It aims to fuse the effective information of each part to extract the results. The addition and subtraction of the contralateral patch and the target patch are effectively fused with the target patch by using the attention module.

2.3. Contextual module

Focusing on multiple adjacent ribs can make the detection more accurate. The normal rib is similar in shape and signal intensity to its adjacent normal rib. By checking on the contextual information of the diagnosed region, the radiologists can intuitively obtain the changes in the appearance of bones due to fractures. A contextual module that takes adjacent contextual ribs as input is designed to imitate a radiologist.

2.3.1. Contextual images extraction

The longitudinal position of each diagnosed rib region needs to be judged according to the rough segmentation result of the spine. Since there are no more than 12 human ribs on each side [40], one of the adjacent upper and lower rib regions should be selected as the input of the context module. If the longitudinal position belongs to the upper region of the spine, the adjacent rib region on the lower side is taken as the input of the context module; otherwise, the adjacent rib region on the upper side is taken as the input of the context module. After the patch direction selection decision is determined, the diagnosed rib patch can be moved along the extension direction of the spine's central axis by half of the boundary length to obtain the context patch. It should be noted that the boundary length refers to the width of the patch, and this study uses 640 pixels as the patch width. Due to the short distance between the adjacent ribs of the human body, half the distance of the moving patch boundary can sufficiently cover the adjacent ribs.

2.3.2. Contextual images fusion

The contextual module is designed to take two adjacent ribs as input patches, including the current diagnosis rib and its adjacent rib. A two-branch structure is used to compare adjacent ribs to identify fractures. This structure uses the same fusion method as in Equation 2 in the previous section. We denote the two patches of the current diagnosis rib and its adjacent rib as R_c and R_a, respectively. These two branches have the same feature extraction network as F. The resulting feature map uses the attention mechanism for feature layer fusion and then is input to the rest of the

convolution of the overall network. The final fusion feature maps M_f can be expressed as:

$$M_f = Attention(F(R_c) + \lambda(F(R_c) - F(R_a))) \quad \text{Eq (3)}$$

2.4. Edge enhanced module

To capture the edge information of the ribs in the image, we extract the edges containing key texture information for visual recognition and integrate the edge map with the neural network model to improve fracture detection. Rib fractures are highly correlated with the appearance of the bone edges of the ribs. For example, a fractured rib will show abnormal curvature of the skeletal cortex. Based on the above experience, we believe that purely enforcing the widening of the receptive field is not sufficient for rib fracture detection, and the introduction of more comprehensive edge information can improve the detection effect. Since the edges reflect local intensity changes and display the boundary information of ribs in the image, maintaining the edges can preserve the structure of the image content and the texture details.

2.4.1. Edge enhanced images extraction

The edge information is added to the neural network to enforce the feature map's integration with the original image's edge information. The Sobel operator can calculate the edge map due to its simplicity. Sobel filters S are used to convolve the original image I_o to generate the edge map. Then, the edge information map is integrated into the input image I by addition pixel by pixel as the following equation:

$$I_e = \lambda_e S(I_o) \oplus I \quad \text{Eq (4)}$$

Where λ_e means the scale factor of edge information, and $\oplus$ means the operation of addition pixel by pixel.

By adding the scale factor calculated by the Sobel operator to the input image, the neural network can integrate the information of texture structure and edge intensity to pay more attention to the rib edge information and reduce the sensitivity to noise. The Sobel filter assigns higher weights to the edge region of the original image and lower weights to other regions, which directly enhances the use of image edges by the neural network.

2.5. Fusion architecture

As shown in Fig.2, all three parts are integrated into the whole framework. Our work aims to integrate each region proposal feature f of contralateral, contextual and edge enhanced modules into the final output O. Integrate by simply connecting operation to make the contribution of each stream equal, which may ignore the high-value features from a module. Feature fusion methods may enhance some feature modules and suppress some feature modules. If we simply combine each feature extraction module mechanically, it is easy to lose the features that should be enhanced. Therefore, we need to design a fusion method that conforms to the characteristics of deep learning. This study needs to select appropriate scale features for different feature extraction modules. The effective fusion of features is achieved by introducing self-learning to select appropriate scales. We link the results of the three modules by

assigning different weight factors to the feature extraction module through the weight distribution module. A weight distribution module is introduced as W that can perform backward propagation and adaptively determine the weights of different modules. The weight distribution module is inspired by the attention mechanism and consists of multiple convolutional and pooling layers. The specific process can be expressed as the following Equation 5:

$$W = w_1 f_1 + w_2 f_2 + w_3 f_3 \quad \text{Eq (5)}$$

Where f_1, f_2 and f_3 mean the feature maps result of contralateral, contextual and edge enhanced modules, w_1, w_2 and w_3 mean the attention mechanism operation.

The final output feature maps of CCE-Net can be expressed as O:

$$O = \lambda W + (1 - \lambda) f_{ori} \quad \text{Eq (6)}$$

Where f_{ori} means the feature maps of original rib patch, λ means the parameter weighting the importance of these three modules.

2.6. Loss function

The detection network architecture that combines the above three modules has been established. The loss function can calculate the error between actual values and predicted values. Cross entropy loss is chosen as the objection function and is given by Eq (7):

$$loss = -\frac{1}{N}\sum (y \times \log(\hat{y}_m) + (1 - y) \times \log(1 - \hat{y}_m)) \quad \text{Eq (7)}$$

where y is the label and y_m is the predicted output vector.

