

Original paper

Assessing the accuracy of artificial intelligence in mandibular canal segmentation compared to semi-automatic segmentation on cone-beam computed tomography images

Julien Issa^{1,2A,B,C,D,E,F}, Marta Dyszkiewicz Konwinska^{1,A,B,C,D,E,F}, Natalia Kazimierczak^{3,E,F}, Raphael Olszewski^{4,5,6,7,A,D,E}

¹Department of Diagnostics, Chair of Practical Clinical Dentistry, Poznan University of Medical Sciences, Poznan, Poland

²Doctoral School, Poznan University of Medical Sciences, Poznan, Poland

³Kazimierczak Private Medical Practice, Bydgoszcz, Poland

⁴Neuro Musculo Skeletal Lab (NMSK), Institut de Recherche Expérimentale et Clinique (IREC), Université Catholique de Louvain (UCLouvain), 1200 Brussels, Belgium

⁵Oral and Maxillofacial Surgery Lab (OMFS Lab), NMSK, IREC, Université Catholique de Louvain (UCLouvain), 1200 Brussels, Belgium

⁶Department of Oral and Maxillofacial Surgery, Cliniques Universitaires Saint-Luc, 1200 Brussels, Belgium

⁷Department of Perioperative Dentistry, L. Rydygiera Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland

Abstract

Purpose: This study aims to assess the accuracy of artificial intelligence (AI) in mandibular canal (MC) segmentation on cone-beam computed tomography (CBCT) compared to semi-automatic segmentation. The impact of third molar status (absent, erupted, impacted) on AI performance was also evaluated.

Material and methods: A total of 150 CBCT scans (300 MCs) were retrospectively analysed. Semi-automatic MC segmentation was performed by experts using Romexis software, serving as the reference standard. AI-based segmentation was conducted using Diagnocat, an AI-driven cloud-based platform. Three-dimensional segmentation accuracy was assessed by comparing AI and semi-automatic segmentations through surface-to-surface distance metrics in Cloud Compare software. Statistical analyses included the intraclass correlation coefficient (ICC) for inter- and intra-rater reliability, Kruskal-Wallis tests for group comparisons, and Mann-Whitney *U* tests for post-hoc analyses.

Results: The median deviation between AI and semi-automatic MC segmentation was 0.29 mm (SD: 0.25-0.37 mm), with 88% of cases within the clinically acceptable limit (≤ 0.50 mm). Inter-rater reliability for semi-automatic segmentation was 84.5%, while intra-rater reliability reached 95.5%. AI segmentation demonstrated the highest accuracy in scans without third molars (median deviation: 0.27 mm), followed by erupted third molars (0.28 mm) and impacted third molars (0.32 mm).

Conclusions: AI demonstrated high accuracy in MC segmentation, closely matching expert-guided semi-automatic segmentation. However, segmentation errors were more frequent in cases with impacted third molars, probably due to anatomical complexity. Further optimisation of AI models using diverse training datasets and multi-centre validation is recommended to enhance reliability in complex cases.

Key words: diagnosis, dentistry, segmentation, inferior alveolar canal, artificial intelligence.

Correspondence address:

Julien Issa, Department of Diagnostics, Chair of Practical Clinical Dentistry, Poznan University of Medical Sciences, 70 Bukowska St., 60-812 Poznan, Poland,
e-mail: 86658@student.ump.edu.pl

Authors' contribution:

A Study design · B Data collection · C Statistical analysis · D Data interpretation · E Manuscript preparation · F Literature search · G Funds collection

Introduction

Technological advancements have transformed dental and maxillofacial workflows, catalysing a paradigm shift toward digitised diagnostic, treatment planning, and postoperative follow-up. Central to this evolution is cone-beam computed tomography (CBCT), a cornerstone imaging modality that overcomes the inherent limitations of conventional two-dimensional radiography by delivering high-resolution, three-dimensional anatomical reconstructions [1,2]. Its capacity to delineate complex craniofacial structures with precision has rendered CBCT indispensable for preoperative planning, particularly in identifying critical regions such as the mandibular canal (MC) [3,4].

