ORIGINAL ARTICLE

Deep learning-based evaluation of the relationship between mandibular third molar and mandibular canal on CBCT

Mu-Qing Liu[1,2,3,4] · Zi-Neng Xu[5] · Wei-Yu Mao[1,2,3,4] · Yuan Li[1,2,3,4] · Xiao-Han Zhang[1,2,3,4] · Hai-Long Bai[5] · Peng Ding[5] · Kai-Yuan Fu[1,2,3,4]

Abstract

Objectives The objective of our study was to develop and validate a deep learning approach based on convolutional neural networks (CNNs) for automatic detection of the mandibular third molar (M3) and the mandibular canal (MC) and evaluation of the relationship between them on CBCT.

Materials and methods A dataset of 254 CBCT scans with annotations by radiologists was used for the training, the validation, and the test. The proposed approach consisted of two modules: (1) detection and pixel-wise segmentation of M3 and MC based on U-Nets; (2) M3-MC relation classification based on ResNet-34. The performances were evaluated with the test set. The classification performance of our approach was compared with two residents in oral and maxillofacial radiology.

Results For segmentation performance, the M3 had a mean Dice similarity coefficient (mDSC) of 0.9730 and a mean intersection over union (mIoU) of 0.9606; the MC had a mDSC of 0.9248 and a mIoU of 0.9003. The classification models achieved a mean sensitivity of 90.2%, a mean specificity of 95.0%, and a mean accuracy of 93.3%, which was on par with the residents.

Conclusions Our approach based on CNNs demonstrated an encouraging performance for the automatic detection and evaluation of the M3 and MC on CBCT.

Clinical relevance

An automated approach based on CNNs for detection and evaluation of M3 and MC on CBCT has been established, which can be utilized to improve diagnostic efficiency and facilitate the precision diagnosis and treatment of M3.

Keywords Mandibular third molar · Mandibular canal · CBCT · Deep learning · Convolutional neural networks

Introduction

Impacted mandibular third molar (M3) extraction, one of the most common operations in oral and maxillofacial surgery, can be associated with postoperative complications, such as pain, bleeding, swelling, opening limitation, and inferior alveolar nerve (IAN) injury, impairing the life quality of patients. The incidence of temporary IAN injury caused by M3 extraction was 0.4–8.4%, while the incidence of permanent injury is less than 1% [1, 2]. However, due to the high incidence of impacted M3, a large number of patients suffer from IAN injury caused by M3 extraction [3]. The most significant risk factor of IAN injury caused by M3 extraction is the proximity of the root of M3 to the mandibular canal (MC) [1, 2, 4, 5]. Therefore, thorough preoperative analysis and evaluation of the anatomical structures are essential before M3 extraction to minimize the IAN injury risk.

✉ Kai-Yuan Fu
kqkyfu@bjmu.edu.cn

1 Center for TMD and Orofacial Pain, Department of Oral and Maxillofacial Radiology, Peking University School and Hospital of Stomatology, No. 22 Zhong Guan Cun South Ave, Beijing 100081, People's Republic of China

2 National Clinical Research Center for Oral Diseases, Beijing, China

3 National Engineering Laboratory for Digital and Material Technology of Stomatology, Beijing, China

4 Beijing Key Laboratory of Digital Stomatology, Beijing, China

5 Deepcare, Inc, Beijing, China

The conventional panoramic radiography presents limitations in displaying the relation between M3 and IAN due to the superimposition of adjacent anatomical structures. The accuracy and specificity of predicting the exposure of IAN during the M3 extraction of panoramic radiographs were unsatisfactory [6]. Cone beam computed tomography (CBCT), a three-dimensional (3D) tomographic imaging modality, provides accurate 3D information at a lower radiation dose level than CT [7]. Researches had demonstrated that CBCT was a better radiographic method than panoramic radiography for evaluating the relationship between M3 and MC [6, 8]. Therefore, CBCT has been advocated as the choice for preoperative examination for complex M3 extraction [9].

The advantage of CBCT examination comes with the increasing workload and responsibility for diagnosis. There are more than one hundred images for a single CBCT scan. It is time-consuming to view the images slice by slice. Moreover, only a small proportion of dentists (12.4%) had experience with CBCT imaging in dental school based on a survey conducted by Buchanan et al. [10]. And most dentists have not received professional training in craniomaxillofacial imaging diagnosis [11], which increases the risk of missed diagnosis and misdiagnosis. It is desired to develop computer-aided diagnosis tools to aid CBCT diagnosis.

Nowadays, deep learning, a subset of artificial intelligence, is undergoing rapid development and has achieved significant success in medical fields. Among deep learning models, supervised learning of the convolutional neural network (CNN) is widely studied, which has been on par with or surpassed human experts' level in many medical imaging fields [12, 13]. After CNN was introduced to the dental field, it was used for the detection, classification, and segmentation of anatomical structures [14–18], the diagnosis of dental caries [19–22], periapical lesions [23], periodontal disease [24], cystic lesions and tumors [25, 26], maxillary sinusitis [27], and cephalometric analysis [28, 29]. Though these applications are still at preliminary stages, promising results have been reported. Recently, deep learning based on CNN models has been used for the M3 and MC detection and segmentation on panoramic radiographs and CBCT [15, 18, 30], the development staging [31, 32], and the angulation measurements of the M3 on panoramic radiographs [33]. Fukuda et al. compared 3 CNNs for classification of the M3-MC relation with panoramic radiographs [34]. Yoo et al. proposed a CNN-based approach to predict the difficulty of the M3 extraction using panoramic radiographs [35]. However, as aforementioned, panoramic radiography has limitations in describing anatomical structures as a two-dimensional (2D) imaging modality. Orhan et al. reported an AI application (Diagnocat, Inc.) based on CNN with high accuracy in detecting the M3 and determining the number of roots and their relation to adjacent anatomical structures

[36]. However, the details of the classification of the M3-MC relation were not elaborated in their report. The objective of our study was to develop and validate an automated approach based on CNNs for the detection of the M3 and MC and the evaluation of the relation between them in CBCT images. We presented a deep learning method based on U-Net [37] and the deep residual network (ResNet) [38].