3. Experimentals and Results

3.1. Dataset

A private dataset named Rib-NJFH is collected to train and validate our proposed method. Due to the protection of patient privacy, please forgive us that the dataset used in this study cannot be publicly available. The dataset contains 1639 DR images from Nanjing First Hospital, among which there are 2703 rib fractures. We use 1311 images for training, 164 images for validating, and 164 images for testing. All images are annotated and examined by experienced radiologists.

3.2. Evaluation metrics

The bounding box AP [41] and recall are calculated to evaluate the performance of the model. Considering rib fractures as a general target, AP50 is used as the main evaluation index. AP25 and AP75 are used as references. We believe that AP75 better reflects the accurate positioning performance of the fracture region due to its strict evaluation criteria, AP25 has relatively loose evaluation criteria to determine whether the test results are misjudged and thus better reflects the recognition performance of rib abnormalities, AP50 is a comprehensive performance index for fracture recognition and regional positioning [42, 43].

$$AP = \int_0^1 P(r)dr \quad \text{Eq (9)}$$

The PR curve with the corresponding AUCs was calculated for validation datasets to evaluate the performance of the model. The Recall is defined by Eq (10):

$$Recall = \frac{TP}{TP + FN} \quad \text{Eq (10)}$$

Where TP and FN are the numbers of positive samples correctly classified and incorrectly classified, TN and FP are the numbers of negative simples correctly classified and incorrectly classified.

3.3. Implementation details

The segmentation of the spine's central axis is needed to use in the contralateral module and the contextual module for our experiment. Specifically, the spine regions of the training set are labeled. An example of spine labeling is shown in Fig.4. The segmentation algorithm is used to complete the segmentation of the spine region in the image preprocessing part. After the segmentation results are obtained, image morphology is used to extract the central axis from the spine mask to obtain the contralateral patch.

Fig.4. The spine line obtention. A red box marks the spine segmentation result. The spine line is used as the position basis of the image extraction of the contralateral module and the contextual module.

The data set is divided into the training set, the validation set and the test set at 8:1:1. The image patches randomly cropped from DR images are resized to 640×640 pixels. Rotation, horizontal and vertical flipping are used for data augmentation. All experiments are implemented using Pytorch on 2 NVIDIA 1080Ti GPUs. ResNet-50 is used as the backbone of the proposed method. We train our model for 50 epochs and test it every five times. For all training, the optimizer is stochastic gradient descent (SGD) with a weight decay of 0.001 and momentum of 0.9 to optimize all models and the batch-size is 2 on each GPU. The learning rate starts at 0.01 and reduces by a factor of 10 after 30 and 40 epochs.

3.4. Models comparison

According to the characteristics of our dataset and detection task, we compare with methods that include Faster RCNN [44], Libra RCNN [45], Dynamic RCNN [46], Cascade RCNN [47] and YOLO v4 [48].

3.4.1. Quantitative Results

To validate the performance of CCE-Net, five mainstream algorithms are selected to implement the comparison experiments. For rib DR images, the performance metrics in the training sets can be achieved with the accuracy of 99.648 and the loss of 0.027. The validation set can achieve AP50 and Recall of 0.910 and 0.938. In the test sets, the model's detection performance yielded the AP50 of 0.911, and the Recall is 0.934, as shown in Table 1.

Table 1 The quantitative comparisons performance of our method.

Method	AP50	AP75	AP25	Recall	AUC
Faster RCNN	0.787	0.349	0.878	0.875	0.816
Libra RCNN	0.825	0.326	0.862	0.886	0.847
Dynamic RCNN	0.887	0.516	0.904	0.903	0.901
Cascade RCNN	0.910	0.781	0.911	0.929	0.933
YOLO v4	0.813	0.689	0.816	0.881	0.840
CCE-Net	0.911	0.794	0.913	0.934	0.941

The five comparison models, including ‘Faster RCNN with ResNet-50’, ‘Libra RCNN’, ‘Dynamic RCNN’, ‘Cascade RCNN with ResNeXt-101’ and ‘YOLO v4’, are re-implemented using our dataset. We are committed to merging our proposed module into a two-stage network of Faster RCNN with ResNet-50 as introduced in Section 3. The performance of our method compared with different methods is presented in Table 1. The experimental results in Table 1 show that our method achieves the best performance in all evaluation metrics. Our method equipped with ‘Faster RCNN’ achieves 15.76% AP50, 127.5% AP75 and 3.99% AP75 higher than the results of the original ‘Faster RCNN’ respectively, the Recall increase by 6.74%, which shows the advantages of the network in the rib fracture detection ability. In addition, due to the ResNet-50 backbone network, our method can achieve fast detection speed while ensuring detection accuracy.

For all indicators, our method can achieve better results than other methods. The improvement of AP50 brought by our method is 10.42% (from 0.825 to 0.911), 2.71% (from 0.887 to 0.911) and 1.1% (from 0.91 to 0.911) when using ‘Libra RCNN’, ‘Dynamic RCNN’ and ‘Cascade RCNN with ResNeXt-101’ to produce rib fracture proposals. Compared with ‘YOLO v4’, which is the one-stage model and achieved outstanding detection performance, the one-stage model with ‘YOLO v4’ achieves 0.813 AP50 and 0.881 Recall. Our method is 0.098 and 0.053 higher than the 'YOLO v4' results, respectively.