Located within the MC, the inferior alveolar nerve (IAN), a component of the mandibular neurovascular bundle, provides sensory innervation to the mandibular dentition, lower lip, chin, and associated soft tissues while also facilitating motor function in masticatory muscles [5]. Precise localisation of the IAN and its spatial orientation relative to adjacent anatomical landmarks is imperative for mitigating iatrogenic injury during procedures such as dental implant placement and third molar extraction [6,7]. Despite advancements, IAN injury persists as a clinically significant complication, with reported incidence rates ranging from 0.4% to 13.4% [8,9]. Such injuries may manifest as transient hypoesthesia or irreversible paraesthesia, profoundly impairing patients' functional and psychosocial well-being [10]. Consequently, meticulous preoperative mapping of the MC, inclusive of anatomical variations, is critical to minimising procedural risk. Nevertheless, achieving consistent and accurate segmentation of the MC, a surrogate for IAN localisation, remains a persistent challenge in clinical practice.

Current segmentation methodologies in medical imaging span manual, semi-automated, and automated approaches. Manual segmentation, traditionally regarded as the gold standard, necessitates expert-driven, slice-by-slice annotation of the MC on CBCT scan [11]. While this technique achieves high anatomical fidelity, it is labour intensive, time prohibitive, and susceptible to inter- and intra-operator variability [12]. Semi-automated algorithms integrated into commercial CBCT software partially alleviate these constraints by allowing users to demarcate discrete points along the MC trajectory, which are interpolated to generate a fixed-diameter cylindrical model [13]. Although this approach reduces segmentation time, it does not eliminate potential operator bias introduced during manual input.

The advent of artificial intelligence (AI), particularly convolutional neural networks (CNNs) and deep learning (DL), has significantly advanced medical image analysis [14,15]. By emulating human cognitive processes through pattern recognition in annotated datasets, AI-driven systems enable rapid, reproducible segmentation of intricate anatomical structures.

Emerging studies highlight the efficacy of AI in segmenting various oral and maxillofacial structures on CBCT images, demonstrating performance metrics comparable to expert manual tracings [16]. In the context of MC segmentation, published research has reported AI-based algorithms achieving accuracy rates of up to 99% [17,18]. However, these studies exhibit several limitations, including inconsistencies in the use and reporting of standardised reference methods [16]. Additionally, while the relationship between the third molar and the MC has been explored with promising results using AI [19], no studies have specifically evaluated the influence of third molar presence (erupted or impacted) or absence on the accuracy of AI-driven segmentation.

This study aims to address these gaps by systematically evaluating the precision and quality of AI-powered MC segmentation in comparison to semi-automated radiologist tracings using a standardised methodology. This segmentation, serving as the ground truth reference, involved manually annotating sequential points along the MC trajectory on cross-sectional CBCT views. Building on calibrated parameters from prior research and addressing methodological limitations identified in our pilot investigations, we seek to assess the accuracy of the AI system in this context. Furthermore, we aim to analyse the impact of third molar status on segmentation accuracy. Given the increasing clinical adoption of AI-based tools, rigorous performance evaluation is essential to validate their reliability and applicability in clinical practice.

Material and methods

Image dataset

A total of 150 anonymised CBCT scans from 150 patients (70 males, 80 females) aged 18 to 71 years were retrospectively retrieved from the database of Poznan University of Medical Sciences, Poland. These scans, originally acquired between 2020 and 2023 for preoperative planning of dental implant placement and third molar extractions, were selected based on the following inclusion criteria: patient age must be equal to or above 18 years, the field of view (FOV) must fully encapsulate the mandible to ensure comprehensive visualisation of the MC and absence of artifacts or pathological anomalies impairing osseous or neurovascular architecture. Scans with insufficient FOV, artifacts, or structural pathologies compromising mandibular integrity were excluded.

All imaging was performed using a CRANEX 3D dental imaging system (Soredex, Milwaukee, USA), with datasets archived in Digital Imaging and Communications in Medicine (DICOM) format. Acquisition parameters adhered to a standardised protocol across all scans: X-ray tube voltage 90 kV, tube current 10 mA, isotropic voxel resolution 0.25 mm, and FOV dimensions ranging from 600 mm × 800 mm to 1600 mm × 1300 mm. These

settings were optimised to balance radiation dose minimisation with diagnostic fidelity, ensuring high-contrast resolution for MC segmentation.

Segmentation protocol

The semi-automated segmentation method, widely used and clinically accepted, was chosen as the reference method to assess AI performance. Semi-automated segmentation of the MC was performed using Romexis software (v. 6.2; Planmeca, Helsinki, Finland). Specific points along the canal pathway were manually placed by the investigator on cross-sectional slices. The software then interpolated these markers to generate a continuous cylindrical pathway (diameter: 1.50 mm).