Materials and methods

Data collection

A CBCT image dataset of 254 patients with a mean age of 29.21 ± 7.60 (age range of 15–64 years, 116 male, and 138 female) diagnosed as unilateral M3 was collected between 2018 and 2019 at the Department of Radiology, Peking University School of Stomatology. The ethical approval from the Biomedical Institutional Review Board of Peking University School of Stomatology was obtained before conducting this study (PKUSSIRB-201949131).

Before CBCT examination, all the patients received periapical film or panoramic radiography, which showed complex root morphology, a close M3-MC relationship (including darkening of the roots, interruption of the MC cortical border, diversion of MC, or superimposition of the root and MC), suspected root resorption of the mandibular second molar, or other specific situations. The CBCT images were acquired with 3D Accuitomo (J Morita Mfg. Corp., Kyoto, Japan). The scan parameters were as follows: tube potential of 85–90kVp, tube current of 5 mA, field of view (FOV) of $6 \mathrm{~cm} \times 6 \mathrm{~cm}$, and a voxel size of 0.125 mm. Slice thickness and interval were both set at 0.2 mm. The inclusion criteria included the complete and clear display of the region of interest (ROI). The exclusion criteria included blurred CBCT images caused by artefact and presence of pathologic conditions or surgical intervention of mandible. The images were exported in digital imaging and communications in medicine (DICOM) format. Patients' private information was masked.

Data annotations

Segmentation of the M3 and MC

For the segmentation task, 229 CBCT images were randomly divided into three subsets: the training set (154, 67.2%), the validation set (30, 13.1%), and the test set (45, 19.7%).

The segmentation of M3 and MC of the 229 cases was performed with a 3D medical image processing software Materialise Mimics (version 22.0; Materialise Inc., Leuven, Belgium). The semi-automatic tool, "Split Mask," was used for preliminary segmentation. Then the segmentation results

were modified manually with multi-planar reformation (MPR) by a trained radiologist, as the ground truth (Fig. 1).

Classification of the M3-MC relation

For the classification task, 254 CBCT images were randomly divided into three subsets: the training set (154, 60.6%), the validation set (30, 11.8%), and the test set (70, 27.6%). The sample size of 70 for the test set was determined a priori using software PASS version 15 (https://www.ncss.com). According to our preliminary results, the sensitivity was about 80%, and the specificity was about 90%. $N_{sensitivty}$ of 70 and $N_{specificity}$ of 44 were calculated using PASS version 15 with α of 0.05 and δ of 0.1. So a minimum total sample size of 70 was determined.

The relation between M3 and MC in CBCT was classified into three types: type I (separation), type II (contact), or type III (invagination) (Fig. 2). For all the 254 cases, the classification of the M3-MC relation was evaluated with the MPR images of Mimics software independently by two radiologists with 10-year experience in CBCT diagnosis. Each axial slice with both the M3 and MC shown was labelled with one of the three types, which was used for networks training. If each axial slice was diagnosed as type I, then the case was diagnosed as type I. As long as one axial slice was diagnosed as type III, the case was diagnosed as type III. Otherwise, it was diagnosed as type II. If there was disagreement between the two radiologists, a consensus was reached after review and discussion. The numbers of types I, II, and III were respectively 51, 52, and 51 in the training set; 10, 10, and 10 in the validation set; and 23, 24, and 23 in the test set.
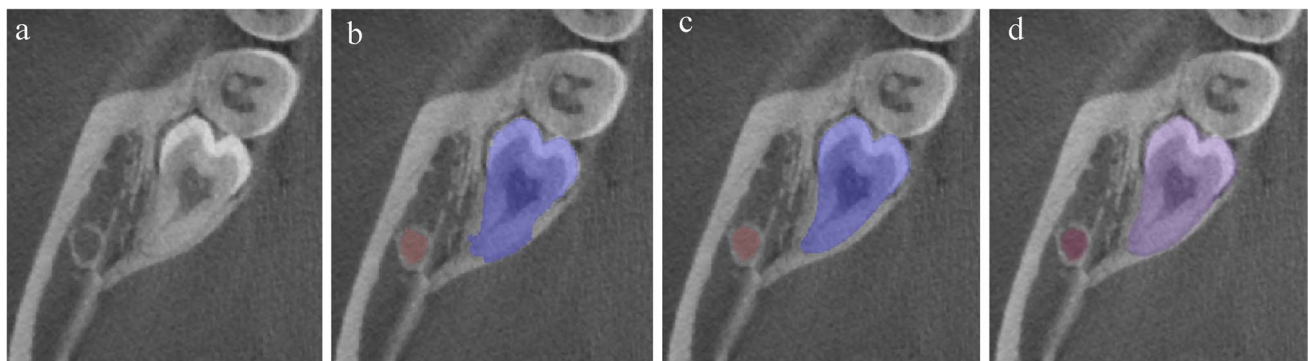
Fig. 1 Segmentation of the M3 and MC in the axial slice. **a** The original image. **b** The preliminary segmentation result using the semi-automatic tool "Split Mask" of Mimics. **c** The ground truth with manual modification of the preliminary segmentation. **d** The segmentation result of our approach based on U-Net

Fig. 2 Classification of the M3-MC relation. **a** Type I: there was the cancellous bone separation between the M3 and the cortical wall of MC. **b** Type II: the root of M3 was in contact with the cortical wall of MC, and the cortical border was still intact. **c** Type III: there was an interruption of the cortical border of MC caused by the root of M3 impinging on it, with anatomical deformation of MC at the contact area, termed "invagination"

Workflow of our deep learning approach

An overview of our proposed approach is shown in Fig. 3. The workflow of the approach consists of two modules: (A) M3 and MC detection; (B) M3-MC relation classification. The training set and validation set were used to train two U-Nets (Fig. 3A) and a ResNet-34 (Fig. 3B). In the testing phase, each CBCT scan in the test set was fed into the trained U-Nets to output detection results of M3 and MC; then the cropped images were fed into the trained ResNet-34 to obtain classification results. Due to the limited training sample, we adopted label smoothing [39], weight decay, and data augmentation to avoid overfitting. Moreover, we adopted logit adjustment softmax cross-entropy loss [40] and oversampling minority class to reduce the influence of imbalanced label distribution.

