Original Article

Anatomical Tracing of Mandibular Third Molars using Deep-Learning Models

Sanicha Yaklai[1], Raknatee Chokluechai[2], Chaloemrit Phrueksotsai[1], Chaklam Silpasuwanchai[2], Soranun Chantarangsu[3], Phonkit Sinpitaksakul[4]

[1]Faculty of Dentistry, Department of Oral and Maxillofacial Surgery, Chulalongkorn University, Bangkok, Thailand
[2]School of Engineering and Technology, Asian Institute of Technology, Pathumthani, Thailand
[3]Faculty of Dentistry, Department of Oral Pathology, Chulalongkorn University, Bangkok, Thailand
[4]Faculty of Dentistry, Department of Radiology, Chulalongkorn University, Bangkok, Thailand

Abstract

The aim of this study was to compare the performance of Convolutional Neural Network (CNN) models in tracing mandibular third molars in panoramic radiographs for surgical difficulty planning. This is a retrospective, cross-sectional study. A total of 343 panoramic radiographs of innovative anatomical landmarks of impacted mandibular third molars and related structures were annotated by a specialist. These were divided into two groups: 276 films for training and 67 films for validation. Four CNN models—AlexNet, ResNet18, ResNet50, and VGG16—were used for comparison with the annotations of the specialist. Loss function, mean squared error and tracing results were used for evaluation. Although ResNet50 achieved a low loss function of 0.001 in the validation set, suggesting it might be the best model, Mean square errors and the tracing results of AlexNet and VGG16 outperform the other models. None of the models generate correct superimposition. The application of deep learning models in this domain is not yet ready for clinical use due to incorrect superimposition tracing; however, they show potency for future development. Further studies are essential.

Keywords: Convolutional neural network, Deep learning, Impacted tooth, Mandibular third molar, Panoramic radiograph

Received date: Dec 2, 2025 **Revised date:** Mar 12, 2026 **Accepted date:** Apr 3, 2026
Doi:

Correspondence to:
Sanicha Yaklai, Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Chulalongkorn University, 34 Henri Dunant Road, Wangmai, Pathumwan, Bangkok 10330, Thailand. Telephone & Fax: (+66)-2-218-8581 Email: sanicha.y@chula.ac.th

Introduction

Mandibular third molar impaction is the most common tooth impaction in the oral cavity. The difficulty of surgically removing the mandibular third molar depends on various factors, such as the space between the anterior ascending ramus and the distal surface of the mandibular second molar, as well as the vertical level of the mandibular third molar relative to the occlusal plane. The criteria for classifying the difficulty of mandibular third molar removal are Pell and Gregory's classification[1] and Winter's classification[2], which are based on panoramic radiographs. Moreover, other factors can influence the difficulty, including the root curvature, root size, number of roots, distance between

the root apex and the mandibular canal, and the width of the periodontal ligament space.[3] However, applying these criteria may require personnel with the ability to analyze the films; and the process to analyze the film seems to be time-consuming.

Convolutional Neural Network (CNN) is a deep learning architecture specialized for processing n-dimensional data, such as images.[4] It works by hierarchically extracting features through layers designed to mimic visual perception, starting with patterns for example edges, textures, shadows and progressing to complex structures. Most CNNs work for image processing in Computer Visions such as object or face detection. Currently, CNNs are used in various fields, including dentistry, for detecting dental caries[5] and periodontal diseases[6], the Kennedy classification of partially edentulous arches[7] as well as classifying jaw relationship types.[8] There are many CNN models that can be used in analyzing the mandibular third molar in the panoramic radiograph,[9-13] but no single models outperform others. Research exploring the use of CNNs for classifying the difficulty of impacted mandibular third molar in radiographic images have not been extensively developed. One study attempted to compare the difficulty score from CNN reading with the specialist reading.[12] While their findings demonstrated good performance, it was observed that their scoring system only has three levels, which may not provide adequate information for surgical planning and cannot fully illustrate the difficulty pattern. Other studies focused on three domains related to difficulty: molar-ramus relationship, depth, and angulation.[10,11] However, it did not include other relevant domains that contribute to difficulty and potential complications, such as the root-inferior alveolar canal (IAC) relationship and root morphology. Moreover, the current research gap is a lack of suitable datasets for segmentation and the inability of existing deep learning models to provide explainable outputs hinder their practical application in classifying mandibular third molar impaction difficulty. These knowledge gaps have driven the design of our study, which aimed to use CNNs for tracing mandibular third molars and related structures that incorporate the existing classification and root-inferior alveolar canal relationship. Four common CNN models—AlexNet, ResNet18, ResNet50, and VGG16— were further evaluated for this tracing task.

