The Saudi Dental Journal 36 (2024) 404–412

Contents lists available at ScienceDirect

The Saudi Dental Journal

journal homepage: www.ksu.edu.sa
www.sciencedirect.com

Review Article

The diagnostic performance of impacted third molars in the mandible: A review of deep learning on panoramic radiographs

Amalia Nur Faadiya [a], Rini Widyaningrum [b,*], Pingky Krisna Arindra [c], Silviana Farrah Diba [b]

[a] Dental Medicine Study Program, Faculty of Dentistry, Universitas Gadjah Mada, Yogyakarta, Indonesia
[b] Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Universitas Gadjah Mada, Yogyakarta, Indonesia
[c] Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Universitas Gadjah Mada, Yogyakarta, Indonesia

ARTICLE INFO

Keywords:
Mandibular canal
Radiograph
Panoramic
Deep learning
Third molar
Impacted

ABSTRACT

Background: Mandibular third molar is prone to impaction, resulting in its inability to erupt into the oral cavity. The radiographic examination is required to support the odontectomy of impacted teeth. The use of computer-aided diagnosis based on deep learning is emerging in the field of medical and dentistry with the advancement of artificial intelligence (AI) technology. This review describes the performance and prospects of deep learning for the detection, classification, and evaluation of third molar-mandibular canal relationships on panoramic radiographs.
Methods: This work was conducted using three databases: PubMed, Google Scholar, and Science Direct. Following the literature selection, 49 articles were reviewed, with the 12 main articles discussed in this review.
Results: Several models of deep learning are currently used for segmentation and classification of third molar impaction with or without the combination of other techniques. Deep learning has demonstrated significant diagnostic performance in identifying mandibular impacted third molars (ITM) on panoramic radiographs, with an accuracy range of 78.91% to 90.23%. Meanwhile, the accuracy of deep learning in determining the relationship between ITM and the mandibular canal (MC) ranges from 72.32% to 99%.
Conclusion: Deep learning-based AI with high performance for the detection, classification, and evaluation of the relationship of ITM to the MC using panoramic radiographs has been developed over the past decade. However, deep learning must be improved using large datasets, and the evaluation of diagnostic performance for deep learning models should be aligned with medical diagnostic test protocols. Future studies involving collaboration among oral radiologists, clinicians, and computer scientists are required to identify appropriate AI development models that are accurate, efficient, and applicable to clinical services.

1. Introduction

Impacted teeth are teeth that are partially or cannot fully erupt due to obstruction by bone or soft tissue or both (Alfadil and Almajed, 2020). The fact that the third molar is the last teeth to errupt often face greater challenges compared to other teeth due to limited space within the mandibular dental arch (Swift and Nelson, 2012). The global prevalence of impacted third molars (ITM) in the mandible is up to 73.5 % (Sujon et al., 2022). Impacted third molars can induce for various dental and oral diseases, such as pericoronitis, caries, periodontal pocket formation, anchorage loss, root damage to adjacent teeth, formation of odontogenic cysts and tumors, and traumatic occlusion (Borle, 2014).

Odontectomy is the most common procedure for treating impacted teeth by removes impacted teeth or the remaining roots that cannot be extracted by conventional treatments (Wayland, 2018). The categorization of ITM in the mandible holds significance in assessing the complexity of odontectomy procedures. The Winter's classifications and Pell & Gregory's are the prevailing method for classifying the ITM (Borle, 2014). Fig. 1 demonstrated Winter's classification that categorizes ITM based on angulation relative to the long axis of mandibular second molars. The classification proposed by Pell & Gregory is utilized to categorize mandibular ITM by considering the distance between the

Peer review under responsibility of King Saud University. Production and hosting by Elsevier

* Corresponding author at: Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Universitas Gadjah Mada, Denta Street, No. 1, Sleman Regency, Yogyakarta, Indonesia.

E-mail address: rinihapsara@ugm.ac.id (R. Widyaningrum).

https://doi.org/10.1016/j.sdentj.2023.11.025
Received 20 June 2023; Received in revised form 21 November 2023; Accepted 23 November 2023
Available online 25 November 2023
1013-9052/© 2023 THE AUTHORS. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

cervical line and occlusal of the second molar and the depth of the ITM, as well as the available space for the crown of the ITM in relation to the mandibular ramus (Fig. 1) (Borle, 2014).

Common practice in dentistry is to utilize panoramic radiography imaging of impacted mandibular third molars, allowing for comprehensive view of the dental arch and their surrounding structures before determining odontectomy. Using panoramic radiograph, one can identify the alignment of ITM in the mandible with nearby anatomical structures (Nagaraj et al., 2016).

Knowledge and technology of artificial intelligence (AI) has rapidly developed; particularly in, assisting in fulfilling human needs, computers and technology are utilized to imitate human intelligent action and analytical thinking (Amisha et al., 2019). In the medical field, there has been a significant advancement in the utilization of deep learning-based computer-aided diagnosis (CAD) systems, which are employed for object detection, image analysis, and data processing. These systems enable direct data analysis without the need of expert (Yasa et al., 2021). Recently, deep learning has been investigated to provide a second opinion for the diagnosis of pneumonia based on X-ray images (Mandeel et al., 2022) and improve the diagnostic performance for skin lesion classification using dermoscopy images (Pratiwi et al., 2021). Convolutional neural networks (CNNs) are subsets of deep learning which emulate human intelligence in learning and problem solving, allowing dentists and radiologists to interpret radiographs more efficiently and accurately for caries detection and classification (Vinayahalingam et al., 2021b, 2021a), perform teeth detection and numbering (Estai et al., 2022), predict third molar eruption (Vranckx et al., 2020), predict time to extraction of third molar (Kwon et al., 2022), detect apical lesion (Ekert et al., 2019), execute segmentation on panoramic radiograph for periodontitis (Widyaningrum et al., 2022) and detect the periodontal disease (Lee et al., 2018), as well as classify maxillary sinus (Murata et al., 2019). Deep learning has also studied in several studies to classify mandibular third molars and analyze the correlation between ITM and mandibular canal (MC) automatically with datasets in the form of panoramic radiograph images using several architectures (Celik, 2022; Choi et al., 2022; Fukuda et al., 2020; Sukegawa et al., 2022a; Yoo et al., 2021; Zhu et al., 2021).

The societal desire for high-quality healthcare has driven technological advancements in the field of health. This technology supports more accurate and faster diagnoses also ensures that patients can receive the best health care with minimum risk. One significant development in pursuit of this goal is the utilization of deep learning in the medical field. Deep learning technology enables the processing of complex data and deeper analysis in disease diagnosis. By utilizing deep learning algorithms, dentists and medical professionals can make more precise and quicker decisions, ultimately enhancing patient care.