M3 and MC detection

Since MC occupies only a small portion of the CBCT images (Fig. 3a), it is difficult to achieve accurate segmentation of MC using the original images. Given the fact that MC is close to M3, we proposed a multi-step method (Fig. 3A) to crop an ROI enclosing M3 and MC from the original CBCT images and then performed the M3 and MC segmentation and classification on the cropped axial images. The sagittal maximum intensity projection (MIP) (Fig. 3b) and transverse MIP (Fig. 3e) were successively generated to obtain the location of M3 and MC in the original CBCT. Then, we could get the cropped ROI according to the detected location. The ROI in the axial image was shown in the green bounding box (Fig. 3f). The number of cropped axial slices was 1.5 times the number occupied by M3 in the sagittal MIP image (Fig. 3g).
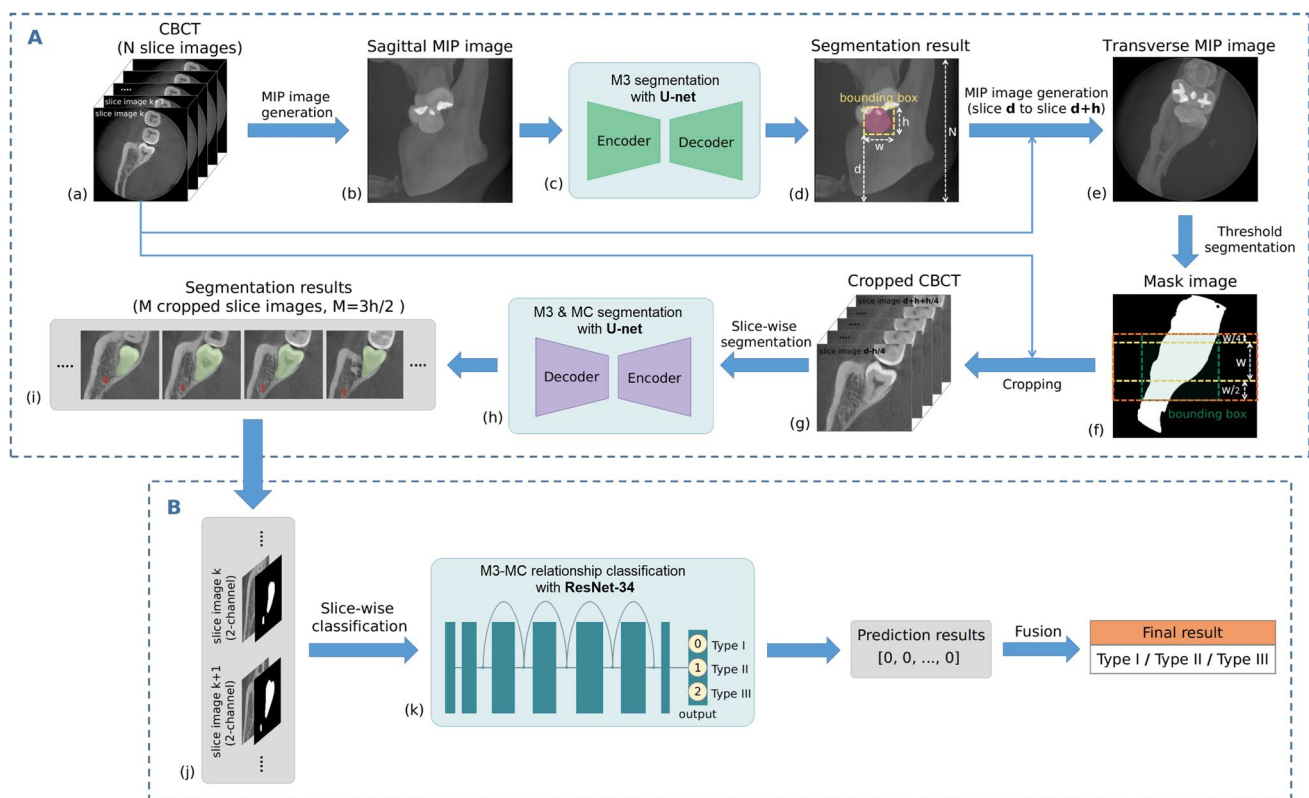
Fig. 3 Workflow of the proposed approach. The approach consisted of two modules: **A** M3 and MC detection module; **B** M3-MC relation classification module. **a, b, c** The sagittal MIP image reconstructed from CBCT was fed into the U-Net model to obtain the segmentation of M3. **d** The slice range (from slice d to slice d+h) of M3 was obtained according to the bounding box coordinates of the segmented M3. **e** A transverse MIP image from slice d to slice d+h of original axial CBCT images was generated. **f** The transverse MIP image was then passed through an Otsu threshold filter combing with morphological operations to obtain the mask image of the ROI, which contains both M3 and MC. The yellow box was converted from the bounding box in Fig. d. To cover the MC, the yellow box was extended to form the orange box. Then, we could obtain the bounding box (the green box) of the segmented mandible located in an orange box. **g** The cropped CBCT (3 h/2 slices) was obtained from the original CBCT according to the two bounding boxes (Fig. d and f). **h and i** The cropped CBCT were fed into the U-Net model to output the segmentation results of M3 and MC. **j and k** The segmented mask image of M3 and MC combined with the cropped grayscale slice image were taken as the input image of the classification model based on ResNet-34. After obtaining the classification result for each slice, a simple fusion approach was developed to obtain the final prediction