3.4.2. Qualitative Results

Visualization images are used to show the difference between the proposed method and other methods. The results of CCE-Net and other methods are shown in Fig.5. The second column is the result of the proposed method. CCE-Net can effectively detect fracture targets. The third to seventh columns are the results of comparison methods. There are missed detections and false detections in the results of comparison methods.

To illustrate the effectiveness of information extraction, feature maps of three different layers are visualized. The visualization includes the result and the feature maps

with sizes of 64, 32, 16 pixels. The feature maps of different sizes are enlarged to the same size and displayed by superimposing the original image. As shown in Fig.6, the proposed method can focus on more reasonable regions. Overall, the correctness of the proposed method is the best because it focuses on more reasonable image features.

The PR curves of different methods on the validation data can be shown in Fig.7. For the results of CCE-Net with rib fractures detection, the method obtains good performance.

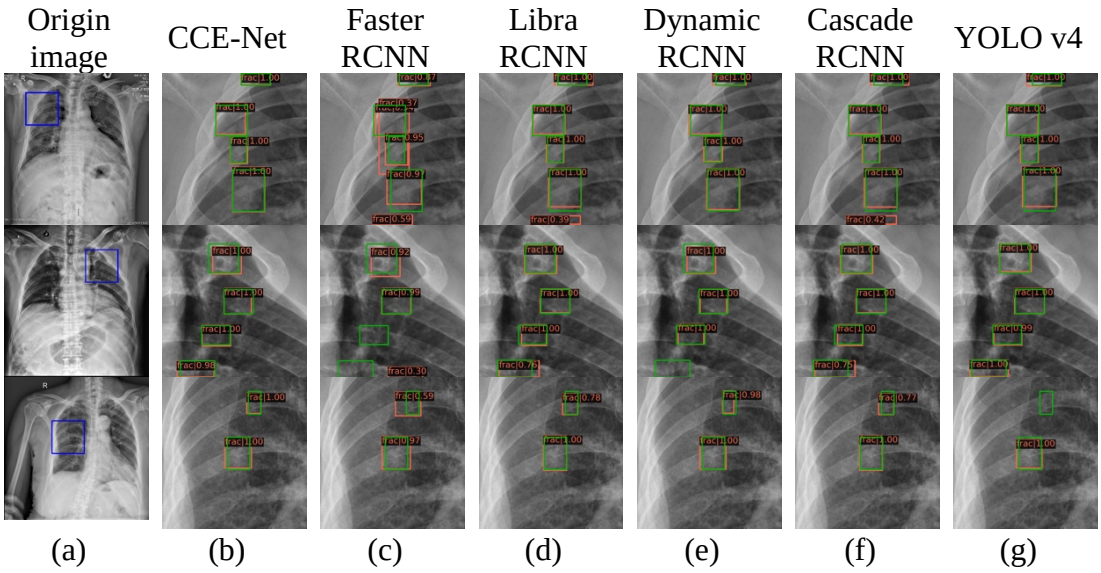


Fig.5. The results demonstrate that the CCE-Net with three modules has better detection Precision and Recall performance. Blue, green, and red boxes stand for the patch position, ground truth, and results, respectively. (a) represent origin images. (b), (c), (d), (e), (f), and (g) represent results obtained via CCE-Net, Faster RCNN, Libra RCNN, Dynamic RCNN, Cascade RCNN, and YOLO v4, respectively.

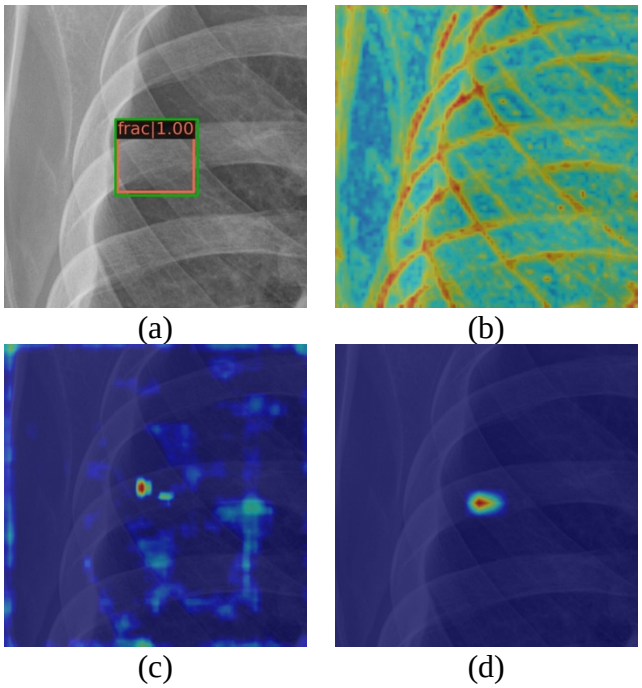


Fig.6. The visualization images show that the CCE-Net can pay more attention to the rib fracture region as the decline in feature maps size. Feature maps are superimposed displayed on the original image through uniform scaling. (a) represent the result. (b), (c), and (d) represent the feature map of different sizes of 64*64, 32*32, and 16*16, respectively.

