



Detection and classification of mandibular fracture on CT scan using deep convolutional neural network

Xuebing Wang¹ · Zineng Xu² · Yanhang Tong¹ · Long Xia³ · Bimeng Jie¹ · Peng Ding² · Hailong Bai² · Yi Zhang¹ · Yang He¹

Received: 25 October 2021 / Accepted: 19 February 2022 / Published online: 26 February 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Objectives This study aimed to evaluate the accuracy and reliability of convolutional neural networks (CNNs) for the detection and classification of mandibular fracture on spiral computed tomography (CT).

Materials and methods Between January 2013 and July 2020, 686 patients with mandibular fractures who underwent CT scan were classified and annotated by three experienced maxillofacial surgeons serving as the ground truth. An algorithm including two convolutional neural networks (U-Net and ResNet) was trained, validated, and tested using 222, 56, and 408 CT scans, respectively. The diagnostic performance of the algorithm was compared with the ground truth and evaluated by DICE, accuracy, sensitivity, specificity, and area under the ROC curve (AUC).

Results One thousand five hundred six mandibular fractures in nine subregions of 686 patients were diagnosed. The DICE of mandible segmentation using U-Net was 0.943. The accuracies of nine subregions were all above 90%, with a mean AUC of 0.956.

Conclusions CNNs showed comparable reliability and accuracy in detecting and classifying mandibular fractures on CT.

Clinical relevance The algorithm for automatic detection and classification of mandibular fractures will help improve diagnostic efficiency and provide expertise to areas with lower medical levels.

Keywords Artificial intelligence · Deep learning · Convolutional neural network · Mandibular fracture · Computed tomography

Introduction

The mandible is the only movable bone located in a prominent part of the maxillofacial region; therefore, it is vulnerable to external forces [1, 2]. Mandibular fracture has the highest incidence in the maxillofacial region [3]. Correct diagnosis is critical for treatment decisions. The treatment protocols, surgical approach, and fixation materials are closely related to the regions and complexity of mandibular fracture. Computed tomography (CT) provides oral and maxillofacial surgeons with a more intuitive understanding about the position, displacement, and relationship of fracture fragments. Therefore, CT scan is considered to be the best diagnostic tool for fracture diagnosis [4]. However, CT contains a large amount of information. Analyzing large numbers of medical images manually is often a time-consuming and laborious process [5, 6].

Convolutional neural networks (CNNs), as the most successful application of deep learning and the most widely

Xuebing Wang and Zineng Xu contributed equally to the work and share first authorship.

✉ Yang He
fridaydust1983@163.com

- ¹ Department of Oral and Maxillofacial Surgery National Engineering Laboratory for Digital and Material Technology of Stomatology; Beijing Key Laboratory of Digital Stomatology National Clinical Research Center for Oral Diseases, Peking University School and Hospital of Stomatology, No 22 Zhongguancun South Road, Beijing 100081, People's Republic of China
- ² Deepcare, Inc, Beijing, China
- ³ Plastic Surgery Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

used model in medical image analysis, have shown excellent image processing capabilities [7]. In recent years, CNN has been widely used in fracture detection and classification on radiographs or CT scans, and some studies have proven that it has equivalent [8–11] or even superior abilities than experts [12]. Moreover, it effectively improves accuracy and decreases the diagnosis time and can be used to assist physicians in diagnosis [13, 14]. All these results have proven the feasibility of CNN in fracture detection.

This study aimed to train and validate the first CNN-based approach for the detection and classification of CT images.

Materials and methods

This retrospective study was approved by the Ethics Committee of Peking University School and Hospital of Stomatology (protocol No. PKUSSIRB-202054056) and was conducted in accordance with the relevant guidelines and regulations.

Demographics of patients

The data of all patients who underwent CT scans using a 16-slice CT scanner (Optima CT 520; GE Healthcare, Waukesha, WI) with 1.25-mm slice thickness at Peking University School and Hospital of Stomatology between January 2013 and July 2020 were extracted according to the following criteria.

The inclusion criteria were (i) Chinese, aged from 18 to 80 years; (ii) history of mandible fracture within 15 days; (iii) no history of maxillofacial tumor; (iv) no systemic bone metabolic disease; (v) no maxillofacial deformity; and (vi) no history of radiotherapy or chemotherapy. The exclusion criteria were (i) congenital facial asymmetry, such as severe jaw deviation and nasal septum deviation; (ii) a history of maxillofacial hard tissue surgery; (iii) delayed or green branch fractures; and (iv) identifiable history of mandibular surgery, bone tumor, bone metabolic disease, etc.

We finally extracted the data of 686 patients, including 497 males and 189 females, with an average age of 35.74 ± 12.86 years. The CT data used in this study were selected from a database with data desensitization technique. Patients' private information, such as name, address, and phone numbers, were masked.

Image annotation

Subregion segmentation

CT data of 278 patients were randomly selected from all included CT data and were synthesized into panoramic radiograph (refer to the following section of mandible

anatomical region segmentation and extraction). Each subregion was drawn on the panoramic radiograph using VGG Image Annotator 2.0.1 (Visual Geometry Group) according to the mandibular fracture classification by Kelly et al. [15] (Fig. 1A–C). The parameter relationship between the original CT and the synthesized panoramic image is shown in Fig. 2.