The cropped CBCT images were fed into the U-Net model to output the final segmentation results of M3 and MC. The U-Net model for M3 segmentation was trained using 154 generated MIP images. Both the MIP images and their corresponding ground truth masks were resized to a resolution of 448×448 for training and validation purposes. The U-Net model for M3 and MC segmentation was trained on 4,287 cropped axial slices of 154 cases, with an average of 27.8 slices per case. The slice images and their corresponding ground truth masks were resized to a resolution of 320×320 for training and validation purposes. To control the class imbalance of the target pixels (small) and background pixels (large), while at the same time enforcing a smooth training, we applied combo loss [41] (a weighted sum of soft Dice loss and cross-entropy loss) as the loss function. Because MC is much smaller than M3, a class weight of 1:5 was used to balance the segmentation of M3 and MC. The U-Net models were trained for 30 epochs (an epoch is a single pass through the full training set), with the initial learning rate being set to 0.0003 using the Adam optimizer [42]. The batch size was set to 4 and 16, respectively, for two U-Nets. The Dice similarity coefficient (DSC) [43] is the most common metric in validating medical image segmentation, calculated as follows:

$$DSC = \frac{2(S_D \cap S_{GT})}{S_D + S_{GT}}$$

where S_D and S_{GT} denote the areas of the segmented mask and its ground truth mask. The DSC is between 0 and 1. A value of 1 indicates a perfect segmentation that the two regions overlap completely, while a value of 0 indicates no overlap.
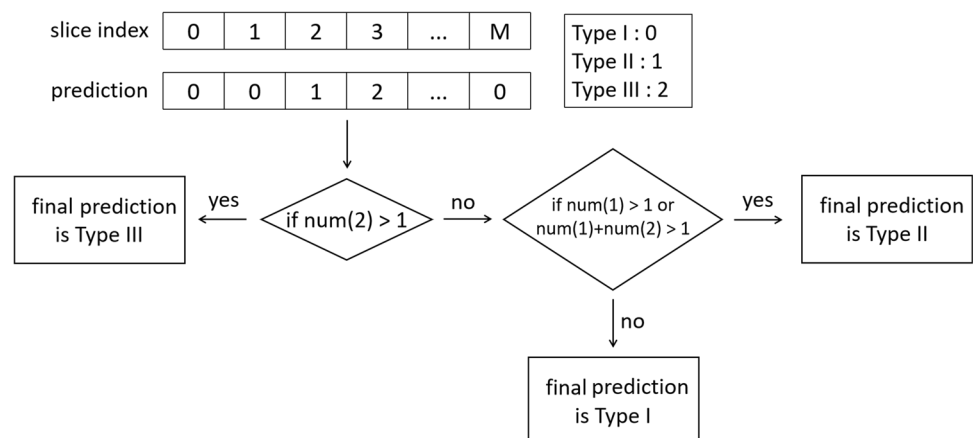
M3-MC relation classification

Intuitively, it is not hard to determine the geometric relation between M3 and MC with adequate segmentation. However, it was not trivial to distinguish between type II and type III. Therefore, we proposed to classify the three types via the supervised deep learning classification technique. We selected ResNet [38], one of the most popular CNN architectures in various computer vision tasks, as the M3-MC spatial relation classification model. Due to the limited amount of CBCT samples, we proposed using slice-wise classification rather than 3D classification to expand the size of the training set. As shown in Fig. 3j, the segmented mask image of M3 and MC combined with the cropped grayscale slice image was taken as the input images of the classification model. After obtaining the classification result for each slice, a simple fusion approach was developed to obtain the final prediction, which was illustrated in Fig. 4. This CNN-based method will be referred to as Model$_{cnn}$ hereafter. As the distance between the M3 and MC is a straightforward differential criterion for type I, we enhanced the Model$_{cnn}$ with an additional morphological classifier. For each slice, we firstly applied morphological dilation to MC segmentation with the kernel size experimentally set to 13 and then calculated the intersection between the dilated MC and M3 masks. If there is no intersection among all slices, we set the prediction as type I. Otherwise, the result of Model$_{cnn}$ was kept. We refer to this classification model as Model$_{comb}$.

As for Model$_{cnn}$, the ResNet-34 was adopted as the classification model. The pre-trained ResNet-34 parameters on the ImageNet dataset [44] were used to accelerate the convergence of the model training. The model was trained on 4,287 2-channel images. The images were resized to a resolution of 320×320 for training and validation purposes. To train the classification model, the logit adjustment softmax cross-entropy loss was adopted as the loss function. The model was trained for 20 epochs, with the initial learning rate set to 0.0002 using the Adam optimizer. The batch size was set to 16, and the weight decay parameter was set to 0.0001.

Fig. 4 Illustration of the algorithm for the final prediction of the M3-MC relation. If more than one slice was predicted as type III, the final prediction was determined as type III. If more than one slice was predicted as type II or type III, the final prediction was determined as type II. Otherwise, the prediction was determined as type I

Comparison with the residents

To evaluate the performance of the M3-MC relation classification method, two residents in oral and maxillofacial radiology were invited to conduct diagnosis on the test dataset based on the multi-planer view of Mimics software for comparison with no time limits. A training session preceded the diagnosis of the test data, which included instructions of Mimics software, presentation of 6 demonstration cases (two per type), and scoring of 30 cases (ten per type) with feedback. The 36 cases were not included in the dataset.

Statistical analysis

The DSC, intersection over union (IoU), and pixel accuracy were employed to test the automatic segmentation model. The values were described as means and standard deviations (mean ± SD). Confidence intervals (CIs) with 95% were calculated. The sensitivity, specificity, accuracy, and confusion matrix with normalization were calculated to evaluate the performance of the classification models. The definitions are as follows:

$$IoU = \frac{TP}{TP+TN+FN}$$
$$Pixel\,accuracy = \frac{TP}{TP+FP}$$
$$Sensitivity = \frac{TP}{TP+FN}$$
$$Specificity = \frac{TN}{TN+FP}$$
$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP}$$

where TP and TN denote true positive and true negative, and FP and FN denote false positive and false negative, respectively.

Weighted kappa was used to assess the reliability between the two residents. Kendall's coefficient of concordance was calculated to evaluate the consistency among the automatic models and two residents. Values of $P < 0.05$ were considered statistically significant.

Results

Segmentation performance

The proposed method exhibited > 0.92 DSC for automatic segmentation of the M3 and MC (Table 1; Fig. 1). The M3 had a mean DSC of 0.9730, a mean IoU of 0.9606, and a mean pixel accuracy of 0.9726. The MC had a mean DSC of 0.9248, a mean IoU of 0.9003, and a mean pixel accuracy of 0.9563. The MC segmentation scored lower than the M3.