In the past decade, deep learning has been investigated for improving the diagnosis of ITM in the mandible using panoramic radiographs. Despite the limitations of panoramic radiograph as a 2-dimensional radiography modality, it remains unknown the significance deep learning is in supporting the treatment of ITM diagnosed using panoramic radiographs. This review further explores the advancement of technologies using AI and provides information on the application and prospects of deep learning for the radiographic assessment for the detection, classification, and evaluation of the positions of ITM in relation to MC.

2. Materials and methods

PubMed, Science Direct, and Google Scholar databases were used to conduct literature searching in this narrative review. Literature search was performed using the keywords, “third molars,” “mandibular nerve,” “impacted tooth,” “dental radiography,” “artificial intelligence,” and “deep learning”. Manual search was also used for literature search. The inclusion criteria were literatures published in English between 2012 and 2023. The exclusion criteria included articles that were not fully accessible, excluded research methods, and contained only abstracts.

In the context of deep learning for computer vision, detection refers to the task of identifying and localizing objects of interest within images. It involves not only recognizing what objects are present in the input data but also determining their precise locations in terms of bounding boxes. In this review, deep learning detects the incidence of mandibular ITM on panoramic radiographs. In deep learning, classification is the process of categorizing input data into predefined classes. The primary goal of classification is to assign labels to input instances based on their features or characteristics. In this review, deep learning relies on the Pell & Gregory and Winter classifications as the fundamental basis for classifying ITM in the mandible. Meanwhile, evaluation refers to the task of categorizing the relationships between ITM and MC, such as whether they are in contact or not.

3. Results

This review was included 49 articles that were elected based on the predetermined exclusion and inclusion criteria. The present discussion focuses on the 12 main articles (Fig. 2). The main articles were original articles reporting studies of detection, classification, and the evaluation of the association between mandibular third molars and MC.

Several studies of deep learning for the detection, classification, and the evaluation of ITM in the MC have been conducted over the past five years. Since 2019, the number of articles has increased annually, and these publications culminated in 2022 with seven articles on the detection, classification, and evaluation of ITM to the MC by deep learning. Fig. 3 illustrates the distribution of studies on deep learning for

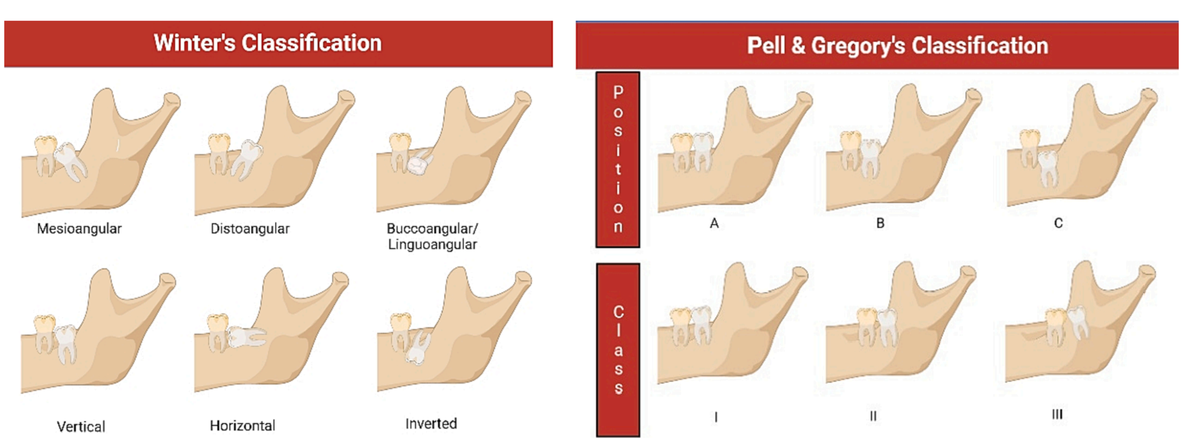

Fig. 1. Winter’s classification and Pell & Gregory’s classifications of ITM in the mandible.

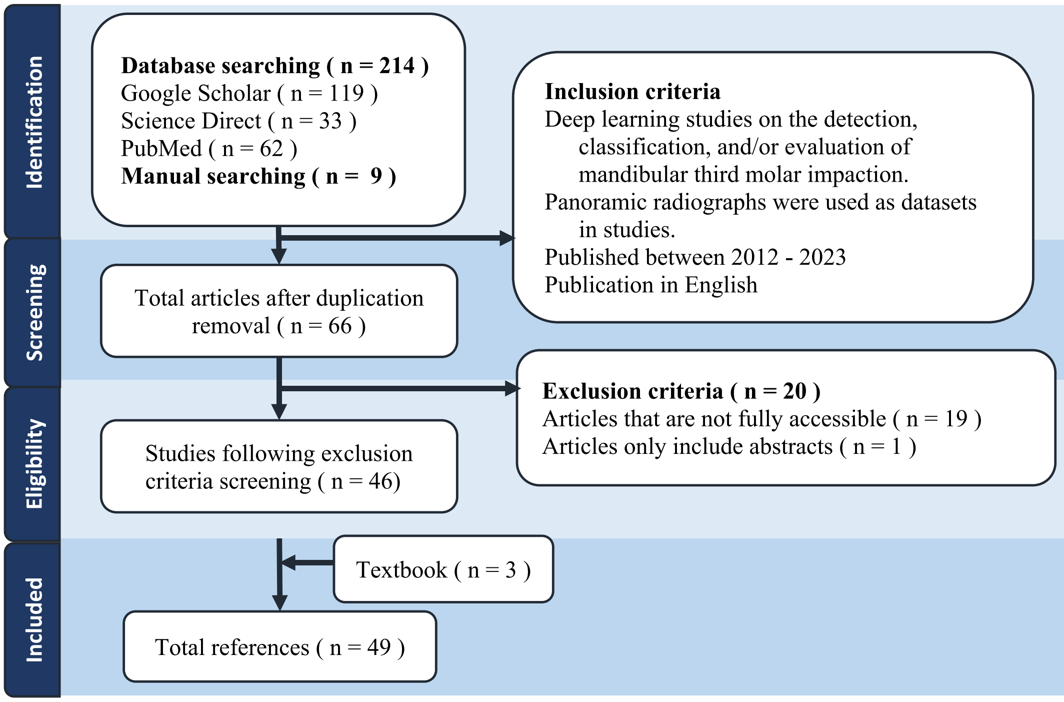

Fig. 2. Literature searching and selection.