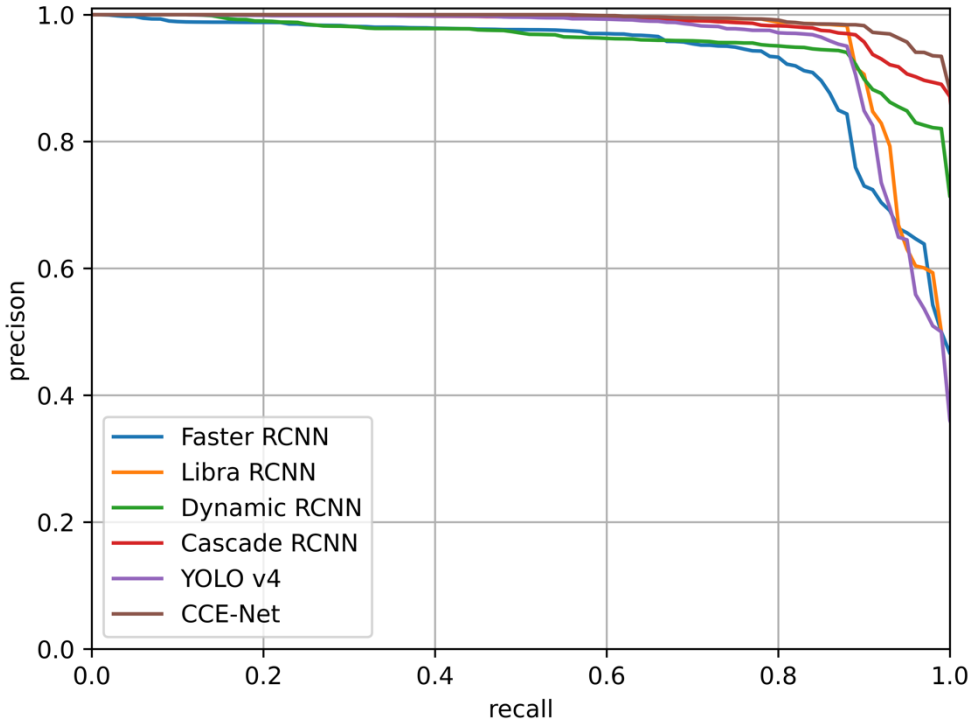


Fig.7. The PR curves show that the CCE-Net has the best performance compared with other methods.

3.5. Ablation study

To validate the effectiveness of each module, an ablation study is conducted on the proposed model. We use the same training, validation, and test sets in all experiments. The 'Faster RCNN with ResNet-50' is used as the baseline model. The contralateral, contextual, and edge enhanced modules are removed from CCE-Net. The results of ablation are shown in Table 2 and Fig.8. The PR curves of the ablation experiment can be seen in Fig.9, which demonstrates the enhancement of our method.

3.5.1. The effect of contralateral module

Compared with CCE-Net, if we remove the contralateral module, the performance is decreased by 1.9% and 0.76% on AP50 and Recall, respectively. The decline on AP50 validates the effectiveness of the contralateral module on overlapping positions.

3.5.2. The effect of contextual module

Removing the contextual module decreases performance by 1.67% and 2.52% on AP50 and Recall, respectively, compared with CCE-Net. It indicates that the contextual module enhances feature representation more for regions of structure repetition. When

adding the contextual module, partial fracture judgment is corrected based on the upper and lower adjacent ribs.

3.5.3. The effect of edge enhanced image module

Compared with the contralateral module and contextual module, adding the edge enhanced module can boost the performance by 7.05% and 7.6% on AP50 and Recall, respectively. The edge enhanced module is encoded by edge information and texture information. This encoding mechanism affects the rib fracture localization, making the proposed method obtain more edge information gains.

Table 2 The ablation study of our method.

	Contralateral module	Contextual module	Edge enhanced module	AP50	Recall	AUC
CCE-Net	*	*	*	0.911	0.934	0.941
	remove	*	*	0.894	0.927	0.933
	*	remove	*	0.896	0.911	0.938
	*	*	remove	0.851	0.868	0.909