Classification performance

The mean sensitivity, specificity, and accuracy for $Model_{cnn}$ were 84.6%, 92.2%, and 89.5%, respectively. $Model_{comb}$ showed higher performance than $Model_{cnn}$, with a mean sensitivity of 90.2%, a mean specificity of 95%, and a mean accuracy of 93.3%. The results were shown in Table 2 and Fig. 5. Type II had the lowest diagnostic sensitivity and highest specificity. Compared with $Model_{cnn}$, the specificity of type I and the sensitivity of type II of $Model_{comb}$ were improved. There was no difference in the diagnostic performance for type III. Both $Model_{comb}$ and Resident A achieved a sensitivity of 100% for type I. The diagnostic performance of $Model_{comb}$, including sensitivity, specificity, and accuracy, was the same as that of the better one of the two residents.

Weighted Kappa of 0.783 ($P < 0.001$) suggested good reliability between the two residents. Kendall's coefficient of concordance was 0.901 ($P < 0.001$), which showed a strong consistency among the $Model_{cnn}$, $Model_{comb}$, and two residents.

Discussion

In recent years, deep learning has gained increasing attention and rapid development in dental imaging [14–36]. Several studies reported the application of CNN in the evaluation of M3, showing promising results in tooth development staging [31, 32], prediction of M3 eruption [33], and the detection and diagnosis of M3 [34–36]. Most of them were conducted on panoramic radiographs. Vranckx et al. proposed an

Table 1 Performance of automatic segmentation of M3 and MC

	Mean ± Std. deviation (95%CI)		
	Mean DSC	Mean IoU	Mean pixel accuracy
M3	0.9730 ± 0.0164 (0.9666–0.9780)	0.9606 ± 0.0186(0.9534–0.9670)	0.9726 ± 0.0201(0.9654–0.9791)
M3 background	0.9998 ± 0.0001 (0.9997–0.9998)	0.9995 ± 0.0002(0.9994–0.9996)	0.9999 ± 0.0001(0.9999–0.9999)
MC	0.9248 ± 0.0330 (0.9121–0.9359)	0.9003 ± 0.0332(0.8891–0.9115)	0.9563 ± 0.0197(0.9497–0.9628)
MC background	0.9999 ± 0.0001 (0.9998–0.9999)	0.9997 ± 0.0002(0.9996–0.9997)	0.9998 ± 0.0001(0.9998–0.9999)

Table 2 Performance of M3-MC relation classification

|	Sensitivity (%)	Specificity (%)	Accuracy (%)
Model$_{cnn}$			
Type I	95.7	83.0	87.1
Type II	62.5	97.8	85.7
Type III	95.7	95.7	95.7
Mean	84.6	92.2	89.5
Model$_{comb}$			
Type I	100.0	91.5	94.3
Type II	75.0	97.8	90.0
Type III	95.7	95.7	95.7
Mean	90.2	95.0	93.3
Resident A			
Type I	100.0	95.7	97.1
Type II	75.0	97.8	90.0
Type III	95.7	91.5	92.9
Mean	90.2	95.0	93.3
Resident B			
Type I	91.3	85.1	87.1
Type II	62.5	87.0	78.6
Type III	78.3	93.6	88.6
Mean	77.4	88.6	84.8

approach based on a fully CNN with ResNet-101 backbone to predict the eruption of M3 through the accurate measurement of the molar angulations on panoramic radiographs [33]. Fukuda et al. compared the performance of 3 CNNs for classification of the M3-MC relation with panoramic radiographs, with the highest area under the receiver operating characteristic curve (AUC) values ranging from 0.88 to 0.93 [34]. Yoo et al. proposed the deep learning model based on ResNet-34 to estimate the difficulty of the M3 removal surgery on panoramic radiographs using the Pederson difficulty score, which achieved accurate prediction results for the depth and angulation of M3 and the ramal relationship (79%, 90%, and 82%) [35].

There is some controversy about whether CBCT would affect the treatment plan of impacted third molar and reduce postoperative IAN injury. According to a statement by Matzen and Berkhout, a previous meta-analysis showed that CBCT did not reduce the incidence of IAN injury [45]. However, recently, some studies suggested that CBCT could help predict the occurrence of IAN injury [46, 47] and change the treatment plan of the high-risk cases [48, 49]. CBCT can elucidate the anatomical structures of M3 and MC and increase the practitioners' confidence before the extraction [48], especially for those who do not have extensive clinical experience. Of course, CBCT could reveal many factors related to the complexity of M3 extraction. In this study, we only evaluated the proximity of the M3 to MC, which is of great clinical concern. For some patients,

the M3-MC relation viewed in CBCT images was not as close as on apical film or panoramic radiograph, as shown in Figs. 1 and 2a, which was due to the overlap of images on 2D films. For them, the use of CBCT would result in fewer coronectomy decisions. Orhan et al. reported a deep CNN-based AI application with high performance in detecting the M3 and determining the number of roots and their relation to adjacent anatomical structures in CBCT, in good agreement with the manual detection (kappa: 0.762) [36]. However, the details of the classification of the M3-MC relation were not elaborated. In our study, we proposed and validated a CNN-based approach, achieving high accuracy of pixel-wise segmentation of M3 and MC, and the classification of the relation between them.

Automated recognition and segmentation of the teeth and MC in CBCT is a fundamental step of computer-aided diagnosis and treatment planning. Manual annotation of ROI is labor-intensive and time-consuming. Automatic segmentation of teeth and MC in CBCT images is challenging due to the imaging features of CBCT and the anatomic structure of the mandible. Compared to multi-slice spiral CT, CBCT had a lower radiation dose but lower contrast resolution and lower signal-to-noise ratio [7]. The segmentation of MC is more difficult than the teeth. The dimension of MC is relatively small, and there are many variations in MC shape and the mandible texture. MC often has a blurry or missing cortical border in CBCT images. In that case, it is difficult to delineate MC from the surrounding cancellous bone. Deep learning, as a data-driven approach, outperformed the previous atlas-based segmentation methods and statistical shape methods for mandibular canal segmentation [30]. Kwak et al. reported MC segmentation models based on multiple CNNs, of which the 3D U-Net showed the highest mean IoU of 0.577 [15]. Jaskari et al. reported a full MC segmentation approach based on fully CNN with a diverse dataset of CBCT images, with a DSC of 0.57–0.58 [30]. We achieved a high segmentation accuracy of MC with the mean DSC of 0.925 and mean IoU of 0.900. However, compared to their studies, we segmented only a portion of the MC around the M3, owing to the purpose of this study and the limitation of the relatively small FOV.