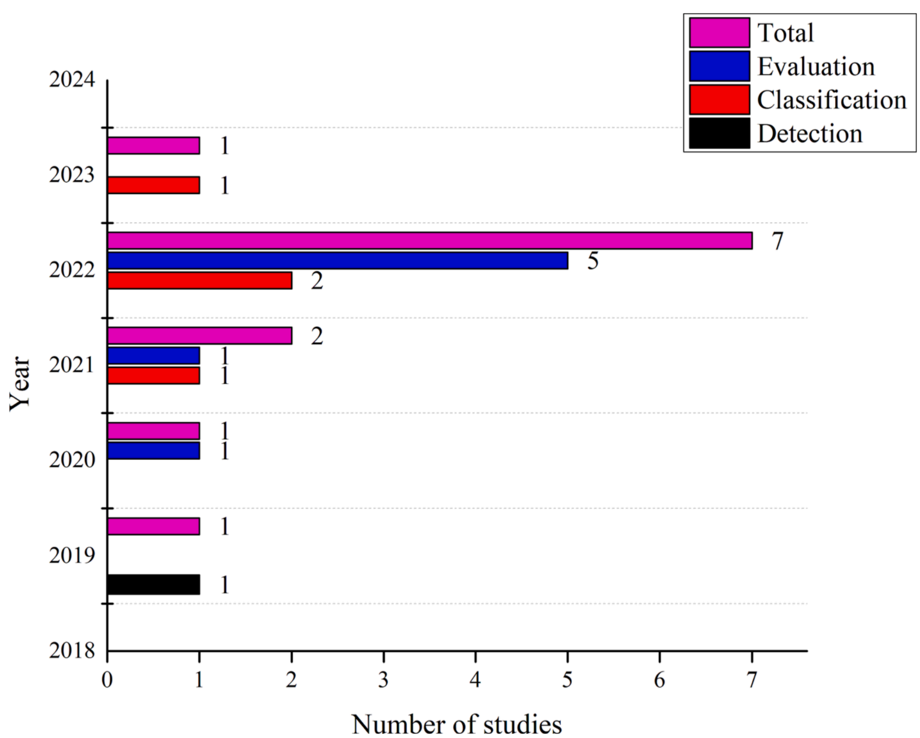

Fig. 3. Distribution of articles on deep learning applications for the detection, classification, and evaluation of ITM on panoramic radiograph during the last five years (2019 to January 2023).

panoramic radiographic evaluation of impacted mandibular third molars over the last five years (2019 to January 2023).

4. Discussion

We subdivided the articles based on the use of deep learning on panoramic radiographs in accordance with the study objective, that is, the detection, classification, and evaluation of ITM. Based on the results of a literature search, only one article discussed the detection (segmentation of panoramic radiographs) of mandibular ITM. Thus, we combined the discussion in this review regarding detection with the classification of ITM.

4.1. Deep learning performance on panoramic radiographs for detection and classification of mandibular ITM

Deep learning is widely used for medical image segmentation to assist the diagnosis of several diseases. Image segmentation is needed to detect and separate objects of interest in an image (Ronneberger et al., 2015). Deep learning can detect and classify impacted teeth, such as canine (Aljabri et al., 2022), mesiodens (Jeon et al., 2022), supernumerary in the maxilla (Kuwada et al., 2020), and dental anomalies (Okazaki et al., 2022). Based on the results of the literature search, one article discussed deep learning using the U-Net model for ITM detection in the mandible, which showed a high performance with an average dice coefficient score of 93.6 % and a Jaccard index of 88.1 % (Vinayahalingam et al., 2019). The dice coefficient was used to measure the level of similarity between the results of the detection of ITM in the mandible, which was performed manually by experts (as ground truth), and the findings of automatic segmentation from deep learning algorithms (Widyaningrum et al., 2022). The Jaccard index is also known as the intersection over union and represents the accuracy prediction of the bounding box. It is determined as the percentage of overlap between the ground truth area and the estimated bounding box area (Celik, 2022). Segmentation can be performed manually, but the automation of image segmentation can reduce the effort and time required to complete the diagnosis process of oral diseases through radiographic images.

Previous studies used deep learning to develop an automatic mandibular third molar impaction classification system with several architectures, such as ResNet-34 (Yoo et al., 2021), VGG-16 (Maruta et al., 2023; Sukegawa et al., 2022a), and YOLOv3 (Celik, 2022) (Table 1). Based on the review results in Table 1, deep learning shows a high performance in the classification of mandibular ITM (>78.91 %) (Yoo et al., 2021). The use of deep learning resulted in an accuracy of > 86 % in Winter's classification, >82.03 % in the Pell & Gregory's classification based on the remaining space, and > 78.91 % in the classification based on position of mandibular ITM (Sukegawa et al., 2022a; Yoo et al., 2021).

As shown in Table 1, a study on the mandibular ITM with Winter's

Table 1
Diagnostic Performance of Deep Learning for Detection and Classification of Mandibular ITM on Panoramic Radiograph.

Authors (Year)	Detection / Classification	Deep Learning Models	Number of datasets	Diagnostic Performance	
				Accuracy	**Others**
Vinayahalingam et al. (2019)	Detection of impacted third molar	U-Net	81		Dice coefficient: 93.6 Jaccard index: 88.1 Sensitivity: 94.7 Specificity: 99.9
Yoo et al. (2021)	Winter (Mesioangular)	ResNet-34	600	90.23	Sensitivity: 94.15 Specificity: 92.67
	(Horizontal)				Sensitivity: 89.53 Specificity: 97.65
	(Vertikal)				Sensitivity: 94.84 Specificity: 95.24
	(Distoangular)*				–
	Pell & Gregory (Class)			82.03	
	I				Sensitivity: 71.69 Specificity: 94.22
	II				Sensitivity: 90.37 Specificity: 69.52
	III				Sensitivity: 61.36 Specificity: 98.29
	Pell & Gregory (Position)			78.91	
	A				Sensitivity: 88.13 Specificity: 92.05
	B				Sensitivity: 72.77 Specificity: 84.12
	C				Sensitivity: 78.63 Specificity: 90.89
Celik, (2022)	Winter (Mesioangular)	YOLOv3	440	86	Precision: 84.9–90.8 AP: 96–98.4 Recall: 95–98.7
	(Horizontal)				Precision: 88–96 AP: 98–99.5 Recall: 83.3–100
Sukegawa et al., (2022a)	Winter	VGG-16	1330	86.63	Precision: 85.59 Recall: 80.03 F1-score: 81.38 AUC: 98.01
	Pell & Gregory (Class)			85.41	Precision: 85.88 Recall: 85.44 F1-score: 85.38 AUC: 96.38
	(Position)			88.95	Precision: 88.24 Recall: 88.77 F1-score: 88.31 AUC: 97.39
Maruta et al. (2023)	Winter	VGG-16	1180	79.59	F1-scrore: 64.23 AUC: 95.49
	Pell & Gregory (Class)			86.09	F1-scrore: 76.24 AUC: 93.34
	(Position)			84.32	F1-scrore: 81.56 AUC: 93.95