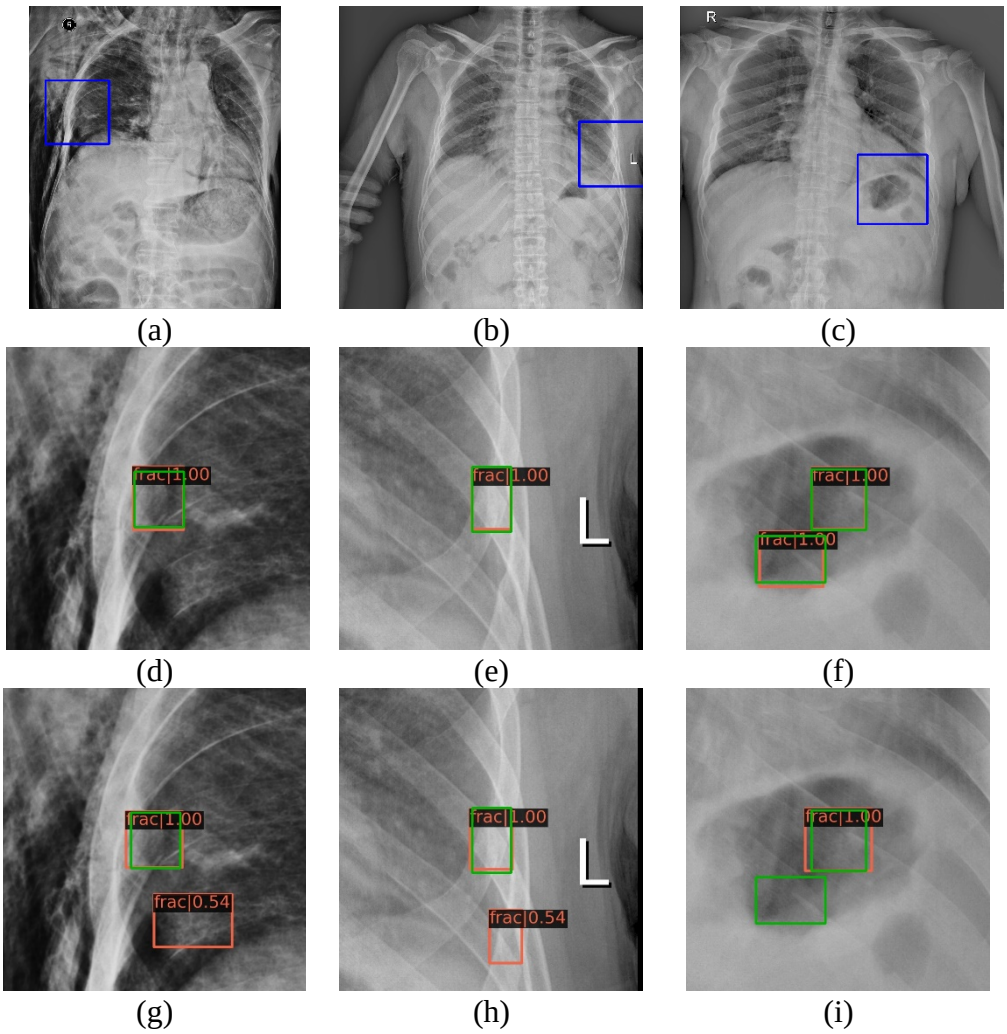


Fig.8. These comparison cases show that the contralateral, contextual, and edge

enhanced module can improve detection capabilities at overlapping positions, structure repetition regions, and complex edge locations, respectively. (a), (b), and (c) represent different origin images. (d), (e), and (f) represent CCE-Net. (g), (h), and (i) represent the results of removing the contralateral module, contextual module, and edge enhanced module, respectively.

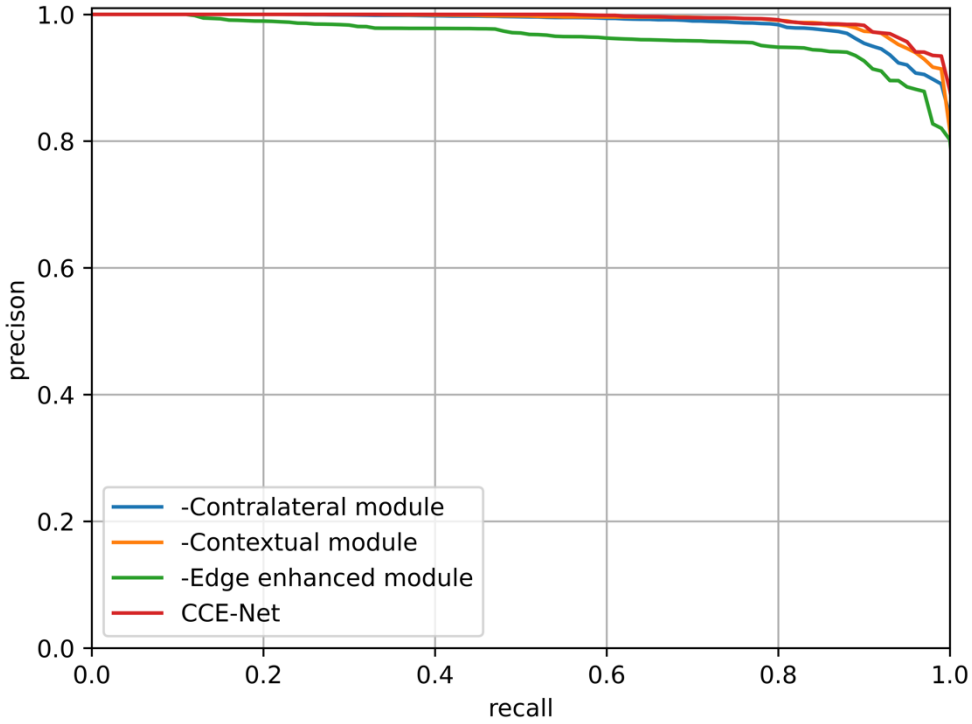


Fig.9. The PR curves show the comparison of the ablation study.