Materials and Methods

Dataset and Deep Learning Model Processing

243 panoramic radiographs of patients with at least one side with a mandibular third molar were collected from the Faculty of Dentistry, Chulalongkorn University. The films were randomly obtained from four machines, i.e., CS 8000C (Carestream Dental LLC, Atlanta, USA), CS 9000C (Carestream Health, NY, USA), Veraview Epocs 3D (J. Morita Corporation, Kyoto, Japan) and NewTom GiANO HR (Cefla s.c., Imola, Italy) according to their user manuals. The images were processed and cropped into the area that was used for training and validating the model.

The region of interest (ROI) was identified from a pilot study using the following steps: cropping the images, reviewing the results with an oral & maxillofacial surgeon (OMFS, SY) who is Thai board certified with ten years of experience, and repeating this process until all the dataset images including the mandibular second and third molar and related structures, such as anterior ascending ramus and inferior alveolar canal. The ROI centers for the right tooth were 0.265 in the width and 0.650 in the height. The ROI centers for the left tooth were 0.735 in the width and 0.650 in the height. The ROI square images had 720 pixels for the width and height (Fig. 1). The cropped images without a mandibular third molar were excluded. Ultimately, 343 cropped images were used as datasets.

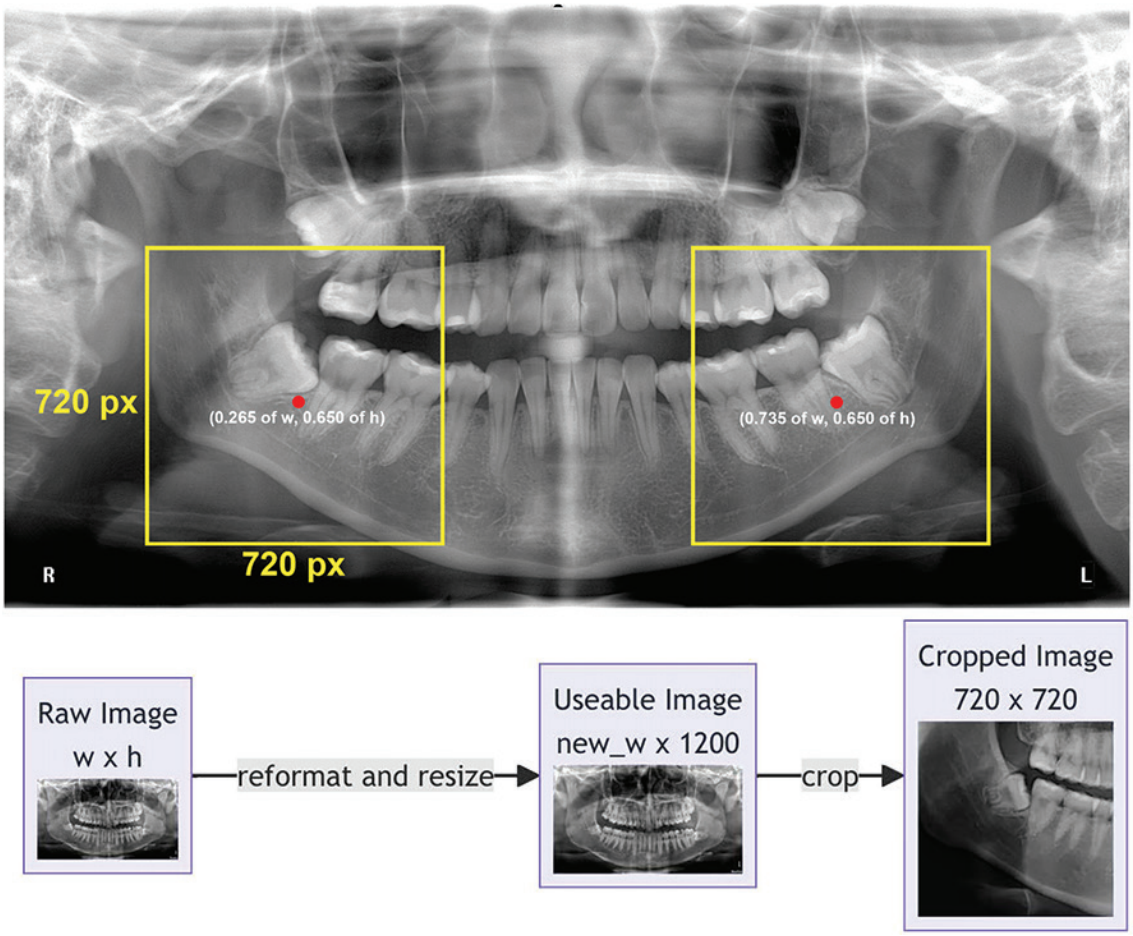

Figure 1 *Panoramic radiographs were processed and cropped. The ROI centers for the left tooth size were 0.265 in the width and 0.650 in the height. The ROI centers for the right tooth size were 0.735 in the width and 0.650 in the height. The ROI square images have 720 pixels for width and height.*