* Reveals that the existing data were insufficient in that circumstance; AP: Average precision; AUC: Area under the Receiver Operating Characteristic (ROC) curve.

and Pell & Gregory's classifications attained the highest accuracy of 90.23 % when using the ResNet-34 architecture but used fewer datasets compared with other research (Yoo et al., 2021). Previous study (Sukegawa et al., 2022a) used a single-stage technique with YOLOv3 compared with two-stage technique with Faster R-CNN and the addition of three backbone architectures, namely ResNet50, AlexNet, and VGG16 (Celik, 2022). The accuracy of the single-stage technique outperformed that of the two-stage technique by 86 %. This study demonstrated that YOLOv3 is effective in classifying ITM based on Winter's classification. The use of data augmentation is one of the ideas proposed to increase the accuracy of deep learning. Flip, rotation, and adjustment of the brightness, sharpness, contrast and additional images are performed using paint software to mimic dental fillings or caries (Maruta et al., 2023).

The performance indicator of deep learning for the classification ITM, such as sensitivity, specificity, precision, recall, and accuracy, cannot be compared with one another because they use the different classification indicators (Table 2) (Celik, 2022; Sukegawa et al., 2022a; Yoo et al., 2021). Sukegawa et al., (2022a) and Maruta et al. (2023) used the same classification indicator and deep learning architecture, thus, these studies can be compared. Using larger datasets, Sukegawa et al., (2022a) reported accuracy better than Maruta et al. (2023).

Research was continually developed to compare a single-task model that classifies impacted mandibular third molars one by one with the Multi 3Task model, which is a deep learning model that investigates multiple classifications, for simultaneous several predictions of diagnosis (Sukegawa et al., 2022a). Multi 3Task classifies ITM based on Winter's classification and Pell & Gregory's class and position simultaneously. By performing several tasks simultaneously, the multitask model can reduce computation costs; however, similar to previous research conducted by Celik (2022), the use of the multitask model did not exceed the performance of the single-task model. A significant drop in performance occurred, particularly in Multi 3Task. Winter's and Pell & Gregory's classifications have different classification characteristics. Thus, differences in the areas of interest resulted led to a decrease in performance of the multitask model in comparison to the single-task model. Deep learning was trained to determine Winter's classification, which assesses the inclination and angulation of the ITM with feature extraction on all mandibular third molars; Winter's classification also has characteristics different from those of Pell & Gregory's classification and uses the area under the mandibular third molars for classification (Sukegawa et al., 2022a).

Nine of the twelve main articles reviewed in this study (Table 2) mentioned the formulas utilized to determine metrics. Some of them used formulas for estimating sensitivity, specificity, NPV, and PPV are different from those prescribed by medical diagnostic test calculation guidelines (Table 2). Yoo et al. (2021) determined the sensitivity and specificity using positive predictive value (PPV) calculations. Furthermore, the studies reviewed in this work did not use a consistent mathematical formula (Ariji et al., 2022; Buyuk et al., 2022; Celik, 2022; Choi et al., 2022; Maruta et al., 2023; Takebe et al., 2022; Vinayahalingam et al., 2019; Yoo et al., 2021; Zhu et al., 2021). Celik (2022), Choi et al. (2022), Takebe et al., 2022, and Zhu et al. (2021), used PPV calculations to determine precision in their research. A sensitivity calculation is used by Zhu et al. (2021), Takebe et al. (2022) and Choi et al. (2022). Meanwhile, Celik (2022) used the true positive formula divided by all the data as a recall. Buyuk et al. (2022), Vinayahalingam et al. (2021a) provided calculations in accordance with the guidelines for medical diagnostic tests in equations (1), (2), (3), (4), and (5) in Table 2 (Leeflang, 2014).

In the discipline of informatic and computer science, the confusion matrix is a common standard used to assess the deep learning performance. This method is similar to the diagnostic test methods used in medical and health studies. Both confusion matrix and diagnostic test methods are concerned with specificity, sensitivity, negative predictive value (NPV), positive predictive value (PPV), and accuracy. The confusion matrix includes other terms, such as recall/true positive rate (which is the same as sensitivity), true negative rate (which is the same as specificity), and precision (which is the same as PPV) (Chicco et al., 2021). In the field of computer science, sensitivity (true positive rate) means how well the deep learning algorithm can identify images in their category. Specificity is the capability of the deep learning algorithm to accurately detect images outside of their class. The PPV parameter indicates the percentage of actually positive data among data that the algorithm deep learning considers to be positive. The NPV denotes the ratio of actual negative data to the data that deep learning algorithms interpret as negative. Accuracy is used to calculate the overall deep learning performance (Buyuk et al., 2022; Vinayahalingam et al., 2021a). Referring to Table 2, the calculation method is the same from a computational and medical perspective but only for the calculation of accuracy.

4.2. Diagnostic performance of deep learning for evaluating the relationship between ITM and the MC on panoramic radiographs

Panoramic radiographs are typically used based on the treatment plan for odontectomy to lessen the chance of alveolar nerve impairment by assessing the association between ITM and MC (Liu et al., 2015). Technology advancements have enabled the integration of AI into CAD systems, which can be used to assess the relationship between the third molars on either side (bucco-lingual) of the mandible and the mandibular canal. GoogLeNet, VGG-16, AlexNet (Buyuk et al., 2022; Fukuda et al., 2020), YOLOv3 (Takebe et al., 2022), YOLOv4 (Zhu et al., 2021),

Table 2
Comparison of the medical diagnostic test and equations for evaluating deep learning performance.