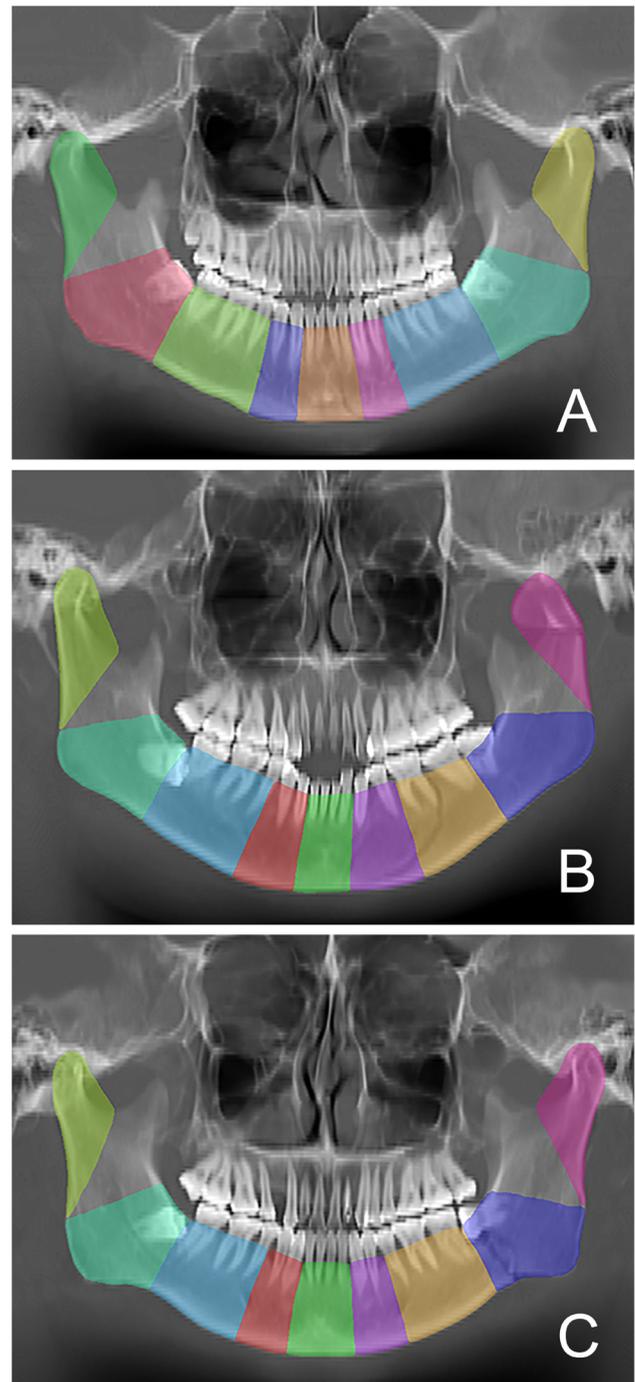


Fig. 1 Annotation of the mandibular subregions. **A** Mandibular symphysis fracture. **B** Left parasymphysis, body, and condyle fracture. **C** Left mandibular angle fracture

Mandibular fracture classification

All the CT data of the included patients were imported into Mimics Medical 21.0 (Materialise HQ Technologielaan, Leuven, Belgium) in DICOM format for observation and annotation by three experienced surgeons.

The annotated content mainly included two parts: (i) the number of axial layers where fracture lines were located and (ii) the mandibular fracture location information according to the classification by Kelly et al. (Table 1) [15].

For any fracture line that spanned more than one sub-region, the fracture was recorded in each area [16]. The majority opinion of three experienced maxillofacial surgeons was selected as the final label result.

Convolutional neural network training

The 686 annotated CT scans were divided into three subsets: a training set with 222 (32.4%) CT scans, a validation set with 56 (8.2%) CT scans, and a test set with 408 (59.5%) CT scans.

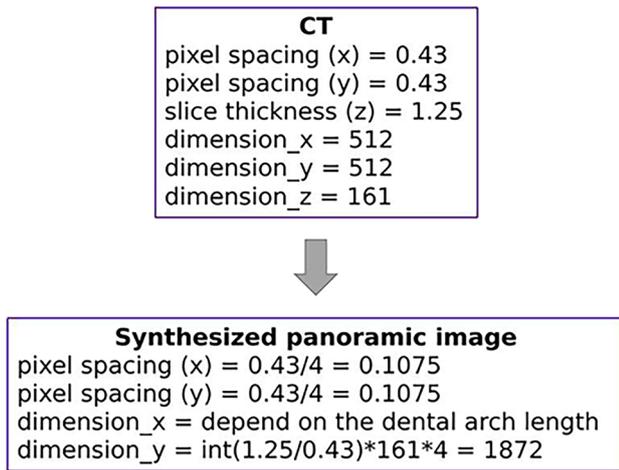


Fig. 2 The parameter relationship between original CT and synthesized panoramic images