The researchers used an image array as the input to the machine learning (ML) model, and the output of the model consisted of landmark points. The original ML architecture was modified by customizing the fully connected layers. Specifically, the custom fully connected layer was revised by replacing its activation function with the rectified linear unit function and incorporating layer normalization and dropout. These modifications were intended to adapt the classification model for regression tasks.

In the current study, Python 3.12 and PyTorch 2.3.1 were used. The ML models including AlexNet, ResNet18, ResNet50, and VGG16 were implemented in Python. Prior to training, all data preprocessing—including tasks such as image resolution resizing—was automated using Python scripts. Data augmentation included random Gaussian blur (5x5 kernel, μ=0.01, σ=0.25) and brightness variation (±20%).[11] These augmentations aimed to simulate inter-patient and imaging variability. Moreover, a sigma range of (0.5, 0.7) to simulate realistic image variations and improve model robustness. For training, the hyperparameters were set as follows: learning rate of 1e-4, L1 regularization rate of 7e-3, L2 regularization rate of 7e-3, and a dropout rate of 0.5 to prevent overfitting. These settings helped enhance generalization by smoothing input images and controlling model complexity during training.

The deep learning model was trained using fine tuning with pretrained CNN models. The pretrained weights from the ImageNet-1K_V1 dataset for model initialization. ImageNet-1K_V1, also known as ILSVRC 2012, is a widely used subset of the larger ImageNet dataset designed for benchmarking computer vision models.[14] The convolutional parameters and fully connected parameters were trained. There were four deep learning models in this experiment comprising AlexNet, ResNet18, ResNet50 and VGG16.

In the experiment, 343 cropped films were split into a training set (276 films) and a validation set (67 films). The training and validation sets were randomly divided manually before being trained. An Adam optimizer was used with a small initial learning rate. To prevent overfitting, weight decay and regularization were applied. The model was trained with a mini-batch size of 64 for 201 epochs, with early stopping implemented when the validation loss converged. Training was performed on an NVIDIA RTX 4070Ti Super.

Anatomical landmark determination

Twenty-nine landmarks (Table 1) were determined by the OMFS. All the landmarks are related to the classification of the mandibular third molar surgical removal difficulty. The landmarks were determined manually and measured using the ImageJ Schneider CA program15 (Fig. 2). To ensure standardization, manual tracing was previously performed by two examiners (SY and CP) during a pilot study.

Table 1 *Anatomical landmarks of the mandibular third molars and related structures.*

Area of Interest	Landmarks
Mandibular third molar	1 Mesial cusp tip
	2 Distal cusp tip
	3 Widest point at mesial crown
	4 Widest point at distal crown
	5 Mesial cemento-enamel junction
	6 Distal cemento-enamel junction
	7 Widest point at mesial root
	8 Widest point at distal root
	9 Mesial root apex
	10 Distal root apex
	11 Point indicated root morphology[a]
	12 Root furcation
	13 Central pit[b]
	21 Highest point where anterior ascending ramus crosses the mandibular third molar.
	22 Point indicated incomplete root formation of mesial root[c]
	23 Point indicated incomplete root formation of distal root[c]
Mandibular second molar	14 Widest point at distal crown
	15 Mesial cusp tip
	16 Distal cusp tip
	17 Mesial cemento-enamel junction
	18 Distal cemento-enamel junction
	19 Central pitb
	20 Root furcation
Inferior alveolar canal (IAN)	24 Upper border of IAN located anterior to the roots of mandibular third molar
	25 Upper border of IAN located near the roots of mandibular third molar
	26 Upper border of IAN located posterior to the roots of mandibular third molar
	27 Lower border of IAN located anterior to the roots of mandibular third molar
	28 Lower border of IAN located near the roots of mandibular third molar
	29 Lower border of IAN located posterior to the roots of mandibular third molar

[a]*In case of root separation, no. 11 will be the same point as no. 12. If the roots do not separate, no. 11 will be at the midpoint where the two roots are attached.*

[b]*The deepest point of the occlusal surface between the mesial and distal cusp tips*

[C]*In case of complete root formation, no. 22 and no. 23 will be the same points as no. 9 and no. 10, respectively.*