Authors (Year)	Sensitivity	Specificity	PPV	NPV	Accuracy	Other metrics
Formula (Leeflang, 2014)	$\frac{TP}{TP+FN}$	$\frac{TN}{TN+FP}$	$\frac{TP}{TP+FP}$	$\frac{TN}{FN+TN}$	$\frac{TP+TN}{TP+FP+TN+FN}$	–
Confusion matrix (Chicco et al., 2021)	Sensitivity = Recall/ True Positive Rate	Specificity = True Negative Rate	PPV = Precision	NPV	Accuracy	–
Yoo et al. (2021)	–	–	Sensitivity	Specificity	Accuracy	–
Zhu et al. (2021)	Recall	–	Precision	–	–	F1-score = 2 × precision × recall / (precision + recall)
Celik (2022)	Recall = TP / All data	–	Precision	–	Accuracy	–
Choi et al. (2022)	Recall	–	Precision	–	Accuracy	F1-score = 2 × precision × recall / (precision + recall)
Buyuk et al. (2022), Vinayahalingam et al. (2021a)	Sensitivity	Specificity	PPV	NPV	Accuracy	–
Ariji et al. (2022)	Sensitivity	–	–	–	–	–
Takebe et al. (2022)	Recall (Sensitivity)	–	Precision	–	Accuracy	F1-score = 2 × precision × recall / (precision + recall)
Maruta et al. (2023)	–	–	–	–	–	F1-score = 2 × precision × recall / (precision + recall)

NPV: Negative predictive value; PPV: Positive predictive value; FN: False negative; FP: False positive; TN: True negative; TP: True Positive.

ResNet-50 (Choi et al., 2022; Sukegawa et al., 2022b) are some deep learning architectures that have been investigated for radiographic evaluation of impacted third molar related to the mandibular canal. In addition, U-Net is also used for image segmentation of the mandibular canal (Ariji et al., 2022; Buyuk et al., 2022; Vinayahalingam et al., 2019).

Fukuda et al. (2020), Zhu et al. (2021), Choi et al. (2022), and Sukegawa et al., (2022b) evaluated the association between ITM and MC, which was categorized into two distinct types: contact and non-contact. In the study by Fukuda et al. (2020), 600 radiographs with the GoogLeNet architecture showed a better performance on all parameters, such as sensitivity (88 ± 6 %), specificity (96 ± 44 %), and accuracy (92 ± 5 %), compared with those in the studies of Zhu et al. (2021) and Choi et al. (2022). High sensitivity means that deep learning can correctly predict positive results if the radiograph shows contact between ITM and MC. High specificity indicates the capability of deep learning to correctly predict negative results if the impacted the ITM and MC is non-contact. The quantity of radiographs as dataset supports the outcomes of this high performance of deep learning because it will enhance the capacity to learn.

Additionally, Fukuda et al. (2020) used a data augmentation process as part of the process for creating the deep learning algorithms, particularly in the case of a restricted quantity of datasets. The use of data augmentation, which is commonly used to improve performance, had an impact on accuracy (Maruta et al., 2023). Fukuda et al. (2020) compared three deep learning architectures—GooLeNet, AlexNet, and VGG-16—to evaluate the ITM in relation to the MC. Each architecture was designed using distinct image from cropped panoramic radiograph pictures (Fukuda et al., 2020). Deep learning performs more accurately in terms of prediction the small-sized image (England and Cheng, 2019). Datasets with larged-size images typically contain more complicated information, in this case, additional anatomical images from panoramic radiographs that are not necessary for the learning process (Fukuda et al., 2020).

Zhu et al. (2021) compared the accuracy of deep learning with dentists (Zhu et al., 2021) In addition, deep learning has been used by dentists to evaluate the relationship between the ITM and the MC. With a precision of 93.88 % and a recall of 92 %, dentists applied deep learning to produce the highest performance results. This finding indicates that dentists can use deep learning to assess how mandibular third molars and MC interact to provide more accurate findings; however, further research is required because of the limited use of radiographs and dentists (Zhu et al., 2021).

Sukegawa et al., (2022b) analyzed the effect of optimizer algorithms with gradient methods, such as sharpness aware minimization (SAM) and stochastic gradient descent (SGD). ResNet 50v2 with an SAM optimizer showed the best performance in this research. Overlapping and poor generalization performance results were obtained with the use of SGD; meanwhile, SAM enhanced robustness against noise and generalization performance (Sukegawa et al., 2022b) and Choi et al. (2022) also performed similar studies, but with six and five oral surgeons, respectively. Compared with oral surgeons, ResNet-50 and YOLOv3 performed better.

The low diagnostic performance of oral surgeons is affected by the difficulty of using panoramic radiographs to determine the bucco/lingual position and vertical relationship of the mandibular third molars to the mandibular canal (Choi et al., 2022). In the evaluation of the relationship between the ITM, radiographic features, such as canal narrowing, dark root apexes, bifid root apexes, root narrowing, interruption of the mandibular canal wall, canal deflection, and root deflection, can be observed (Nasser et al., 2018; Whaites and Drage, 2014). The position of impacted mandibular third molars that are buccal or lingual to the mandibular canal but not in contact with it can be revealed by the superimposed radiographic appearance (Choi et al., 2022). Buyuk et al. (2022) addressed the needs of dentists because deep learning can assess the association between ITM and MC using standards that are nearly identical to those used in dentistry by classifying panoramic radiographs of ITM into four categories (Table 3). The fused deep learning method was utilized, segmentation was performed using the U-Net architecture, and the AlexNet architecture was used to assess the relationship between the ITM and MC (Buyuk et al., 2022). This study also compared the performance of fused deep learning with dentists, revealing nearly the same performance, and showed that fused deep learning can be used to help in diagnosis.

Ariji et al. (2022) use the U-Net and transfer learning technique. Transfer learning techniques are performed using the learning processes in deep learning models on large datasets (source models) and then transferring knowledge to conduct learning on other deep learning models with smaller datasets (target models) for the same task (Alzubaidi et al., 2021). This study conducted by Ariji et al. (2022) shows the target model's deep learning performance is on par with the source model, especially in 200 datasets indicate that the transfer learning technique can be a method for transferring knowledge in a relatively small dataset. Transfer learning techniques can be a solution for developing deep learning using large datasets obtained from several institutions while still protecting the security of patient medical record data (Ariji et al., 2022). The study's findings demonstrate that the target model's deep learning performance is comparable to that of the source model, especially in small datasets. While maintaining the integrity of patient medical record data, transfer learning techniques can be used to construct deep learning using massive datasets collected from many institutions (Ariji et al., 2022).

4.3. *Prospects for deep learning applications in the Detection, Classification, and evaluation of mandibular ITM on panoramic radiographs*

Deep learning has developed because of today's increasingly powerful hardware and software improvements, and it has been used in many fields of studies, including oral radiology (Corbella et al., 2021). CBCT is widely regarded as the preferred method as well as the gold standard for supporting examinations of ITM. CBCT involves maxillofacial imaging that uses the most advanced technology; it can display three-dimensional images from a cross-sectional view to evaluate the incidence or nonexistence of corticalization between the ITM and MC (Nasser et al., 2018). Deep learning can also identify and categorize ITM with the use of CBCT (Borgonovo et al., 2017; Liu et al., 2022; Orhan et al., 2021). On the other hand, CBCT is expensive and requires larger doses of radiation, thus, panoramic radiography remains a widely used an alternative in the examination of impacted teeth (Whaites and Drage, 2014), in addition as a preliminary step to determining whether CBCT imaging is necessary. In diagnosis, deep learning may be used in conjunction with CBCT (Takebe et al., 2022). However, deep learning improves the visualization of the pseudo contact of ITM with the MC on panoramic radiographs. Therefore, the utilization of CBCT can be diminished. (Zhu et al., 2021).