Table 1 Mandibular subregions and definitions [15]

Fracture types	Definition
Symphysis fracture	The fracture occurred in the incisor area, extending from the alveolar process to the lower edge of the mandible
Parasymphysis fracture	The fracture occurred between the distal of mandibular incisor and the mental foramen
Body fracture	The fracture occurred between the mental foramen and the distal of the mandibular second molar
Angle fracture	The fracture occurred in the distal of mandibular second molar, the fracture lines extending from any point on the curve between the distal of mandibular second molar and the anterior edge of ramus to any point on the curve formed by the curve between mandibular lower edge and the posterior edge of the ramus
Ramus fracture	The fracture lines cross the anterior to posterior edge of the mandibular ramus horizontally or run vertically downward from the sigmoid notch to the mandibular lower edge
Condyle fracture	The fracture line extended from the sigmoid notch along the upper part of the mandibular ramus to the posterior edge of the mandibular ramus
Coronoid process fracture	The fracture occurred in the coronoid process

As shown in Fig. 3, the workflow consisted of four steps. The first step was to obtain the synthesized panoramic radiograph (Fig. 3b) and straighten the mandible region slices (Fig. 3c) from the original CT scan. The second step was to detect nine anatomical regions (middle/right/left symphysis, right/left body, right/left angle, right/left condylar process) of the mandible on the synthesized panoramic radiograph image by applying a U-Net segmentation model. Thereafter, the image patches (Fig. 3f) of these nine mandible anatomical regions were extracted based on the generated straightened mandible region slices and segmentation results. The last step was to detect the fracture by applying a patch-wise based classification model (ResNet-50). The final detection result of a specific anatomical region of the mandible was obtained by fusing multiple patch-level detection results.

Segmentation and extraction of mandible anatomical regions

The CT scans used in this study had a large field of view and covered most of the craniofacial skeleton (mandible, maxilla, zygoma, teeth, etc.). Therefore, mandible region detection was crucial for accurate mandible fracture detection. To achieve accurate localization of nine anatomical regions of the mandible, we first obtained the synthesized panoramic radiograph and corresponding straightened mandible slices of a given CT scan as previously described [17]. Then, we trained a semantic segmentation model, which automatically extracted different anatomical regions of the mandible on the synthesized panoramic radiograph. As U-Net [18] architecture has achieved outstanding performance on many medical image segmentation tasks [18, 19], we selected U-Net as the segmentation model.

The U-Net model was trained using 222 panoramic radiographs, which were synthesized from 222 training CT scans. Both the panoramic radiographs and their corresponding ground-truth annotation masks were resized to a resolution of 768 × 768 for training and validation purposes. To control