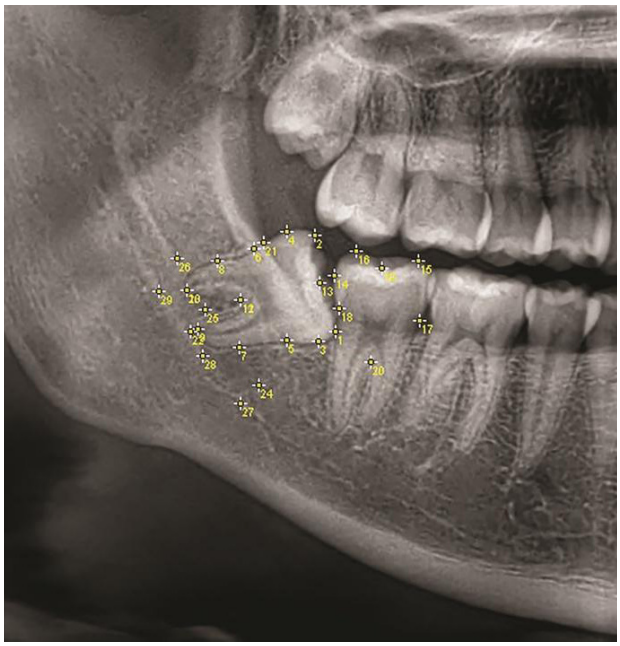

Figure 2 *29 landmarks were determined and located by the OMFS.*

Tracing the Models

The validation sets of four models were traced by connecting each selected reference point using an automated Python script and were sent to the OMFS. We selected the points that referred to three anatomical structures comprising the mandibular third molar (1 → 3 → 5 → 7 → 9 → 12 → 10 → 8 → 6 → 4 → 2 → 13 → 1), mandibular second molar (15 → 17 →20 → 18 → 14 → 16 → 19 → 15) and inferior alveolar canal (24 → 25 → 26 and 27 → 28 → 29). Four landmarks which are nos. 11, 21, 22, 23 were excluded for tracing visualization. Then, the accuracy of the superimpositions was categorized into four levels based on criteria established by the OMFS as Acceptable tracing or not, correct side or not and correct superimposition or not. The definition of the categorization is shown in Table 2.

Table 2 *The tracing categorization by the OMFS.*

Categorization	Definition
Acceptable tracing	The model generates a tracing of all three structures—the mandibular second molar, the mandibular third molar, and the inferior alveolar canal—in proper alignment.
Correct side	The model generates a tracing corresponding to the same side as the teeth and canal in the film
Correct superimposition	The model generates a tracing that matches the position of the teeth and canal in the film

Statistical analysis

Prior to the experiment, inter-observer reliability was established between two OMFS specialists (SY and CP) through the evaluation of 29 landmarks in 15 randomly selected radiographs. Intra-observer reliability was assessed by one OMFS (SY) at 15-day intervals. The intraclass correlation coefficients (ICCs) were then calculated, resulting in values of 0.973 for inter-observer reliability and 0.934 for intra-observer reliability, indicating excellent agreement.

The 29 landmarks were compared between the OMFS and the models using Mean Squared Error (MSE), and Loss function (L_{Total}) were calculated using the following equations:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (1)$$

Where:

N is the number of data points

y_i is the actual value for the i-th data point

$\hat{y}_i$ is the predicted value for the i-th data point

$$L_{Total} = MSE + \lambda\sum_{j=1}^{M}|w_j| \qquad (2)$$

Where:

L_{Total} is the overall loss function

λ is the regularization parameter

w_j is the parameters of the model

Differences of MSEs among the models were analyzed using one way Analysis of Variance (ANOVA)

followed by the Bonferroni Post hoc test. The tracing results between the models were analyzed using the Chi-square test.

The data were analyzed using SPSS version 29.0 (IBM Corp., Armonk, N.Y., USA). *P*-value < 0.05 indicated statistically significant difference.

Study Protocol and Consent Form Approval

This project was approved by The Human Research Ethics Committee of the Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand (HREC-DCU 2022-057)

Results

After randomization, the distribution of mandibular third molar in a training set (276 films) were 124, 60, 70 and 22 of mesio-, horizontal, vertical and disto- angulation, respectively. A validation set (67 films) were 31, 14, 15 and 7 of mesio-, horizontal, vertical and disto- angulation, respectively. L_{Total} results of the training and validation sets were shown in Table 3. The L_{Total} for the ResNet50 validation set was 0.001, which shows the best value among other models. The MSEs of AlexNet, ResNet18, ResNet50, and VGG16 were 0.0213 ± 0.0130, 0.0194 ± 0.0154, 0.0186 ± 0.0136 and 0.0074 ± 0.0067, respectively. The MSE results of the validation sets were shown in Figure 3. The MSE of VGG-16 was significantly different from other models. The validation set tracings were evaluated and categorized by the OMFS (Table 4). VGG-16 and AlexNet performed better in the 'Acceptable tracing' and 'Correct side' categories than ResNet18 and ResNet50. However, none of the models generated a tracing with correct superimposition. Examples of the validation sets categorized by the OMFS are shown in Figure 4.