AI based on deep learning can be applied in clinical practice to assist dentists in making a diagnosis quickly and accurately, minimize misinterpretations of diagnoses due to high workloads, and speed up the interpretation process (Zhu et al., 2021). AI will greatly facilitate the screening process in regions with a shortage of radiologists (Celik, 2022). In the future, the detection, classification, and evaluation of ITM in the mandible using panoramic radiographs can be integrated in recent applications to support the diagnosis in dental practice, as well as to increase the diagnostic accuracy of radiographic examinations. As the process of making a diagnosis is still handled by the dentist in charge of the patient, dentists and oral radiologists are still in charge of patient care. Deep learning or any other AI cannot fully replace dentists or radiologists because of the probability of misdiagnosis, that is, either underdiagnosis or overdiagnosis (Roosanty et al., 2022).

The findings presented in Table 2 demonstrated the inconsistent performance of deep learning in facilitating the diagnosis of impacted

Table 3
Diagnostic Performance of Deep Learning in Evaluating the Relationship Between ITM and MC.

Authors (Year)	Evaluation Relationship with Mandibular Canal	Deep Learning Models	Numbers of Dataset	Diagnostic Performance	
				Accuracy	**Others**
Vinayahalingam et al., (2019)	Segmentation of mandibular canals	U-Net	81	–	Dice coefficient: 80.5 Jaccard index: 68.7 Sensitivity: 84.7 Specificity: 96.7
Fukuda et al., (2020)	Contact/ Superimposed and Non-contact/ Non-superimposed	GoogLe Net	600	92 ± 5	Sensitivity: 88 ± 6 Specificity: 96 ± 4 AUC: 93
Zhu et al. (2021)	Contact and Non-contact	YOLOv4	503	–	Precision: 88.71 Recall: 91.67 AP: 83.02 F1-score: 90.16
		YOLOv4 and Dentist			Precision: 93.88 Recall: 92.00 AP: 88.06 F1-score: 92.93
Choi et al. (2022)	Contact and Non-contact	ResNet-50	571	72.32	Sensitivity: 84.62 Specificity: 55.32
	Buccolingual position			80.65	Sensitivity: 86.67 Specificity: 75
Buyuk et al. (2022)	Segmentation mandibular canals	U-Net	1880	99	Dice coefficient: 91 Jaccard index: 98
	Non-contact	AlexNet		80	Sensitivity: 74 Specificity: 92 PPV: 79 NPV: 90
	Contact				Sensitivity: 83 Specificity: 95 PPV: 88 NPV: 94
	Superimposed				Sensitivity: 86 Specifisity: 88 PPV: 80 NPV: 93
	Mandibular third molar roots that crossed the mandibular canal.				Sensitivity: 67 Specifisity: 96 PPV: 68 NPV: 96
Takebe et al. (2022)	Contact	YOLOv3	579	89.4	Precision: 89.1 Recall: 92.5 F1-score: 90.8
Sukegawa et al., (2022b)	Contact and Non-contact	ResNet50 with SAM Optimizer	1279	85.5	Precision: 81 Recall: 78.5 F1-score: 79.4 AUC: 88.3
		ResNet50 with SGD Optimizer		85	Precision: 80.4 Recall: 78.1 F1-score: 78.9 AUC:87.5
		ResNet50v2 with SAM Optimizer		86	Precision: 81.6 Recall: 79.1 F1-score: 80 AUC: 89
		ResNet50v2 with SGD Optimizer		85.3	Precision: 80.9 Recall: 78.2 F1-score: 79.2 AUC: 88.4
Ariji et al. (2022)	Segmentation mandibular canal	U-Net with Transfer Learning	881	–	Dice coefficient: 74.9–85.7 Jaccard index: 60.7–79.1 Sensitivity: 71.4–87.1

AP: Average precision; AUC: Area under the ROC curve; NPV: Negative predictive value; PPV: Positive predictive value.

teeth across different studies. Regarding the results presented in Table 2, further studies on deep learning for medical and dental applications should consider using diagnostic performance calculations that are more consistent with the diagnostic test commonly employed in medical and health studies.

Bringing together the fields of computer science, radiology, oral surgery, and statistical science through collaborative efforts will facilitate the development of AI studies that are tailored to satisfy the demands of medical and dental practices. However, dentists and dental specialists participating in deep learning research are still very few (Choi et al., 2022; Zhu et al., 2021). To produce high performance deep learning-based CAD that can be applied clinically in dental practice, computing experts and dentists should collaborate more in the future to develop deep learning. In further research, deep learning can be used to estimate the difficulty level of ITM extraction and radiographs can be used to evaluate the depth of impaction of mandibular ITM. The depth of mandibular third molars is assessed using two methods: Winters line and second molar guides (Whaites and Drage, 2014). Moreover, third molars, which are still partly erupted and far from the mandibular canal, minimize the risk of complications in odontectomy. Further research on deep learning to predict the complete eruption position of mandibular ITM that are still in dentition will be beneficial in clinical practice (Zhu et al., 2021).

Deep learning can enhance assistance for clinicians by incorporating predictive factors such as the complexity of wisdom tooth extraction (including the potential need for root division), estimating extraction time, assessing the likelihood of root adhesion, and considering their relationship with the mandibular canal. Nevertheless, acquiring this data presents challenges as it relies on clinical information. Considering this issue, the development of AI for CAD necessitates multidisciplinary collaboration among clinicians, radiologists, and computer scientists (Montagnon et al., 2020). From a clinical standpoint, radiologists should collaborate with oral surgeons to determine the problem formulation, which demands precise diagnosis and treatment planning for patients. Meanwhile, oral radiologists and computer scientists must collaborate for the development of imaging examinations with the best visualization, beginning with annotation, segmentation, training, testing, and validation, ensuring that deep learning works precisely based on the needs of clinicians in dental practice.

5. Conclusion

Deep learning-based AI applications show high performance in the detection, classification, and evaluation of mandibular ITM to the MC on panoramic radiographs. With or without the combined usage of other techniques, deep learning architectures used are continually being developed. Deep learning-based AI can be improved and used in clinical practice to aid dentists in diagnosing patients more accurately. A multidisciplinary approach involving clinicians, radiologists, and computer scientists must be developed to accomplish this goal.