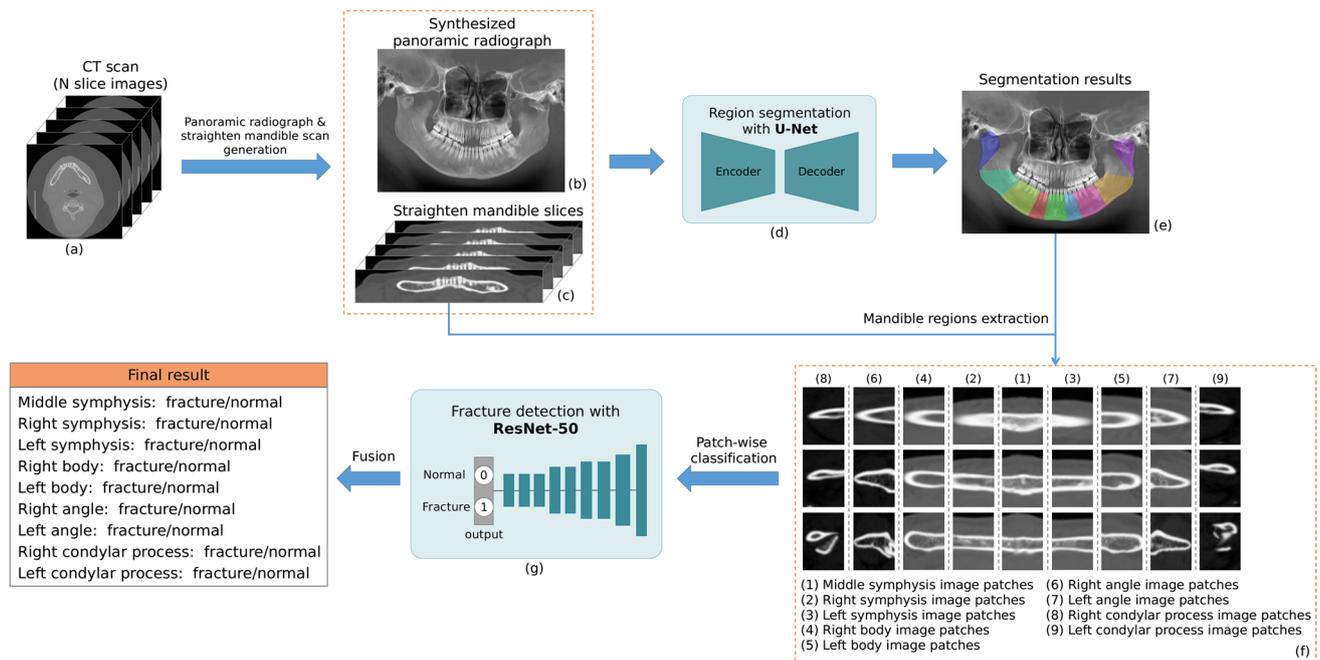


Fig. 3 The pipeline of our proposed mandibular fracture detection method

the tradeoff between the false negatives and positives while simultaneously enforcing a smooth training, we applied combo loss [20] (a weighted sum of soft Dice loss and cross-entropy loss) as the loss function. The U-Net model was trained for 50 epochs (an epoch is a single pass through the full training set) with the initial learning rate set to 0.0003 using the AdamW optimizer [21]. The batch size was set to 2 and the weight decay parameter was set to $1e-5$. The Dice coefficient [22] (DICE), also called the overlap index, is the most commonly used metric for validating medical image segmentation. The DICE is calculated as follows:

$$\text{DICE} = \frac{2(S_D \cap S_{GT})}{S_D + S_{GT}} \quad (1)$$

where S_D and S_{GT} represent the areas of the segmented mask and its corresponding ground-truth mask. The average Dice coefficient on 56 validation images was 0.943 after the model was appropriately trained.

After obtaining segmentation results of anatomical regions, we applied morphological dilation to obtain the dilated segmentation masks. Furthermore, the dilated segmentation masks were masked to the generated straightened mandible slices to obtain image patches (Fig. 3f) of different anatomical regions.

Mandible fracture detection

Once the image patches of anatomical regions were acquired, they were fed into a convolutional neural network (CNN) model for fracture detection (fracture or normal) training. Given that the semantic features of fractures in different regions are similar, we only trained one CNN model. The deep residual network (ResNet) [23] is one of the most popular architectures in various computer vision tasks. Therefore, we selected ResNet as the fracture detection model. As the fracture line can be very small in some cases, we proposed using patch-wise classification rather than 3D classification to improve the sensitivity of the model. The pre-trained ResNet-50 model on the ImageNet dataset [24] was used in this study to accelerate the convergence of the model.

The ResNet-50 model was trained on 33,966 (generated from 222 training CT scans) patch images. The patch images were padded to square and then resized to a resolution of 400×400 for training and validation purposes. To train the classification model, the cross-entropy loss was adopted as the loss function with the label smoothing technique [25]. The model was trained for 30 epochs with the initial learning rate being set to 0.0001 using the AdamW optimizer. The batch size was set to 32, and the weight decay parameter was set to $5e-4$. The F1 score was used to evaluate the performance. It can be calculated as follows:

$$\text{Precision} = \frac{N_{TP}}{N_{TP} + N_{FP}} \tag{2}$$

$$\text{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{3}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

where N_{TP} represents the number of true positives, N_{FP} represents the number of false positives, and N_{FN} represents the number of false negatives. After training, F1 of 0.9512 was achieved on 8613 validation patch images (generated from 56 validation CT scans).

At the inference phase, every extracted patch image was sequentially fed into the trained ResNet-50 model to obtain nine binary sequences (corresponding to nine anatomical regions of the mandible). For each binary sequence, we designed a simple fusion rule to generate the final fracture detection result. The rule was as follows: for each sequence, if the number of 1 was greater than K , the final result was fracture; otherwise, the result was normal. When K was set to 2, we achieved the highest F1 (0.9607) on 56×9 regions.

Results

Mandible fracture classification

The number of mandible fractures of each subregion obtained after an annotation is shown in Table 2.

Performance of CNNs

The average Dice coefficient on 56 validation images for mandibular subregion segmentation was 0.943, after the model was appropriately trained.

ResNet was used to detect mandible fracture lines in the test set, which consisted of 408 mandible fracture cases. The number of mandible fractures used for training, validating, and testing in each subregion is shown in Table 3. The area under the ROC curve (AUC), accuracy, sensitivity, and specificity were measured. The performance is summarized in Table 4. The accuracies of nine subregions were all above 90%, with an average AUC of 0.956.

Discussion

The mandible is the only movable bone in the maxillofacial region, and its position is relatively prominent. It is susceptible to fractures caused by forces including traffic accidents, attacks, falls, and sports injuries [26]. According to statistics

Table 2 Distribution of mandibular fractures according to the classification [15] in 686 patients

Fracture types	Number
Symphysis fracture	375
Left parasymphysis fracture	153
Right parasymphysis fracture	199
Left body fracture	66
Right body fracture	75
Left angle fracture	67
Right angle fracture	47
Left condyle fracture	251
Right condyle fracture	273
Left ramus fracture	24
Right ramus fracture	7
Left coronoid process fracture	13
Right coronoid process fracture	8
Total	1558

from different countries, the incidence of maxillofacial fractures varies from 24.30 to 81.28% [26–28].