Our proposed approach based on ResNet-34 demonstrated a satisfying diagnostic performance with an accuracy of 93.3%, which was in strong agreement with the residents in oral and maxillofacial radiology. Although we obtained a good segmentation result, it was still not sufficient to be directly used to distinguish type II from type III using the morphological dilation method. As shown in Fig. 2b and c, root tips are in contact with the MC for both type II and type III. As for type I, there was a certain amount of false positive diagnoses even by residents. In addition to the interference caused by the unclear images mentioned above, it might be also due to the fact that for type II and type III, the root tip
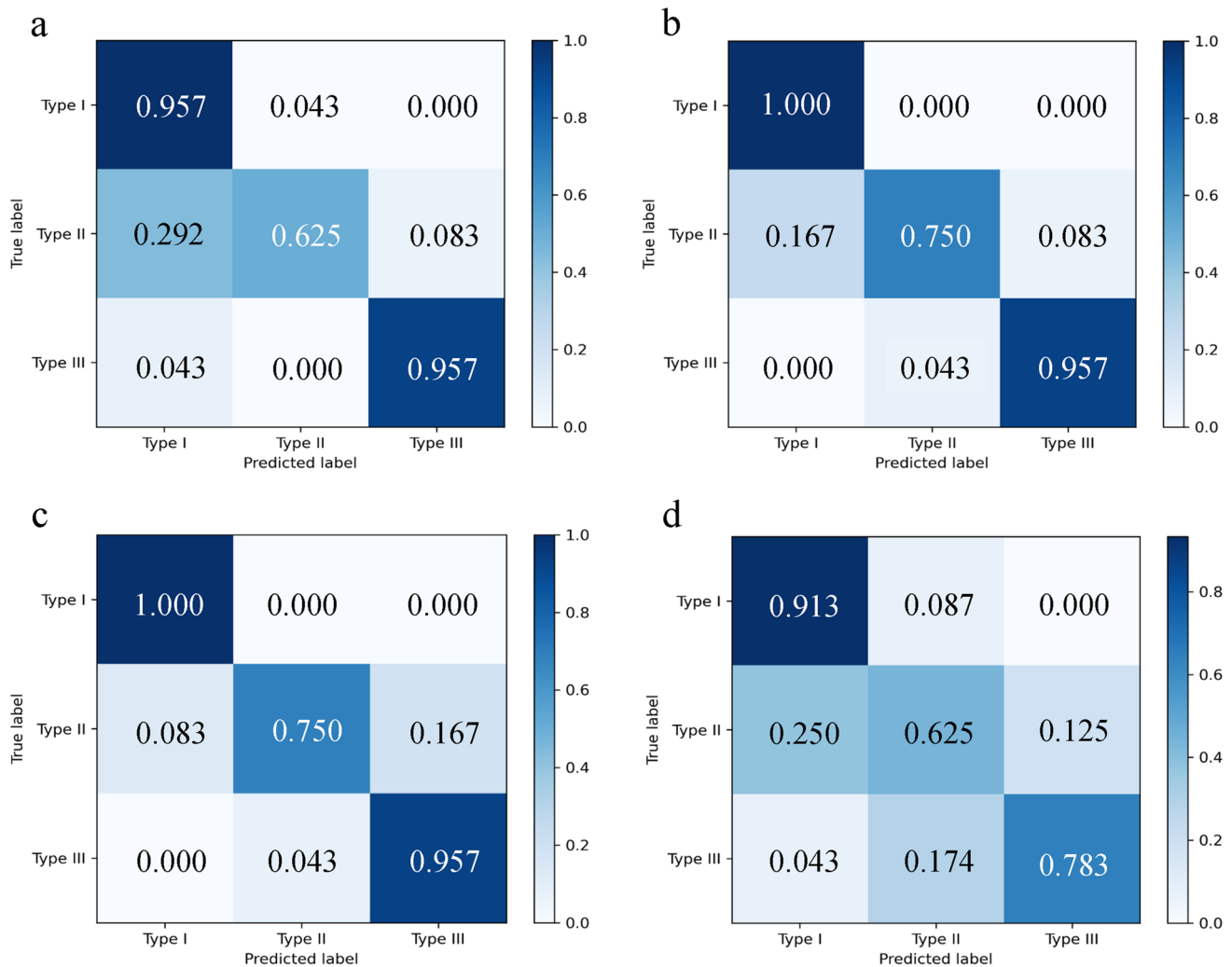
Fig. 5 Confusion matrix with normalization, showing the classification results of the M3-MC relation. **a** Model$_{cnn}$; **b** Model$_{comb}$; **c** Resident A; **d** Resident B

of M3 was in contact with MC at only a few slices, which could easily be missed when performing multi-planer viewing and lead to a misdiagnosis as type I. We explored different models for classifying the M3-MC relation in CBCT images. Model$_{comb}$ improved the diagnostic specificity of type I and sensitivity of type II based on the segmentation results of the M3 and MC. By combining the ResNet-based classification model with the morphological dilation method based on segmentation results, we obtained better diagnostic performance.

Our approach based on deep learning could perform the accurate detection and diagnosis of M3 and MC quickly. In the present study, we did not record the exact duration of time used to manually segment the M3 and MC and classify the M3-MC relation. The manual segmentation was performed using a semi-automatic tool of Mimics software, followed by a manual modification, which usually took

1–2 hrs. In the diagnostic test, there was no time limit for the residents, which usually took several minutes. In contrast, the models based on deep learning took 6.1 s $\pm$ 1.0 s to segment the M3 and MC and 7.4 $\pm$ 1.0 s to classify the M3-MC relation (including the segmentation time) per case. Clearly, the models based on deep learning had a great advantage in terms of efficiency.

Automatic diagnosis of the M3-MC relation will aid in the preoperative risk assessment and the surgical plan, preventing the postoperative IAN injury. The technique of automatic image segmentation combined with 3D display will allow the dentists to visualize the M3 and MC clearly and intuitively, saving the time spent on the slice-by-slice reading of CBCT images in the routine workflow. In addition, it could also be used in the image-guided removal of deeply impacted M3 using a navigation system. On the other hand, it might be helpful in doctor-patient communication.