Table 3 *Loss function (LTotal) Results of the Training set and Validation Set.*

Pretrain Model	Training set	Validation set
VGG16	0.025	0.008
AlexNet	0.027	0.008
ResNet18	0.029	0.008
ResNet50	0.032	0.001

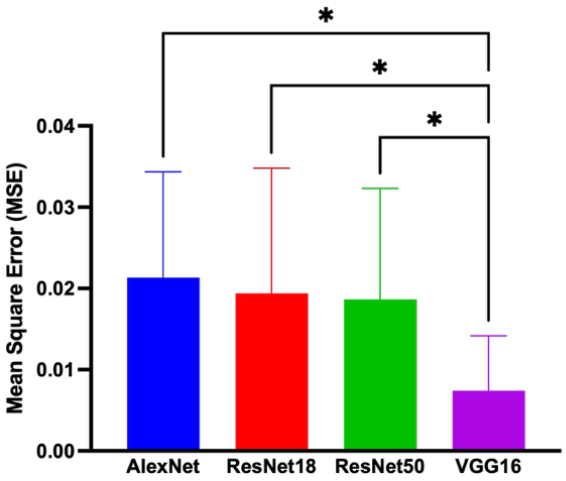

Figure 3 *Mean Square Error (MSE) Results of the Validation Set Among Four Models. Asterisk demonstrated significant difference (p-value < 0.001)*

Table 4 *Tracing Results of Validation Sets.*

	VGG16	AlexNet	ResNet18	ResNet50
Acceptable Tracing	64 (95.5%) $_a$	64 (95.5%) $_a$	66 (98.5%) $_a$	48 (71.6%) $_b$
Correct side	65 (97%) $_a$	66 (98.5%) $_a$	34 (50.7%) $_b$	50 (74.6%) $_c$
Acceptable tracing and correct side	63 (94%) $_a$	64 (95.5%) $_a$	33 (49.3%) $_b$	35 (52.2) $_b$
Correct superimposition	0 (0%) $_a$	0 (0%) $_a$	0 (0%) $_a$	0 (0%) $_a$

Different subscript letters indicate a statistically significant difference between models in the same row (p < 0.05).