Ethical statement

No ethical issues were raised during the study presentation.

CRediT authorship contribution statement

Amalia Nur Faadiya: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Rini Widyaningrum:** Methodology, Formal analysis, Investigation, Data curation, Writing – original draft. **Pingky Krisna Arindra:** Writing – review & editing, Supervision. **Silviana Farrah Diba:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This review is part of the first author’s thesis for undergraduate final project, under supervision of other authors. The authors acknowledge Universitas Gadjah Mada, Indonesia, for supporting this publication under grant of Final Project Recognition Grant Number 5075/UN1.P.II/Dit-Lit/PT.01.01/2023.

References

Alfadil, L., Almajed, E., 2020. Prevalence of impacted third molars and the reason for extraction in saudi arabia. Saudi Dent. J. 32, 262–268. https://doi.org/10.1016/j.sdentj.2020.01.002.

Aljabri, M., Aljameel, S.S., Min-Allah, N., Alhuthayfi, J., Alghamdi, L., Alduhailan, N., Alfehaid, R., Alqarawi, R., Alhareky, M., Shahin, S.Y., Al Turki, W., 2022. Canine impaction classification from panoramic dental radiographic images using deep learning models. Inform. Med. Unlocked 30, 100918. https://doi.org/10.1016/j.imu.2022.100918.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: concepts, cnn architectures, challenges, applications, Future Directions. J. Big Data 8, 1–74. https://doi.org/10.1186/s40537-021-00444-8.

Amisha, M.P., Pathania, M., Rathaur, V., 2019. Overview of artificial intelligence in medicine. J. Family Med. Prim. Care 8, 2328–2331. https://doi.org/10.4103/jfmpc.jfmpc_440_19.

Ariji, Y., Mori, M., Fukuda, M., Katsumata, A., Ariji, E., 2022. automatic visualization of the mandibular canal in relation to an impacted mandibular third molar on panoramic radiographs using deep learning segmentation and transfer learning techniques. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. 134, 749–757. https://doi.org/10.1016/j.oooo.2022.05.014.

Borgonovo, A.E., Rigaldo, F., Maiorana, C., Grossi, G.B., Augusti, D., Re, D., 2017. CBCT evaluation of the tridimensional relationship between impacted lower third molar and the inferior alveolar nerve position. Minerva Stomatol. 66, 9–19. https://doi.org/10.23736/S0926-4970.16.03976-X.

Borle, R.M., 2014. Textbook of Oral and Maxillofacial Surgery. Jaypee Brothers Medical Publisher (P), New Delhi.

Buyuk, C., Akkaya, N., Arsan, B., Unsal, G., Aksoy, S., Orhan, K., 2022. A fused deep learning architecture for the detection of the relationship between the mandibular third molar and the mandibular canal. Diagnostics 12, 2018. https://doi.org/10.3390/diagnostics12082018.

Celik, M.E., 2022. Deep learning based detection tool for impacted mandibular third molar teeth. Diagnostics 12, 942. https://doi.org/10.3390/diagnostics12040942.

Chicco, D., Tötsch, N., Jurman, G., 2021. The matthews correlation coefficient (mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min. 14, 1–22. https://doi.org/10.1186/s13040-021-00244-z.

Choi, E., Lee, S., Jeong, E., Shin, S., Park, H., Youm, S., Son, Y., Pang, K.M., 2022. Artificial intelligence in positioning between mandibular third molar and inferior alveolar nerve on panoramic radiography. Sci. Rep. 12, 2456. https://doi.org/10.1038/s41598-022-06483-2.

Corbella, S., Srinivas, S., Cabitza, F., 2021. Applications of deep learning in dentistry. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. 132, 225–238. https://doi.org/10.1016/j.oooo.2020.11.003.

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.

England, J.R., Cheng, P.M., 2019. Artificial intelligence for medical image analysis: a guide for authors and reviewers. Am. J. Roentgenol. https://doi.org/10.2214/AJR.18.20490.

Estai, M., Tennant, M., Gebauer, D., Brostek, A., Vignarajan, J., Mehdizadeh, M., Saha, S., 2022. Deep learning for automated detection and numbering of permanent teeth on panoramic images. Dentomaxillofacial Radiol. 51, 20210296. https://doi.org/10.1259/dmfr.20210296.

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.

Jeon, K.J., Ha, E.G., Choi, H., Lee, C., Han, S.S., 2022. Performance comparison of three deep learning models for impacted mesiodens detection on periapical radiographs. Sci. Rep. 12, 1–8. https://doi.org/10.1038/s41598-022-19753-w.

Kuwada, C., Ariji, Y., Fukuda, M., Kise, Y., Fujita, H., Katsumata, A., Ariji, E., 2020. Deep learning systems for detecting and classifying the presence of impacted supernumerary teeth in the maxillary incisor region on panoramic radiographs. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. 130, 464–469. https://doi.org/10.1016/j.oooo.2020.04.813.

Kwon, D., Ahn, J., Kim, C.S., Kang, D.O., Paeng, J.Y., 2022. A deep learning model based on concatenation approach to predict the time to extract a mandibular third molar tooth. BMC Oral Health 22, 571. https://doi.org/10.1186/s12903-022-02614-3.

Lee, J.H., Kim, D.H., Jeong, S.N., Choi, S.H., 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.

Leeflang, M.M.G., 2014. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin. Microbiol. Infect. 20, 105–113. https://doi.org/10.1111/1469-0691.12474.

Liu, M.-Q., Xu, Z.-N., Mao, W.-Y., Li, Y., Zhang, X.-H., Bai, H.-L., Ding, P., Fu, K.-Y., 2022. Deep learning-based evaluation of the relationship between mandibular third molar and mandibular canal on CBCT. Clin. Oral Invest. 26, 981–991. https://doi.org/10.1007/s00784-021-04082-5.

Liu, W., Yin, W., Zhang, R., Li, J., Zheng, Y., 2015. Diagnostic value of panoramic radiography in predicting inferior alveolar nerve injury after mandibular third molar extraction: a meta-analysis. Aust. Dent. J. 60, 233–239. https://doi.org/10.1111/adj.12326.

Mandeel, T.H., Awad, S.M., Naji, S., 2022. Pneumonia binary classification using multi-scale feature classification network on chest X-ray images. IAES Int. J. Artif. Intell. (IJ-AI) 11, 1469–1477. https://doi.org/10.11591/ijai.v.