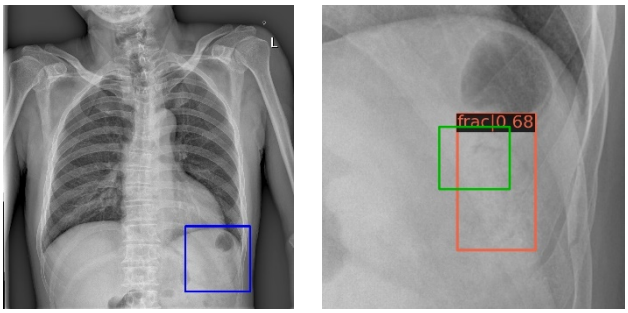
4. Discussion

This study proposes a novel network architecture for image detection of rib fractures. Three modules are integrated into the Faster RCNN [44] framework. The contralateral module is used to obtain contralateral information of the same structure on both sides of the spine. The contextual module is added to extract image features of the upper and lower positions of the ribs. The edge enhanced module can stress rib image edges and obtain the target details more effectively. Experimental results show that combining these three modules improves the rib fracture detection performance in evaluation indicators. Compared to the performance of the Faster RCNN, AP50 is observed to increase by 15.76% from 0.787 to 0.911 and Recall increases by 6.74% from 0.875 to 0.934. Even compared with many mainstream detection algorithms, including Faster RCNN [44], Libra RCNN [45], Dynamic RCNN [46], Cascade RCNN [47] and YOLO v4 [48] in our experiment, the proposed method exhibits a certain performance improvement in detection effect and fewer training parameters. The ablation experiments are also conducted on CCE-Net by removing different modules. The methods after removing different modules have a certain degree attenuation of the detection effect in different scenarios. It is verified that the proposed method can detect

various rib fracture features more effectively and comprehensively.

Rib image data are more conducive to detecting fractures when combined with the visual features of contralateral, contextual and edge images. One elaborate fusion way to accomplish the features fusion process is to fuse image features from three modules at different network stages. The feature fusion process at the neck of basic two-stage network architecture is designed for the contralateral and contextual modules. They have different original images requiring feature extraction at the backbone stage. The edge image is added to the rib image straightforwardly before feature extraction for the edge enhanced module. In the feature fusion process details, we innovatively used a mixture of numerical operations and attention mechanisms to achieve a better feature fusion strategy. Besides, the different information contained in the three modules is needed to combine to make the final decision. The reasonable weight control of each feature channel is designed at the neck stage of the network, expecting that the proposed method can simulate the clinical diagnosis ideas of radiologists.

Although the proposed network architecture has good detection capabilities in rib fracture images, the detection results still have several limitations due to the complexity of medical data. Some failure cases are shown in Fig.10. The limitations of our method are as follows: (1) For some DR images, due to the curvature of the spine, the accurate extraction of curved spine lines still needs to solve; (2) Whether the texture and structural similarity of rib images can be better integrated is also an interesting research topic. To settle these issues, it deserves further study to design a more efficient module to explore the contralateral contextual information and enhance the extraction of information.



(a) (b)

Fig.10. These failure cases demonstrate that our proposed method is still worth improving. (a) The origin image, (b) CCE-Net.

5. Conclusion

This study proposed a CCE-Net based on contralateral, contextual, and edge enhanced modules to detect rib fracture. A rib fracture is a kind of small target object, which is difficult to detect integrally. Following the valuable experience of radiologists in diagnosis, new modules were added to the design of the detection network. CCE-Net unified contralateral, contextual and edge information together. Compared with the traditional detection network, CCE-Net can capture more effective information. We established the rib fracture database with 1639 DR rib images to train the CCE-Net

model. A total of 2703 rib fractures were in our database labeled by experienced radiologists. Based on experiments, compared with other methods, its performance can improve the medical image detection ability of rib fracture targets. The detection performance of the CCE-Net was significantly improved than the current methods. CCE-Net attained AP50 0.911, AP75 0.794, AP25 0.913, and Recall 0.934. It can reduce workload for radiologists and assist radiologists in rib fracture diagnosis.

Acknowledgments

This work was supported in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grants SJCX21_0635; in part by the Xinghuo Talent Program of Nanjing First Hospital; in part by the State's Key Project of Research and Development Plan under Grants 2017YFA0104302, 2017YFC0109202, and 2017YFC0107900, in part by National Natural Science Foundation under Grants 81530060 and 61871117, in part by the Science and Technology Program of Guangdong under Grant 2018B030333001.

CRediT authorship contribution statement

Yuan Gao: Conceptualization, Methodology, Formal analysis, Data curation, Validation, Writing. Hongzhi Liu: Conceptualization, Methodology, Writing, Validation, Investigation. Yang Chen: Funding acquisition, Writing, Validation, Investigation. Liang Jiang: Funding acquisition, Conceptualization, Methodology, Supervision, Writing. Chunfeng Yang: Data curation, Writing. Xindao Yin: Formal analysis, Data curation, Funding acquisition, Supervision. Jean-Louis Coatrieux: Data curation, Writing.

Conflict of 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.