Although the high diagnostic accuracy was obtained, there are some limitations in our study. First, one major limitation is that we only collected images from one CBCT facility with fixed parameters. CBCT may exhibit variation with different brands, models, and exposure conditions. More diverse datasets from different CBCT brands and models with varying exposure conditions will be gathered in future study to improve the generalization ability and robustness of our CNN models. Second, the current diagnostic model only evaluates the M3-MC relation. More factors, such as the direction of the obstruction, the depth of impaction, and the morphology of roots, need to be added and analyzed to thoroughly evaluate the difficulty and risk of M3 extraction.

Conclusion

Our proposed models based on CNNs demonstrated an encouraging performance for detecting M3 and MC and classifying the relation between them in CBCT, which can be utilized to improve diagnostic efficiency and facilitate the precision diagnosis and treatment of M3.

Declarations

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Board of Peking University School and Hospital of Stomatology and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the Institutional Review Board of Peking University School and Hospital of Stomatology (PKUSSIRB-201949131).

Informed consent Written informed consent was not required for this study because all the included patients were collected retrospectively. Exemption of informed consent will not affect the rights and health of included patients. The application for free informed consent has been approved by the Institutional Review Board.

Conflict of interest The authors declare no competing interests.

References

1. Gulicher D, Gerlach KL (2001) Sensory impairment of the lingual and inferior alveolar nerves following removal of impacted mandibular third molars. Int J Oral Maxillofac Surg 30:306–312. https://doi.org/10.1054/ijom.2001.0057
2. Leung YY, Cheung LK (2011) Risk factors of neurosensory deficits in lower third molar surgery: a literature review of prospective studies. Int J Oral Maxillofac Surg 40:1–10. https://doi.org/10.1016/j.ijom.2010.09.005
3. Ghaeminia H, Meijer GJ, Soehardi A, Borstlap WA, Mulder J, Berge SJ (2009) Position of the impacted third molar in relation to the mandibular canal. Diagnostic accuracy of cone beam computed tomography compared with panoramic radiography. Int J Oral Maxillofac Surg 38:964–971. https://doi.org/10.1016/j.ijom.2009.06.007
4. Tay AB, Go WS (2004) Effect of exposed inferior alveolar neurovascular bundle during surgical removal of impacted lower third molars. J Oral Maxillofac Surg 62:592–600. https://doi.org/10.1016/j.joms.2003.08.033
5. Kim JW, Cha IH, Kim SJ, Kim MR (2012) Which risk factors are associated with neurosensory deficits of inferior alveolar nerve after mandibular third molar extraction? J Oral Maxillofac Surg 70:2508–2514. https://doi.org/10.1016/j.joms.2012.06.004
6. Tantanapornkul W, Okouchi K, Fujiwara Y, Yamashiro M, Maruoka Y, Ohbayashi N, Kurabayashi T (2007) A comparative study of cone-beam computed tomography and conventional panoramic radiography in assessing the topographic relationship between the mandibular canal and impacted third molars. Oral Surg Oral Med Oral Pathol Oral Radiol 103:253–259. https://doi.org/10.1016/j.tripleo.2006.06.060
7. De VW, Casselman J, Swennen GR (2009) Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. Int J Oral Maxillofac Surg 38:609–625. https://doi.org/10.1016/j.ijom.2009.02.028
8. Patel PS, Shah JS, Dudhia BB, Butala PB, Jani YV, Macwan RS (2020) Comparison of panoramic radiograph and cone beam computed tomography findings for impacted mandibular third molar root and inferior alveolar nerve canal relation. Indian J Dent Res 31:91–102. https://doi.org/10.4103/ijdr.IJDR_540_18
9. Ghaeminia H, Meijer GJ, Soehardi A, Borstlap WA, Mulder J, Vlijmen OJ, Berge SJ, Maal TJ (2011) The use of cone beam CT for the removal of wisdom teeth changes the surgical approach compared with panoramic radiography: a pilot study. Int J Oral Maxillofac Surg 40:834–839. https://doi.org/10.1016/j.ijom.2011.02.032
10. Buchanan A, Thachil K, Haggard C, Kalathingal S (2017) Predoctoral and postdoctoral education on cone-beam computed tomography. J Evid Based Dent Pract 17:310–316. https://doi.org/10.1016/j.jebdp.2017.05.002
11. Rumpa G, Aruna R (2014) Systematic interpretation of CBCT scans: why do it? J Mass Dent Soc 62:68–70
12. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402–2410. https://doi.org/10.1001/jama.2016.17216
13. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118. https://doi.org/10.1038/nature21056
14. Hiraiwa T, Ariji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, Fujita H, Ariji E (2019) A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. Dentomaxillofac Radiol 48:20180218. https://doi.org/10.1259/dmfr.20180218
15. Kwak GH, Kwak EJ, Song JM, Park HR, Jung YH, Cho BH, Hui P, Hwang JJ (2020) Automatic mandibular canal detection using a deep convolutional neural network. Sci Rep 10:5711. https://doi.org/10.1038/s41598-020-62586-8
16. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, Fujita H (2017) Classification of teeth in cone-beam CT using

deep convolutional neural network. Comput Biol Med 80:24–29. https://doi.org/10.1016/j.compbiomed.2016.11.003

17. Silva G, Oliveira L, Pithon M (2018) Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives. Expert Syst Appl 107:15–31. https://doi.org/10.1016/j.eswa.2018.04.001

18. Vinayahalingam S, Xi T, Berge S, Maal T, de Jong G (2019) Automated detection of third molars and mandibular nerve by deep learning. Sci Rep 9:9007. https://doi.org/10.1038/s41598-019-45487-3

19. Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-Rizzini A, Schürmann F, Krejci I, Markram H (2019) Caries detection with near-infrared transillumination using deep learning. J Dent Res 98:1227–1233. https://doi.org/10.1177/0022034519871884

20. Lee JH, Kim DH, Jeong SN, Choi SH (2018) Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 77:106–111. https://doi.org/10.1016/j.jdent.2018.07.015

21. Patil S, Kulkarni V, Bhise A (2019) Algorithmic analysis for dental caries detection using an adaptive neural network architecture. Heliyon 5:e01579. https://doi.org/10.1016/j.heliyon.2019.e01579