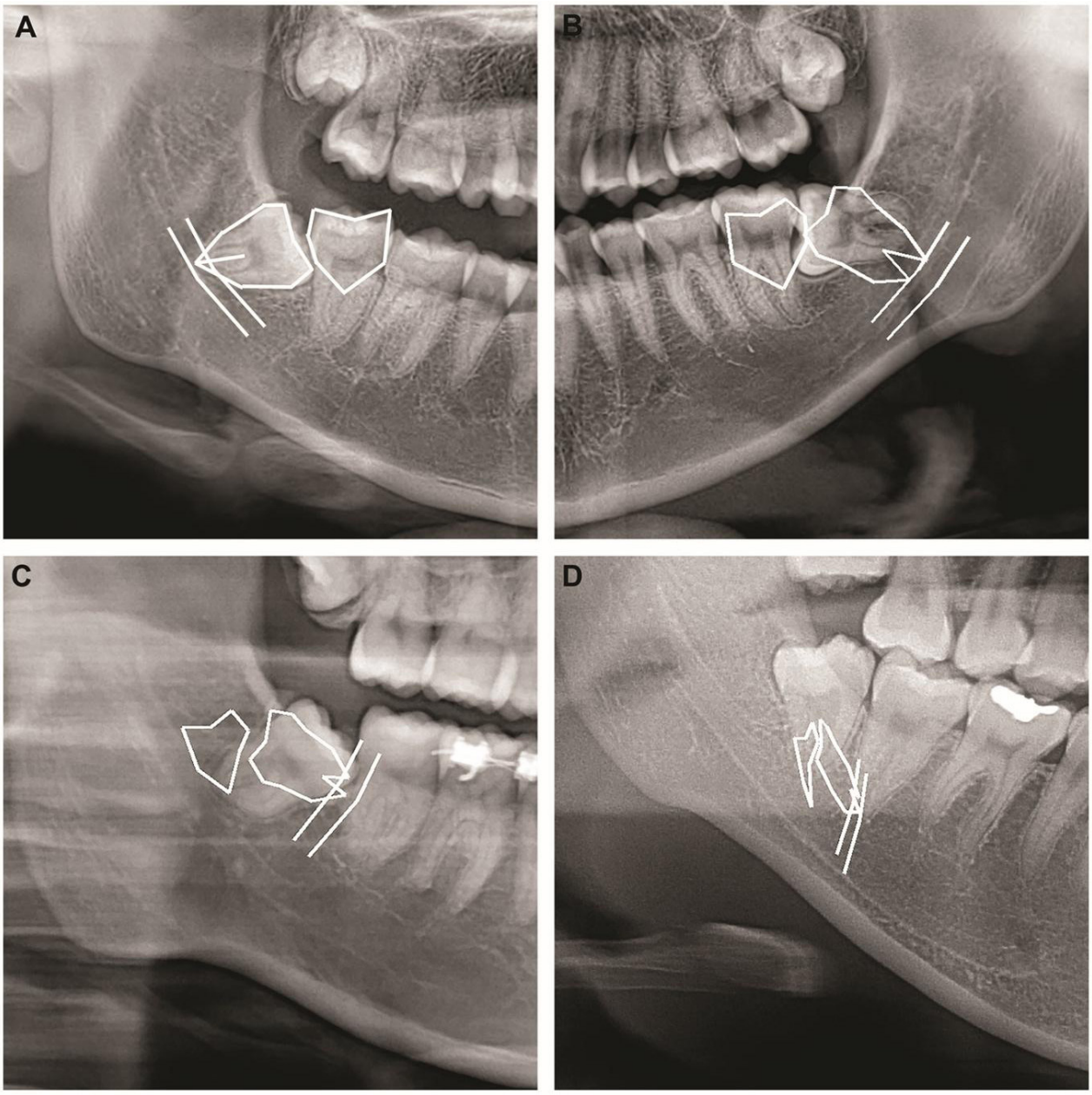

Figure 4 *The examples of tracing validation sets categorized by OMFS. Correct superimposition: A gold-standard reference achieved via the OMFS manual landmark determination (A), Acceptable tracing and correct side with incorrect superimposition (B), Acceptable tracing with incorrect side (C), Unacceptable tracing and incorrect side (D).*

Discussion

This study introduced innovative anatomical landmarks and anatomical tracing of mandibular third molar and surrounding structures related to surgical difficulty for CNN models to train. This novel method may overcome black-box predictions and enhance clinical utility. All landmarks were established based on significant anatomical markers used in difficulty classification systems.[1-3] For instance, Winter's classification utilizes the root furcation and central pit of the mandibular second and third molars (landmarks 12, 13, 19, 20) to determine angulation. The ramal relationship in the Pell and Gregory classification relies on the mesial and distal widest points of the mandibular third molar crown, as well as the distance from the distal widest point of the second molar crown to the highest point where the anterior ascending ramus crosses the third molar (landmarks 3, 4, 14, 21) to assess

space availability. Additionally, landmarks 9, 10, 22, and 23 are employed to evaluate root formation. While the concept of anatomical landmark identification was inspired by lateral cephalometric tracing[16,17], several of these landmarks are prone to geometric distortion, which may increase task complexity. Consequently, utilizing a reduced or hierarchical landmark set may enhance diagnostic performance and clinical interpretability.

The selection of CNN models in this study was informed by previous literature[9,11] and our available computational resources. Although AlexNet, VGG16, and ResNet were initially designed for image classification, their effectiveness as hierarchical feature extractors is well-established for spatial tasks through transfer learning. These architectures were selected to establish a robust baseline by leveraging proven pre-trained weights; furthermore, their utility extends to landmark-based regression tasks, demonstrating their versatility across diverse computer vision applications.[18]

The present study found that ResNet50, in which the L_{Total} of the validation set was 0.001, appeared to be the best CNN model for tracing mandibular third molar impaction compared with AlexNet, ResNet18, and VGG16. However, the MSE of VGG16 is significantly lower value than the other models. Furthermore, the tracing results demonstrated that VGG16 and AlexNet were similar in performance and significantly better than ResNet50. While ResNet18's performance in illustrating molars and IAC is comparable to AlexNet and VGG16, its side identification is inferior. After the researchers verified their performances, the ResNet50 model was found to exhibit signs of overfitting. These findings imply that relying on a single parameter is insufficient to identify the optimal CNN model for analyzing mandibular third molars in real-world scenarios.