The anatomical region and complexity of mandible fractures significantly affect the choice of treatment and fixation materials. Imageology is one of the most important methods to diagnose mandible fractures, and CT scanning is considered the gold standard tool for correct diagnosis [4]. Nevertheless, fracture is one of the most easily missed types in radiological diagnosis [29]. The lack of professional knowledge and long working hours can lead to human diagnosis errors [30, 31], and the daily real-time error rate can reach 3–5% [32]. However, machines are not limited in this respect, and they can make quick and accurate diagnosis even while processing a large amount of data [9]. Therefore, we wanted to use artificial intelligence to achieve efficient and accurate automatic detection of mandibular fractures based on the anatomical classification by Kelly et al. [15].

Convolutional neural network (CNN) has been used for medical image analysis since the 1990s [7]. Since the development of computer technology provides technical support for deeper neural networks, CNN can deal with more complex problems [33] and is currently a leader in image processing problems [7, 34]. Many studies have used CNN in the diagnosis and classification of fractures in recent years [8–14]. This was the first study to propose a deep learning framework to detect and classify mandibular fractures in CT scans.

In this study, U-Net was used for segmentation, and ResNet was used for fracture detection. U-Net [18] is a U-shaped architecture network proposed by Ronneberger et al. based on the “full convolutional network.” It is the most famous network currently used in medical image segmentation tasks and has shown good performance in multiple

Table 3 Distribution of mandibular fractures in nine subregions in different datasets

Fracture types	Training set (<i>N</i> = 222)		Validation set (<i>N</i> = 56)		Test set (<i>N</i> = 408)	
	Positive	Negative	Positive	Negative	Positive	Negative
Symphysis	122	100	30	26	222	186
Left parasymphysis	53	169	15	41	85	323
Right parasymphysis	58	164	14	42	126	282
Left body	24	198	7	49	35	373
Right body	29	193	9	47	37	371
Left angle	20	202	8	48	40	368
Right angle	21	201	6	50	19	389
Left condyle	77	145	17	39	158	250
Right condyle	86	136	24	32	164	244
Total	490	1508	130	374	886	2786

Table 4 Results of the test set

Fracture types	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (95% CI)
Symphysis	93.87	95.95	91.40	0.937 (0.909–0.965)
Left parasymphysis	94.36	92.06	95.39	0.931 (0.893–0.970)
Right parasymphysis	94.61	90.59	95.67	0.937 (0.906–0.968)
Left body	97.79	97.30	97.84	0.972 (0.939–1.000)
Right body	97.30	97.14	97.32	0.976 (0.944–1.000)
Left angle	98.28	94.74	98.46	0.950 (0.902–0.999)
Right angle	97.06	92.50	97.55	0.966 (0.907–1.000)
Left condyle	97.79	95.73	99.18	0.958 (0.933–0.983)
Right condyle	96.57	92.41	99.20	0.975 (0.955–0.994)
Mean AUC				0.956

bone segmentation tasks [35–37]. ResNet was used in previous studies related to fractures. Chung et al. [9] trained the ResNet to detect and classify proximal humeral fractures on the X-ray. The accuracy of fracture detection was 96%, and the CNN also showed superiority for the classification of fractures compared to the human groups. Tomita et al. [10] used a combination of recurrent neural network (RNN) and ResNet to detect osteoporotic vertebral fractures (OVFs) in CT and achieved an accuracy of 89.2%. Pranata et al. [38] compared the results of detecting calcaneal fractures in CT with ResNet and VGG, the accuracies were both 98%, and ResNet showed better performance as it has deeper neural network architecture. Olaczak et al. [39] applied a modified ResNet model to classify ankle fractures and obtained an average AUC of 0.90 (95% CI 0.82–0.94).

The mandible is an irregular bone. In this study, we used the synthesized panoramic radiographs and corresponding new CT scans of the straightened mandible in order to make more intuitive region segmentation. Thereafter, we input the new CT data into U-Net and trained for mandible subregion segmentation and ResNet for fracture line detection for obtaining the final results. The results showed that the automatic segmentation of the mandible using U-Net achieved a high DICE of 0.943. The accuracies of the mandible fracture

detection in the nine subregions were from 93.87 to 98.28%, with an average AUC of 0.956. As shown in Table 4, the results proved the feasibility of U-Net and ResNet in automatically detecting and classifying mandibular fractures in CT.

We analyzed the following possible causes for misdiagnosed cases. First, there were no clear boundaries between the adjacent subregions of the mandible. According to the classification of mandible fractures by Kelly et al. [15], the symphysis and parasymphysis are divided by distal mandibular lateral incisors, and the mental foramen divides the parasymphysis and mandibular body. The dividing lines between adjacent regions of the mandible are imaginary lines based on some anatomical landmarks (such as mental foramen, lateral incisor, the second molar), and there are no specific anatomical structures for demarcating condyles, coronoid process, and mandibular ramus. Although the algorithm segmentation results coincide with the artificial segmentation results, they cannot achieve 100% replication. Hence, there was no controversy about the fracture line that is entirely in a certain region or clearly across the dividing lines. However, the fracture lines on the dividing line cause divergence between the physician and the algorithm. In addition, in the cases of incorrect diagnosis of mandibular angle