22. Srivastava M, Kumar P, Pradhan L, Varadarajan S (2017) Detection of tooth caries in bitewing radiographs using deep learning. arXiv preprint arXiv:1711.07312. https://arxiv.org/abs/1711.07312

23. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, Schwendicke F (2019) Deep learning for the radiographic detection of apical lesions. J Endod 45:917-922.e5. https://doi.org/10.1016/j.joen.2019.03.016

24. Lee JH, Kim DH, Jeong SN, Choi SH (2018) Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 48:114–123. https://doi.org/10.5051/jpis.2018.48.2.114

25. Lee JH, Kim DH, Jeong SN (2020) Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. Oral Dis 26:152–158. https://doi.org/10.1111/odi.13223

26. Kwon O, Yong TH, Kang SR, Kim JE, Huh KH, Heo MS, Lee SS, Choi SC, Yi WJ (2020) Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network. Dentomaxillofac Radiol 49:20200185. https://doi.org/10.1259/dmfr.20200185

27. Murata M, Ariji Y, Ohashi Y, Kawai T, Fukuda M, Funakoshi T, KiMase Y, Nozawa M, Katsumata A, Fujita H, Ariji E (2019) Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. Oral Radiol 35:301–307. https://doi.org/10.1007/s11282-018-0363-7

28. Arık S, Ibragimov B, Xing L (2017) Fully automated quantitative cephalometry using convolutional neural networks. J Med Imaging (Bellingham) 4:014501. https://doi.org/10.1117/1.jmi.4.1.014501

29. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J (2020) Automated skeletal classification with lateral cephalometry based on artificial intelligence. J Dent Res 99:249–256. https://doi.org/10.1177/0022034520901715

30. Jaskari J, Sahlsten J, Järnstedt J, Mehtonen H, Karhu K, Sundqvist O, Hietanen A, Varjonen V, Mattila V, Kaski K (2020) Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes. Sci Rep 10:5842. https://doi.org/10.1038/s41598-020-62321-3

31. Banar N, Bertels J, Laurent F, Boedi RM, De Tobel J, Thevissen P, Vandemeulen D (2020) Towards fully automated third molar development staging in panoramic radiographs. Int J Legal Med 134:1831–1841. https://doi.org/10.1007/s00414-020-02283-3

32. Merdietio Boedi R, Banar N, De Tobel J, Bertels J, Vandermeulen D, Thevissen PW (2020) Effect of lower third molar segmentations on automated tooth development staging using a convolutional neural network. J Forensic Sci 65:481–486. https://doi.org/10.1111/1556-4029.14182

33. Vranckx M, Van Gerven A, Willems H, Vandemeulebroucke A, Ferreira Leite A, Politis C, Jacobs R (2020) Artificial intelligence (AI)-driven molar angulation measurements to predict third molar eruption on panoramic radiographs. Int J Environ Res Public Health 17:3716. https://doi.org/10.3390/ijerph17103716

34. Fukuda M, Ariji Y, Kise Y, Nozawa M, Kuwada C, Funakoshi T, Muramatsu C, Fujita H, Katsumata A, Ariji E (2020) Comparison of 3 deep learning neural networks for classifying the relationship between the mandibular third molar and the mandibular canal on panoramic radiographs. Oral Surg Oral Med Oral Pathol Oral Radiol 130:336–343. https://doi.org/10.1016/j.oooo.2020.04.005

35. Yoo JH, Yeom HG, Shin WS, Yun JP, Lee JH, Jeong SH, Lim HJ, Lee J, Kim BC (2021) Deep learning based prediction of extraction difficulty for mandibular third molars. Sci Rep 11:1954. https://doi.org/10.1038/s41598-021-81449-4

36. Orhan K, Bilgir E, Bayrakdar IS, Ezhov M, Gusarev M, Shumilov E (2020) Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans. J Stomatol Oral Maxillofac Surg S2468–7855(20):30303–30307. https://doi.org/10.1016/j.jormas.2020.12.006

37. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

38. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90

39. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826. https://doi.org/10.1109/CVPR.2016.308

40. Menon AK, Jayasumana S, Rawat AS, Jain H, Veit A, Kumar S (2020) Long-tail learning via logit adjustment. arXiv preprint arXiv:2007.07314. https://arxiv.org/abs/2007.07314

41. 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. Comput Med Imaging Graph 75:24–33. https://doi.org/10.1016/j.compmedimag.2019.04.005

42. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980. https://arxiv.org/abs/1412.6980

43. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302. https://doi.org/10.2307/1932409

44. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, BergLi ACFF (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115:211–252. https://doi.org/10.1007/s11263-015-0816-y

45. Matzen LH, Berkhout E (2019) Cone beam CT imaging of the mandibular third molar: a position paper prepared by the European Academy of DentoMaxilloFacial Radiology (EADMFR). Dentomaxillofac Radiol 48(5):20190039. https://doi.org/10.1259/dmfr.20190039

46. Qi W, Lei J, Liu YN, Li JN, Pan J, Yu GY (2019) Evaluating the risk of post-extraction inferior alveolar nerve injury through the relative position of the lower third molar root and inferior alveolar canal. Int J Oral Maxillofac Surg 48:1577–1583. https://doi.org/10.1016/j.ijom.2019.07.008

47. Sklavos A, Delpachitra S, Jaunay T, Kumar R, Chandu A (2021) Degree of compression of the inferior alveolar canal on cone-beam computed tomography and outcomes of postoperative nerve injury in mandibular third molar surgery. J Oral Maxillofac Surg 79:974–980. https://doi.org/10.1016/j.joms.2020.12.049

48. Szalma J, Vajta L, Lovász BV, Kiss C, Soós B, Lempel E (2020) Identification of specific panoramic high-risk signs in impacted third molar cases in which cone beam computed tomography changes the treatment decision. J Oral Maxillofac Surg 78:1061–1070. https://doi.org/10.1016/j.joms.2020.03.012

49. Mendonça LM, Gaêta-Araujo H, Cruvinel PB, Tosin IW, Azenha MR, Ferraz EP, Oliveira-Santos C, Tirapelli C (2020) Can diagnostic changes caused by cone beam computed tomography alter the clinical decision in impacted lower third molar treatment plan? Dentomaxillofac Radiol 50(4):20200412. https://doi.org/10.1259/dmfr.20200412