VGG16 and AlexNet demonstrated the ability to identify patterns associated with mandibular third molars. Notably, the models correctly categorized the orientation of panoramic radiographs (left vs. right sides) despite the absence of explicit directional labels. The deep learning algorithms consistently identified third molar features across different radiographic equipment types, suggesting potential for standardized dental image analysis. This autonomous recognition of spatial orientation indicates that the models learned to utilize anatomical landmarks and radiographic cues to determine position, a capability that could assist in streamlining clinical diagnostic workflows. However, none of the models provide a correct superimposition in the validation set. This finding demonstrated that the models trained in this study failed to be applicable in clinical situations.

There were several limitations in this study. First, the sample size was restricted due to unforeseen technical and resource constraints during the model implementation stage; consequently, the dataset remained fixed at 343 images. While the study protocol initially aimed for expansion, the current results represent the finalized analytical framework, and further data augmentation or synthetic expansion could not be performed. Based on previous studies[10-13], the number of datasets should be over thousand. Second, the datasets had unequal numbers of mesio-, horizontal, vertical and disto-angulated molars, which should be addressed in subsequent studies. However, this class imbalance may not have as significant an impact as the overall small sample size, primarily because this study utilized CNNs for landmark identification and tracing rather than for classification. Lastly, Annotations in this study were performed using the ImageJ program, a method that entails more manual steps for pipeline integration and can be difficult to reproduce. These limitations can be attributed to the inadequate accuracy of the tracing results and will be a subject of investigation in future studies.

The use of anatomical landmarks and tracing was explored to classify the difficulty and visualize mandibular third molars. This approach differs from another study that employed a difficulty scoring system.[12] While their ML demonstrated good accuracy, the concept of difficulty was not truly visualized. Furthermore, this study also incorporates landmarks related to root morphology and the root-IAC relationship. These factors appear to contribute to the perceived difficulty and potential complications, an aspect that might not be fully addressed in other studies.[10,11] A direct landmark-based regression approach[19] was utilized to obtain discrete geometric coordinates for all 29 landmarks, providing an efficient input for our secondary difficulty

classification. However, it was noted that some landmarks, such as numbers 24 through 29 which define the IAC, represent a continuous anatomical structure. While direct regression is computationally lean, heatmap-based CNN architectures might better preserve the anatomical continuity and spatial relationships of the canal by leveraging local image context.[20] Addressing these complex structural dependencies remains a valuable direction for future optimization. To improve clinical reliability, future iterations will transition toward heatmap-based architectures and the implementation of connectivity-preserving loss functions. These methods leverage local image context to ensure that linear structures are predicted as a single, uninterrupted path, providing a more intuitive and reliable visual guide for surgical planning. While heatmap-based architectures often offer superior spatial localization by preserving high-resolution feature maps[21], coordinate-based regression was chosen due to the specific nature of our landmark identification task. These architectures provide a more direct mapping from global image features to coordinate outputs, thereby reducing computational overhead and post-processing complexity compared to high-resolution heatmap generation.[22,23]

The inability of the models to generate tracings with correct superimposition represents a notable limitation of this study. Successful superimposition is critical, as it facilitates a comprehensive classification of mandibular third molar impactions and permits the precise evaluation of the molar-ramus relationship, impaction depth, and angulation. Furthermore, such accuracy enables the assessment of other vital diagnostic parameters—specifically root morphology and anatomical proximity to the IAC which are decisive factors in determining surgical difficulty and predicting potential complications. Although this objective was not fully realized within the scope of the present investigation, the findings establish a foundational framework for future longitudinal research.

Conclusion

The current study introduced a new idea for analyzing the mandibular third molars and surrounding structures related to surgical difficulty classifications. However, the result did not meet expectations. By addressing the limitations, further refinement of the deep learning approach and subsequent studies may achieve more robust and generalizable results. It is suggested to enhance the study by incorporating additional feature extraction engineering. Specifically, this may include the integration of multi-modal data, the utilization of larger and more balanced datasets, and the exploration of advanced model architectures. These improvements are expected to further augment the performance and generalizability of the proposed approach. Additionally, the knowledge gained from this research provides a foundation for future investigations, potentially accelerating progress in this area of study before clinical application can be considered.

Acknowledgements

This research project is supported by grants for development of new faculty staff, Ratchadaphiseksomphot Fund, Chulalongkorn University (DNS-65045320021).

Conflict of Interest

The authors declare no conflict of interest.

References

1. Pell G, Gregory B. Impacted mandibular third molars: classification and modified techniques for removal. ***Dent Digest*** 1933;39:330–8.

2. Winter G. Principles of Exodontia as Applied to the Impacted Mandibular Third Molar : A Complete Treatise on the Operative Technic with Clinical Diagnoses and Radiographic Interpretations. St. Louis: American Medical Books; 1926.