Segmentations were conducted by an experienced oral and maxillofacial radiologist (with more than 10 years of clinical practice) and a trainee in oral and maxillofacial radiology. To ensure consistency, the operators were calibrated and used the following settings: slice thickness 0.25 mm, sharpness set to zero, contrast set to 1200, and brightness value fixed to 1808. These parameters were chosen based on a study by Issa *et al.* [20], which identified them as optimal for MC visualisation on CBCT images using Romexis software (v. 6.2; Planmeca, Helsinki, Finland). To assess reproducibility, the segmentation process for all scans was repeated on 2 separate occasions with a 10-day interval.

All segmentations were performed on an NEC MultiSync EA245Wmi-2 display screen (Sharp NEC Display Solutions, Tokyo, Japan) under optimal ambient lighting conditions. Bilateral MCs from each scan were then exported as a unified stereolithography (STL) file for subsequent analysis (Figure 1).

For the automated segmentation, the same CBCT datasets were processed using Diagnocat (DGNCT LLC, Miami, USA), a cloud-based AI platform utilising a U-Net-like CNN architecture. The CBCT images were uploaded to the software, and the automatic segmentation function was activated to segment the entire scanned region. The segmented bilateral MCs in each image, generated automatically, were then exported as an STL file for analysis (Figure 1).

Evaluation of 3D models

Cloud Compare v.2.13.alpha (open-source software, <http://www.cloudcompare.org/>) was used to perform the overlap and visualisation of the segmented MCs generated by both the investigators and the AI-based software for the same CBCT image. This process aimed to assess the segmentation accuracy of the AI in comparison to the semi-automatic segmentations performed by the investigators. The evaluation focused on quantifying volumetric deviations to determine the level of agreement between the segmentation methods.

As part of this analysis, each investigator's segmented MC (STL file) was individually overlapped with the AI-generated segmentation from the corresponding scan. The computed distances between the models were then averaged to obtain a global average measure of segmentation deviation. The process began with a pre-registration step using the 3-points method, in which 3 anatomically relevant landmarks, the mandibular foramen, molar region, and premolar region, were manually placed on both 3D models to ensure proper spatial alignment (Figure 2).

Following this alignment, the "compute cloud/mesh distance" function in Cloud Compare was applied to calculate the spatial deviations between the overlapping 3D models. This function provided numerical outputs, including the average distance and maximum distances between the 2 surfaces. The software's default overlap parameter was set to 100, ensuring a full 3D alignment for precise deviation analysis.

Statistical analysis

Statistical analyses were conducted using SPSS 29.0 (SPSS Inc., Chicago, IL, USA). The inter-rater and intra-rater reliability of the semi-automatic segmentation of the MCs performed by both investigators was assessed using a two-way mixed model intraclass correlation coefficient (ICC) for absolute agreement, evaluating agreement both between investigators and within each investigator across the 2 evaluation sessions.

To compare the MC segmentation methods, volumetric deviations were assessed by calculating the overall average distance and standard deviation, which provided

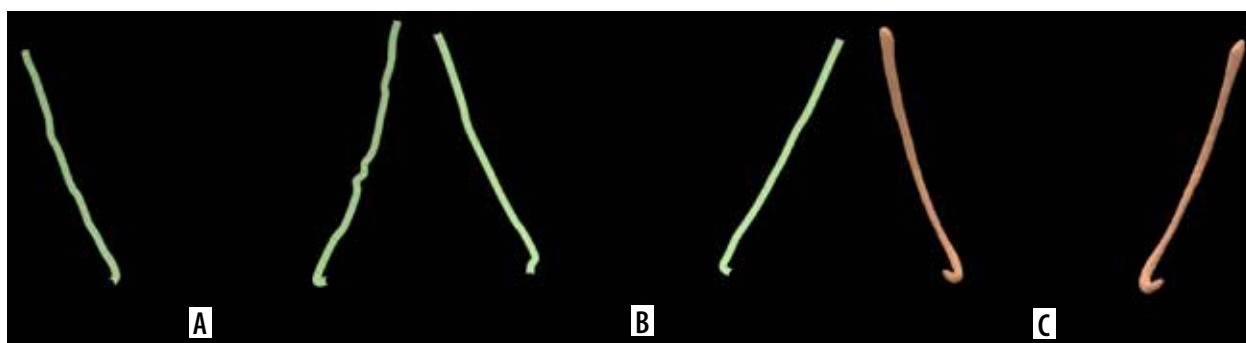


Figure 1. Mandibular canal segmentation. A) Semi-automatic segmentation performed by the first investigator. B) Semi-automatic segmentation performed by the second investigator. C) Automatic segmentation performed by Diagnocat