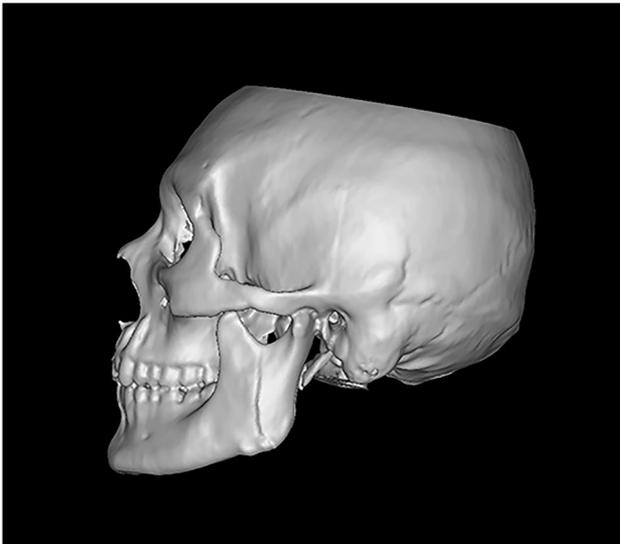


Fig. 4 Mandibular ramus fracture

fractures, we have found another pattern. The condition that the fracture lines run vertically downward from the sigmoid notch to the lower edge of the mandible is defined as the ramus fracture (Fig. 4). The lower fracture line of this type often passes through the mandibular angle region. Based on our algorithm, once a fissure is detected in a certain region, this region is considered to have a fracture. Consequently, some ramus fractures were misidentified as angle fractures.

Doctors need time to acquire professional knowledge. Using algorithms with learned professional knowledge for automatic diagnosis of fractures would be useful in regions lacking experienced doctors. The existence and location of mandibular fracture lines were illustrated in this preliminary study. On the basis of this study, the morphology of mandibular fractures will be described in future studies. Condylar fracture is a unique type of mandibular fracture whether from clinical management or the development of artificial intelligence algorithms. More factors should be considered, such as dislocation, angulation, and ramus height loss. For this reason, the work from classification to the labeling of condylar fracture and mandibular fracture is totally different, and the morphology of both will be described, respectively.

This study had several limitations. In terms of the scope of application, the data used in this study were all obtained from Peking University School and Hospital of Stomatology. This algorithm requires a comparable level of accuracy for data obtained from other machines. For deep learning, the larger the training sample size, the higher the credibility of the algorithm [5, 11, 34]. However, due to the low incidence of the coronoid process fractures and the mandibular ramus fractures, the sufficient number of datasets has not been established in this study. Furthermore, the alveolar process fracture is different from any

other site of the mandible. It refers specifically to a fractured segment that is bordered between two distinct vertical fracture lines at a variable distance from each other and by an interconnecting horizontal fracture line running through the apical base and is documented by the FDI numbers of the involved teeth, to be used for providing information about the location and extent of the fracture [40]. Therefore, this algorithm based on the segmentation of the mandibular basal bone is unsuitable for alveolar process fractures. Collecting more cases of coronoid process fractures and mandibular ramus fractures and developing new algorithms for alveolar process fractures are needed in the future.

In conclusion, this study demonstrated that CNN showed comparable reliability and accuracy in detecting and classifying mandibular fractures. The algorithm proposed in this study can be useful for the automated diagnosis and classification of mandibular fractures.

Acknowledgements The authors acknowledge the assistance of the radiological department of Peking University School and Hospital of Stomatology in Beijing, China, for database establishment.

The first author Dr. Xuebing Wang acknowledges the care and assistance from Dr. Ruiliu Li, Dr. Lihang Shen, Dr. Hang Wang, Dr. Chengyi Wang, Dr. Huiyu Peng, Dr. Xiyue Wang, and Feier Wang.

Author contribution Yang He conceived the ideas. Xuebing Wang participated in the study design and collected and annotated all the data. Zineng Xu developed the algorithm. Xuebing Wang and Zineng Xu contributed equally to the article writing. Yanhang Tong and Long Xia participated in data annotation. Bimeng Jie contributed to the polish of the article. Hailong Bai and Peng Ding provided technical support. Yi Zhang provided the research platform. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This study was supported by the National Key Research and Development Program of China (2019YFF0302401). The funding agencies had no role in study design, implementation procedures, analysis of results, or preparation of the manuscript.

Declarations

Ethics approval All the procedure was approved by the Ethics Committee of Peking University School and Hospital of Stomatology (protocol No. PKUSSIRB-202054056) and was conducted in accordance with the relevant guidelines and regulations.

Informed consent For this type of study, formal consent is not required.

Conflict of interest The authors declare no competing interests.