Reference

1. S.H. Cho, Y.M. Sung, M.S. Kim, Missed rib fractures on evaluation of initial chest CT for trauma patients: Pattern analysis and diagnostic value of coronal multiplanar reconstruction images with multidetector row CT, Br J Radiol. 85 (2012) 845-850.
2. M.P. Fredric, M. Sarah, A.O. Francis, A. Darwin, D. Andrew, G.E. John, F. Bruce, G. Mario, M. Silvana, M. Christian, S. Babak, T. William, H. V. Don, W.W. Thomas, Consensus statement: surgical stabilization of rib fractures rib fracture colloquium clinical practice guidelines, Injury. 48 (2017) 307-321.
3. J.C. Mayberry, D.D. Trunkey, The fractured rib in chest wall trauma, Chest Surg

Clin N Am. 7 (1997) 239-261.
4. B.S. Talbot, C.P. Gange, A. Chaturvedi, N. Klionsky, S.K. Hobbs, A. Chaturvedi, Traumatic rib injury: Patterns, imaging pitfalls, complications, and treatment, Radiographics. 37 (2017) 628-651.
5. F.C. Lin, R. Li, Y. Tung, K. Jeng, S.C. Tsai, Morbidity, mortality, associated injuries, and management of traumatic rib fractures, J Chinese Med Assoc. (2016) 6-11.
6. B. Eric, L. Andre, C. David, M. Lynne, R. Sebastien, T. Stephane, L. Jacques, M. Marcel, Elderly trauma patients with rib fractures are at greater risk of death and pneumonia, Journal of Trauma and Acute Care Surgery. 54 (2003) 478-485.
7. S. RobTodd, M.M. McNally, J.B. Holcomb, R. A. Kozar, L.S. Kao, E.A. Gonzalez, C.S. Cocanour, G.A.Vercruysse, M.H. Lygas, B. K. Brasseaux, F.A. Moore, A multidisciplinary clinical pathway decreases rib fracture–associated infectious morbidity and mortality in high-risk trauma patients, The American Journal of Surgery. 192 (2006) 806-811.
8. R.M. Shorr, A. Rodriguez, M.C. Indeck, M.D. Crittenden, S. Hartunian, R.A. Cowley, Blunt chest trauma in the elderly, The Journal of Trauma. 29 (1989) 234-237.
9. F. Birse, H. Williams, D. Shipway, E. Carlton, Blunt chest trauma in the elderly: an expert practice review, Emergency Medicine Journal. 37 (2020) 73-78.
10. D. Stephanie, A. Affatato, Blunt chest trauma: utility of radiological evaluation and effect on treatment patterns, The American journal of emergency medicine. 26 (2006) 482-486.
11. L. Fabricant, B. Ham, R. Mullins, J. Mayberry, Prolonged pain and disability are common after rib fractures, The American Journal of Surgery. 205 (2013) 511-516.
12. J.B. Holcomb, N.R. McMullin, R.A. Kozar, M.H. Lygas, F.A. Moore, Morbidity from rib fractures increases after age 45, Journal of the American College of Surgeons. 196 (2003) 549-555.
13. B.T. Flagel, F.A. Luchette, R.L. Reed, T.J. Esposito, K.A. Davis, J.M. Santaniello, R.L. Gamelli, Half-a-dozen ribs: The breakpoint for mortality, Surgery. 138 (2005) 717-725.
14. R. Kent, W. Woods, O. Bostrom, Fatality risk and the presence of rib fractures, Annals of Advances in Automotive Medicine/Annual Scientific Conference, Association for the Advancement of Automotive Medicine. 52 (2008).
15. W.M. Wu, Y. Yang, Z.L. Gao, T.C. Zhao, W.W. He, Which is better to multiple rib fractures, surgical treatment or conservative treatment?, International journal of clinical and experimental medicine. 8 (2015) 7930.
16. M. Sirmali, H. Türüt, S. Topçu, E. Gülhan, Ü. Yazici, S. Kaya, I. Taştepe, A comprehensive analysis of traumatic rib fractures: morbidity, mortality and management, European Journal of Cardio-Thoracic Surgery. 24 (2003) 133-138.
17. S.W. Ho, Y.H. Teng, S.F. Yang, H.W. Yeh, Y.H. Wang, M.C. Chou, C.B. Yeh, Risk of pneumonia in patients with isolated minor rib fractures: a nationwide cohort study, BMJ open. 7 (2017) e013029.