3. Yuasa H, Kawai T, Sugiura M. Classification of surgical difficulty in extracting impacted third molars. ***Br J Oral Maxillofac Surg*** 2002;40(1):26-31.

4. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. ***J Big Data*** 2021;8(1):53.

5. Choi J, Eun H, Kim C. Boosting Proximal Dental Caries Detection via Combination of Variational Methods and Convolutional Neural Network. ***J Signal Process Syst*** 2018;90(1):87-97.

6. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, *et al.* Deep Learning for the Radiographic Detection of Periodontal Bone Loss. ***Sci Rep*** 2019;9(1):8495.

7. Khurshid Z, Waqas M, Hasan S, Kazmi S, Faheemuddin M. Deep Learning Architecture to Infer Kennedy Classification of Partially Edentulous Arches Using Object Detection Techniques and Piecewise Annotations. ***Int Dent J*** 2025;75(1):223-35.

8. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J. Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. ***J Dent Res*** 2020;99(3):249-56.
9. Celik ME. Deep Learning Based Detection Tool for Impacted Mandibular Third Molar Teeth. ***Diagnostics*** 2022;12(4)942.
10. Sukegawa S, Matsuyama T, Tanaka F, Hara T, Yoshii K, Yamashita K, *et al.* Evaluation of multi-task learning in deep learning-based positioning classification of mandibular third molars. ***Sci Rep*** 2022;12(1):684.
11. Yoo JH, Yeom HG, Shin W, Yun JP, Lee JH, Jeong SH, *et al.* Deep learning based prediction of extraction difficulty for mandibular third molars. ***Sci Rep*** 2021;11(1):1954.
12. Trachoo V, Taetragool U, Pianchoopat P, Sukitporn-udom C, Morakrant N, Warin K. Deep Learning for Predicting the Difficulty Level of Removing the Impacted Mandibular Third Molar. ***Int Dent J*** 2025;75(1):144-50.
13. Chindanuruks T, Jindanil T, Cumpim C, Sinpitaksakul P, Arunjaroensuk S, Mattheos N, *et al.* Development and validation of a deep learning algorithm for the classification of the level of surgical difficulty in impacted mandibular third molar surgery. ***Int J Oral Maxillofac Surg*** 2025;54(5):452-60.
14. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, *et al.* ImageNet Large Scale Visual Recognition Challenge. ***Int J Comput Vis*** 2015;115(3):211-52.
15. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. ***Nat Methods*** 2012;9(7):671-5.
16. Durão APR, Morosolli A, Pittayapat P, Bolstad N, Ferreira AP, Jacobs R. Cephalometric landmark variability among orthodontists and dentomaxillofacial radiologists: a comparative study. ***Imaging Sci Dent*** 2015;45(4):213-220.
17. Lindner C, Wang C-W, Huang C-T, Li C-H, Chang S-W, Cootes TF. Fully Automatic System for Accurate Localisation and Analysis of Cephalometric Landmarks in Lateral Cephalograms. ***Sci Rep*** 2016;6(1):33581.
18. Zhao X, Wang L, Zhang Y, Han X, Deveci M, Parmar M. A review of convolutional neural networks in computer vision. ***Artif Intell Rev*** 2024;57(4):99.
19. Laitenberger F, Scheuer HT, Scheuer HA, Lilienthal E, You S, Friedrich RE. Cephalometric landmark detection using vision transformers with direct coordinate prediction. ***J Cranio-Maxillofac Surg*** 2025;53(9):1518-29.
20. Payer C, Štern D, Bischof H, Urschler M. Regressing heatmaps for multiple landmark localization using CNNs. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Proceedings, Part II; 2016 Oct 17-21; Athens, Greece. Cham: Springer; 2016. p. 230-238. (Lecture Notes in Computer Science; vol. 9901).
21. Bulat A, Tzimiropoulos G. How far are we from solving the 2D & 3D Face Alignment problem? (and a dataset of 230,000 3D facial landmarks). In: Proceedings of the IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22-29; Venice, Italy. IEEE; 2017. p. 1021-1030.
22. Yin S, Wang S, Chen X, Chen E, Liang C. Attentive One-Dimensional Heatmap Regression for Facial Landmark Detection and Tracking. In: Proceedings of the 28th ACM International Conference on Multimedia; 2020 Oct 12-16; Seattle, WA, USA. New York: ACM; 2020. p. 1436-1444.
23. Feng Zhen-Hua, Kittler J, Awais M, Huber P, Wu Xiao-Jun. Wing Loss for Robust Facial Landmark Localisation with Convolutional Neural Networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2018 Jun 18-23; Salt Lake City, UT, USA. IEEE; 2018. p. 2235-2244.