References

1. Chrcanovic BR, Abreu MH, Freire-Maia B, Souza LN (2012) 1,454 mandibular fractures: a 3-year study in a hospital in Belo Horizonte, Brazil. *Journal of cranio-maxillo-facial surgery* :

- official publication of the European Association for Cranio-Maxillo-Facial Surgery 40:116–123. <https://doi.org/10.1016/j.jcms.2011.03.012>
2. Boffano P, Kommers SC, Roccia F, Forouzanfar T (2015) Mandibular trauma treatment: a comparison of two protocols. *Medicina oral, patologia oral y cirugia bucal* 20:e218–e223. <https://doi.org/10.4317/medoral.20263>
 3. Boffano P, Roccia F, Zavattero E, Dediol E, Uglešić V, Kovačić Ž, Vesnaver A, Konstantinović VS, Petrović M, Stephens J, Kanzaria A, Bhatti N, Holmes S, Pechalova PF, Bakardjiev AG, Malanchuk VA, Kopchak AV, Galteland P, Mjøen E, Skjelbred P, Koudougou C, Mouallem G, Corre P, Løes S, Lekven N, Laverick S, Gordon P, Tamme T, Akermann S, Karagozoglu KH, Kommers SC, Forouzanfar T (2015) European Maxillofacial Trauma (EUR-MAT) project: a multicentre and prospective study. *Journal of cranio-maxillo-facial surgery : official publication of the European Association for Cranio-Maxillo-Facial Surgery* 43:62–70. <https://doi.org/10.1016/j.jcms.2014.10.011>
 4. He D, Zhang Y, Ellis E 3rd (2007) Panfacial fractures: analysis of 33 cases treated late. *Journal of oral and maxillofacial surgery : official journal of the American Association of Oral and Maxillofacial Surgeons* 65:2459–2465. <https://doi.org/10.1016/j.joms.2007.06.625>
 5. Shen D, Wu G, Suk HI (2017) Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng* 19:221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
 6. Kalmet PHS, Sanduleanu S, Primakov S, Wu G, Jochems A, Refaee T, Ibrahim A, Hulst LV, Lambin P, Poeze M (2020) Deep learning in fracture detection: a narrative review. *Acta Orthop* 91:215–220. <https://doi.org/10.1080/17453674.2019.1711323>
 7. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak J, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88. <https://doi.org/10.1016/j.media.2017.07.005>
 8. Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, Sköldenberg O, Gordon M (2017) Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop* 88:581–586. <https://doi.org/10.1080/17453674.2017.1344459>
 9. Chung SW, Han SS, Lee JW, Oh KS, Kim NR, Yoon JP, Kim JY, Moon SH, Kwon J, Lee HJ, Noh YM, Kim Y (2018) Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthop* 89:468–473. <https://doi.org/10.1080/17453674.2018.1453714>
 10. Tomita N, Cheung YY, Hassanpour S (2018) Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. *Comput Biol Med* 98:8–15. <https://doi.org/10.1016/j.compbiomed.2018.05.011>
 11. Adams M, Chen W, Holdorf D, McCusker MW, Howe PD, Gailard F (2019) Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures. *J Med Imaging Radiat Oncol* 63:27–32. <https://doi.org/10.1111/1754-9485.12828>
 12. Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N (2019) Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. *Skeletal Radiol* 48:239–244. <https://doi.org/10.1007/s00256-018-3016-3>
 13. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, Hanel D, Gardner M, Gupta A, Hotchkiss R, Potter H (2018) Deep neural network improves fracture detection by clinicians. *Proc Natl Acad Sci USA* 115:11591–11596. <https://doi.org/10.1073/pnas.1806905115>
 14. Zhou QQ, Wang J, Tang W, Hu ZC, Xia ZY, Li XS, Zhang R, Yin X, Zhang B, Zhang H (2020) Automatic Detection and Classification of Rib Fractures on Thoracic CT Using Convolutional Neural Network: Accuracy and Feasibility. *Korean J Radiol* 21:869–879. <https://doi.org/10.3348/kjr.2019.0651>
 15. Kelly DE and Harrigan WF (1975) A survey of facial fractures: Bellevue Hospital, 1948–1974. *Journal of oral surgery (American Dental Association : 1965)* 33:146–9.
 16. Corneliussen CP, Audigé L, Kunz C, Rudderman R, Buitrago-Téllez CH, Frodel J, Prein J (2014) The Comprehensive AOCMF Classification System: Mandible Fractures- Level 2 Tutorial. *Cranio-maxillofac Trauma Reconstr* 7:S015-30. <https://doi.org/10.1055/s-0034-1389557>
 17. Yun Z, Yang S, Huang E, Zhao L, Yang W, Feng Q (2019) Automatic reconstruction method for high-contrast panoramic image from dental cone-beam CT data. *Comput Methods Programs Biomed* 175:205–214. <https://doi.org/10.1016/j.cmpb.2019.04.024>
 18. Ronneberger O, Fischer P, Brox TJC (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation.
 19. Liu L, Cheng J, Quan Q, Wu F-X, Wang Y-P, Wang J (2020) A survey on U-shaped networks in medical image segmentations. *Neurocomputing* 409:244–258. <https://doi.org/10.1016/j.neucom.2020.05.070>
 20. Taghanaki SA, Zheng Y, Kevin Zhou S, Georgescu B, Sharma P, Xu D, Comaniciu D, Hamarneh G (2019) Combo loss: handling input and output imbalance in multi-organ segmentation. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society* 75:24–33. <https://doi.org/10.1016/j.compmedimag.2019.04.005>
 21. Loshchilov I and Hutter F (2017) Decoupled Weight Decay Regularization.
 22. Dice LR (1945) Measures of the Amount of Ecologic Association Between Species 26:297–302. <https://doi.org/10.2307/1932409>
 23. He K, Zhang X, Ren S and Sun JJI (2016) Deep Residual Learning for Image Recognition.
 24. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vision* 115:211–252. <https://doi.org/10.1007/s11263-015-0816-y>
 25. Szegeedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z (2016) Rethinking the Inception Architecture for Computer Vision.
 26. Ellis E 3rd, Moos KF, el-Attar A, (1985) Ten years of mandibular fractures: an analysis of 2,137 cases. *Oral Surg Oral Med Oral Pathol* 59:120–129. [https://doi.org/10.1016/0030-4220\(85\)90002-7](https://doi.org/10.1016/0030-4220(85)90002-7)
 27. Gassner R, Tuli T, Hächl O, Rudisch A, Ulmer H (2003) Cranio-maxillofacial trauma: a 10 year review of 9,543 cases with 21,067 injuries. *Journal of cranio-maxillo-facial surgery : official publication of the European Association for Cranio-Maxillo-Facial Surgery* 31:51–61. [https://doi.org/10.1016/s1010-5182\(02\)00168-3](https://doi.org/10.1016/s1010-5182(02)00168-3)
 28. Erol B, Tanrikulu R, Görgün B (2004) Maxillofacial fractures. Analysis of demographic distribution and treatment in 2901 patients (25-year experience). *Journal of cranio-maxillo-facial surgery : official publication of the European Association for Cranio-Maxillo-Facial Surgery* 32:308–313. <https://doi.org/10.1016/j.jcms.2004.04.006>
 29. Pinto A, Acampora C, Pinto F, Kourdioukova E, Romano L, Verstraete K (2011) Learning from diagnostic errors: a good way to improve education in radiology. *Eur J Radiol* 78:372–376. <https://doi.org/10.1016/j.ejrad.2010.12.028>
 30. Krupinski EA, Berbaum KS, Caldwell RT, Schartz KM, Kim J (2010) Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol* 7:698–704. <https://doi.org/10.1016/j.jacr.2010.03.004>