18. H. Tanaka, T. Yukioka, Y. Yamaguti, S. Shimizu, H. Goto, H. Matsuda, S. Shimazaki, Surgical stabilization of internal pneumatic stabilization? a prospective randomized study of management of severe flail chest patients, J Trauma. 52 (2002) 727-732.
19. M.S. Lu, Y.K. Huang, Y.H. Liu, H.P. Liu, C.L. Kao, Delayed pneumothorax complicating minor rib fracture after chest trauma, Am J Emerg Med. 26 (2008) 551-554.
20. M. Bemelman, M.W. de Kruijf, B.M. van, L. Leenen, Rib fractures: To fix or not to fix? An evidence-based algorithm, Korean J Thorac Cardiovasc Surg. 50 (2017) 229-234.
21. M.B. de Jong, M.C. Kokke, F. Hietbrink, L.P.H. Leenen, Surgical management of rib fractures: Strategies and literature review, Scand J Surg. 103 (2014) 120-125.
22. E.G. Hwang, Y. Lee, Simple X-ray versus ultrasonography examination in blunt chest trauma: effective tools of accurate diagnosis and considerations for rib fractures, J Exerc Rehabil. 12 (2016) 637-641.
23. J. Malghem, B.C. Vande Berg, F.E. Lecouvet, B.E. Maldague, Costal cartilage fractures as revealed on CT and sonography. Am J Roentgenol, 176 (2001) 429-432.
24. M. Kara, E. Dikmen, H.H. Erdal, I. Simsir, S.A. Kara, Disclosure of unnoticed rib fractures with the use of ultrasonography in minor blunt chest trauma, Eur J Cardio-thoracic Surg. 24 (2003) 608-613.
25. L.I.G. Worthley, Thoracic epidural in the management of chest trauma, Intensive Care Med. 11 (1985) 312-315.
26. Y. Barnea, H. Kashtan, Y. Skornick, N. Werbin, Isolated rib fractures in elderly patients: Mortality and morbidity, Can J Surg. 45 (2002) 43-46.
27. T.J. Ellis, Hip fractures in the elderly, Curr Womens Health Rep. 3 (2003) 75-80.
28. T. Weikert, L.A. Noordtzij, J. Bremerich, S. Bram, P. Victor, C. Joshy, S. Gregor, W.S. Alexander, Assessment of a deep learning algorithm for the detection of rib fractures on whole-body trauma computed tomography, Korean J Radiol. 21 (2020) 891-899.
29. Q.Q. Zhou, W. Tang, J. Wang, Z.C. Hu, Z.Y. Xia, Z. Rongguo, X. Fan, W. Yong, X. Yin, B. Zhang, H. Zhang, Automatic detection and classification of rib fractures based on patients' CT images and clinical information via convolutional neural network, Eur Radiol. 31 (2021) 3815-3825.
30. A. Urbaneja, J. De Verbizier, A.S. Formery, C. Tobon-Gomez, L. Nace, A. Blum, P. A. G. Teixeira, Automatic rib cage unfolding with CT cylindrical projection reformat in polytraumatized patients for rib fracture detection and characterization: Feasibility and clinical application, Eur J Radiol. 110 (2019) 121-127.
31. L. Jin, J. Yang, K. Kuang, B. Ni, Y. Gao, Y. Sun, P. Gao, W. Ma, M. Tan, H. Kang, J. Chen, M. Li, Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet, EBioMedicine.

62 (2020) 103106.
32. X.H. Meng, D.J. Wu, Z. Wang, X.L. Ma, X.M. Dong, A.E. Liu, L. Chen, A fully automated rib fracture detection system on chest CT images and its impact on radiologist performance, Skeletal Radiol. (2021) 1821-1828.
33. R. Lindsey, A. Daluiski, S. Chopra, A. Lachapelle, M. Mozer, S. Sicular, D. Hanel, M. Gardner, A. Gupta, R. Hotchkiss, H. Potter, Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A. 115 (2018) 11591-11596.
34. E. Yahalomi, M. Chernofsky, M. Werman, Detection of Distal Radius Fractures Trained by a Small Set of X-Ray Images and Faster R-CNN, Intelligent Computing-Proceedings of the Computing Conference. (2019) 971-981.
35. Y.L. Thian, Y. Li, P. Jagmohan, D. Sia, V.E.Y. Chan, R.T. Tan, Convolutional Neural Networks for Automated Fracture Detection and Localization on Wrist Radiographs, Radiol Artif Intell. 1 (2019) e180001.
36. D.H. Kim, T. MacKinnon, Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks, Clin Radiol. 73 (2018) 439-445.
37. G. Kitamura, C.Y. Chung, B.E. Moore, Ankle Fracture Detection Utilizing a Convolutional Neural Network Ensemble Implemented with a Small Sample, De Novo Training, and Multiview Incorporation, J Digit Imaging. 32 (2019) 672-677.
38. O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention, Springer (2015) 234–241.
39. J. Liu, G. Zhao, Y. Fei, M. Zhang, Y. Wang, Y. Yu, Align, attend and locate: Chest x-ray diagnosis via contrast induced attention network with limited supervision, Proc IEEE Int Conf Comput Vis. (2019) 10631-10640.
40. A.F. Tredgold, Variations of ribs in the primates, with especial reference to the number of sternal ribs in man, Journal of anatomy and physiology. 31 (1897) 288.
41. M. Everingham, L.V. Gool, C.K.I. Williams, J. Winn, A. Zisserman, The pascal visual object classes (VOC) challenge, Int J Comput Vis. 88 (2010) 303-338.
42. J. Lian, J. Liu, S. Zhang, K. Gao, X. Liu, D. Zhang, Y. Yu, A Structure-Aware Relation Network for Thoracic Diseases Detection and Segmentation, IEEE transactions on medical imaging. 2021.
43. R. Padilla, S. Netto, E. Silva, A Survey on Performance Metrics for Object-Detection Algorithms, international conference on systems, signals and image processing (IWSSIP), IEEE. (2020) 237-242.
44. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, IEEE transactions on pattern analysis and machine intelligence. 39 (2016) 1137-1149.
45. J. Pang, K. Chen, J. Shi, H. Feng, W. Ouyang, D. Lin, Libra R-CNN: Towards balanced learning for object detection, Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. (2019) 821-830.

673 46. H. Zhang, H. Chang, B. Ma, N. Wang, X. Chen, Dynamic R-CNN: Towards
674 High Quality Object Detection via Dynamic Training, European Conference on
675 Computer Vision. 2020.
676 47. Z. Cai, N. Vasconcelos, Cascade R-CNN: Delving into High Quality Object
677 Detection, Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. (2018)
678 6154-6162.
679 48. A. Bochkovskiy, C.Y. Wang, H.Y.M. Liao, YOLOv4: Optimal Speed and
680 Accuracy of Object Detection, 2020.
681