Maruta, N., Morita, K.I., Harazono, Y., Anzai, E., Akaike, Y., Yamazaki, K., Tonouchi, E., Yoda, T., 2023. Automatic machine learning-based classification of mandibular third molar impaction status. J. Oral Maxillofac. Surg. Med. Pathol. 35, 327–334. https://doi.org/10.1016/j.ajoms.2022.12.010.

Montagnon, E., Cerny, M., Cadrin-Chênevert, A., Hamilton, V., Derennes, T., Ilinca, A., Vandenbroucke-Menu, F., Turcotte, S., Kadoury, S., Tang, A., 2020. deep learning workflow in radiology: a primer. Insights Imaging. https://doi.org/10.1186/s13244-019-0832-5.

Murata, M., Ariji, Y., Ohashi, Y., Kawai, T., Fukuda, M., Funakoshi, T., Kise, 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.

Nagaraj, T., Keerthi, I., James, L., Shruthi, R., Balraj, L., Bhavana, T.V., 2016. Visibility of mandibular anatomical landmarks in panoramic radiography: a retrospective study. J. Med. Radiol. Pathol. Surg. 2, 14–17. https://doi.org/10.15713/ins.jmrps.57.

Nasser, A., Alomar, A., AlOtaibi, N., Altamimi, A., 2018. Correlation of panoramic radiograph and CBCT findings in assessment of relationship between impacted mandibular third molars and mandibular canal in Saudi population. Dent. Oral Craniofac. Res. 4, 1–5. https://doi.org/10.15761/docr.1000256.

Okazaki, S., Mine, Y., Iwamoto, Y., Urabe, S., Mitsuhata, C., Nomura, R., Kakimoto, N., Murayama, T., 2022. Analysis of the feasibility of using deep learning for multiclass classification of dental anomalies on panoramic radiographs. Dent. Mater. J. 41, 889–895. https://doi.org/10.4012/dmj.2022-098.

Orhan, K., Bilgir, E., Bayrakdar, I.S., Ezhov, M., Gusarev, M., Shumilov, E., 2021. Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans. J. Stomatol. Oral Maxillofac. Surg. 122, 333–337. https://doi.org/10.1016/j.jormas.2020.12.006.

Pratiwi, R.A., Nurmaini, S., Rini, D.P., Rachmatullah, M.N., Darmawahyuni, A., 2021. Deep ensemble learning for skin lesions classification with Convolutional Neural Network. IAES International. J. Artif. Intell. 10, 563–570. https://doi.org/10.11591/ijai.v10.i3.pp563-570.

Ronneberger, O., Fischer, P., Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015. Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

Roosanty, A., Widyaningrum, R., Diba, S.F., 2022. Artificial intelligence based on convolutional neural network for detecting dental caries on bitewing and periapical radiographs. J. Radiol. Dentomaksilof. Indonesia (JRDI) 6, 89–94. https://doi.org/10.32793/jrdi.v6i2.867.

Sujon, M.K., Alam, M.K., Rahman, S.A., Noor, S.N.F.M., 2022. Third molar impactions prevalence and pattern among adults using 5923 digital orthopantomogram. Bangladesh J. Med. Sci. 21, 717–729.

Sukegawa, S., Matsuyama, T., Tanaka, F., Hara, T., Yoshii, K., Yamashita, K., Nakano, K., Takabatake, K., Kawai, H., Nagatsuka, H., Furuki, Y., 2022a. Evaluation of multi-task learning in deep learning-based positioning classification of mandibular third molars. Sci. Rep. 12, 1–10. https://doi.org/10.1038/s41598-021-04603-y.

Sukegawa, S., Tanaka, F., Hara, T., Yoshii, K., Yamashita, K., Nakano, K., Takabatake, K., Kawai, H., Nagatsuka, H., Furuki, Y., 2022b. Deep learning model for analyzing the relationship between mandibular third molar and inferior alveolar nerve in panoramic radiography. Sci. Rep. 12, 16925. https://doi.org/10.1038/s41598-022-21408-9.

Swift, J.Q., Nelson, W.J., 2012. The nature of third molars: are third molars different than other teeth? Atlas Oral Maxillof. Surg. Clin. 20, 159–162. https://doi.org/10.1016/j.cxom.2012.07.003.

Takebe, K., Imai, T., Kubota, S., Nishimoto, A., Amekawa, S., Uzawa, N., 2022. Deep learning model for the automated evaluation of contact between the lower third molar and inferior alveolar nerve on panoramic radiography. J. Dent. Sci. https://doi.org/10.1016/j.jds.2022.12.008.

Vinayahalingam, S., Xi, T., Bergé, 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.

Vinayahalingam, S., Kempers, S., Limon, L., Deibel, D., Maal, T., Hanisch, M., Bergé, S., Xi, T., 2021a. Classification of caries in third molars on panoramic radiographs using deep learning. Sci. Rep. 11 https://doi.org/10.1038/s41598-021-92121-2.

Vinayahalingam, S., Kempers, S., Limon, L., Maal, T., Bergé, S., Xi, T., Hanisch, M., 2021b. The automatic detection of caries in third molars on panoramic radiographs using deep learning: a pilot study. Res Sq. https://doi.org/10.21203/rs.3.rs-379636/v1.

Vranckx, M., Van Gerven, A., Willems, H., Vandemeulebroucke, A., Leite, A.F., 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.

Wayland, J., 2018. Impacted Third Molars. John Wiley & Sons, Pondicherry.

Whaites, E., Drage, N., 2014. Essentials of Dental Radiography and Radiology, fifth ed. Elsevier, China.

Widyaningrum, R., Candradewi, I., Aji, N.R.A.S., Aulianisa, R., 2022. Comparison of Multi-Label U-Net and Mask R-CNN for panoramic radiograph segmentation to detect periodontitis. Imaging Sci. Dent. 52, 383–391. https://doi.org/10.5624/isd.20220105.

Yasa, Y., Çelik, Ö., Bayrakdar, I.S., Pekince, A., Orhan, K., Akarsu, S., Atasoy, S., Bilgir, E., Odabaş, A., Aslan, A.F., 2021. An artificial intelligence proposal to automatic teeth detection and numbering in dental bite-wing radiographs. Acta Odontol. Scand. 79, 275–281. https://doi.org/10.1080/00016357.2020.1840624.

Yoo, J.H., Yeom, H.G., Shin, W.S., Yun, J.P., Lee, J.H., Jeong, S.H., Lim, H.J., Lee, J., Kim, B.C., 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.

Zhu, T., Chen, D., Wu, F., Zhu, F., Zhu, H., 2021. Artificial intelligence model to detect real contact relationship between mandibular third molars and inferior alveolar nerve based on panoramic radiographs. Diagnostics 11, 1664. https://doi.org/10.3390/DIAGNOSTICS11091664.