31. Brady AP (2017) Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging* 8:171–182. <https://doi.org/10.1007/s13244-016-0534-1>
32. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE (2013) Cognitive and system factors contributing to diagnostic errors in radiology. *AJR Am J Roentgenol* 201:611–617. <https://doi.org/10.2214/AJR.12.10375>
33. Erickson BJ, Korfiatis P, Akkus Z, Kline TL (2017) Machine learning for medical imaging. *radiographics : a review publication of the Radiological Society of North America, Inc* 37:505–515. <https://doi.org/10.1148/rg.2017160130>
34. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, Kadoury S and Tang A (2017) Deep learning: a primer for radiologists. *Radiographics : a review publication of the Radiological Society of North America, Inc* 37:2113–2131. doi: <https://doi.org/10.1148/rg.2017170077>
35. Deniz CM, Xiang S, Hallyburton RS, Welbeck A, Babb JS, Honig S, Cho K, Chang G (2018) Segmentation of the proximal femur from MR images using deep convolutional neural networks. *Sci Rep* 8:16485. <https://doi.org/10.1038/s41598-018-34817-6>
36. Qiu B, Guo J, Kraeima J, Glas HH, Borra RJH, Witjes MJH, van Ooijen PMA (2019) Automatic segmentation of the mandible from computed tomography scans for 3D virtual surgical planning using the convolutional neural network. *Phys Med Biol* 64:175020. <https://doi.org/10.1088/1361-6560/ab2c95>
37. Noguchi S, Nishio M, Yakami M, Nakagomi K, Togashi K (2020) Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques. *Comput Biol Med* 121:103767. <https://doi.org/10.1016/j.compbiomed.2020.103767>
38. Pranata YD, Wang KC, Wang JC, Idram I, Lai JY, Liu JW, Hsieh IH (2019) Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images. *Comput Methods Programs Biomed* 171:27–37. <https://doi.org/10.1016/j.cmpb.2019.02.006>
39. Olczak J, Emilson F, Razavian A, Antonsson T, Stark A, Gordon M (2021) Ankle fracture classification using deep learning: automating detailed AO Foundation/Orthopedic Trauma Association (AO/OTA) 2018 malleolar fracture identification reaches a high degree of correct classification. *Acta Orthop* 92:102–108. <https://doi.org/10.1080/17453674.2020.1837420>
40. Cornelius CP, Audigé L, Kunz C, Rudderman R, Buitrago-Téllez CH, Frodel J, Prein J (2014) The comprehensive AOCMF classification system: mandible fractures-level 3 Tutorial. *Cranio-maxillofac Trauma Reconstr* 7:S031–43. <https://doi.org/10.1055/s-0034-1389558>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.