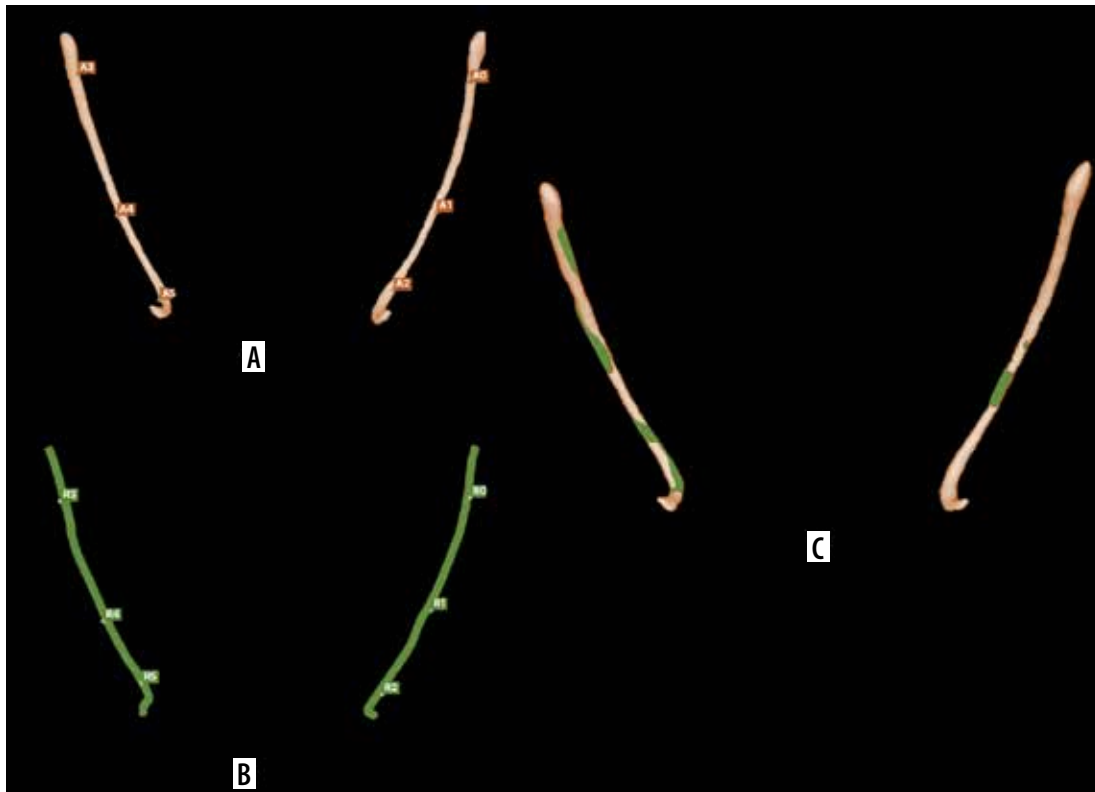


Figure 2. Alignment of mandibular canal segmentation (STL file). A) Automatic segmentation performed by Diagnocat with 3 points for alignment. B) Semi-automatic segmentation performed by the second investigator with 3 points for alignments. C) Aligned and overlapped structures

a quantitative measure of segmentation accuracy for each image. The computed distance between each operator and the AI was recorded, and the average distance for each scan was then calculated.

Average distance per scan

For each scan n , the average distance was determined by taking the mean of the distances obtained from both operator-AI comparisons:

$$\bar{d}_n = \frac{\bar{d}_{1n} + \bar{d}_{2n}}{2}$$

- $\bar{D}_{1n}$ is the computed mean distance between the first investigator and AI for scan n ,
- $\bar{D}_{2n}$ is the computed mean distance between the second investigator and AI for scan n .

Descriptive statistics were applied to the numerical results of the 3D evaluation, with data summarised as average $\pm$ standard deviation.

The normality of numerical variables was assessed using the Shapiro-Wilk test ($p < 0.05$).

To compare segmentation deviations across third molar status categories (absent, erupted, impacted), the Kruskal-Wallis test was applied. Post hoc pairwise comparisons were conducted using the Mann-Whitney U test with Bonferroni correction to adjust for multiple comparisons.

Results

Regarding third molar status, 60 scans on the left side showed the absence of a third molar, 34 had an erupted third molar (with the crown positioned above the alveolar ridge), and 56 presented with an impacted third molar (with the crown partially or fully located below the alveolar ridge). On the right side, 70 scans showed no third molar, 25 had an erupted third molar, and 55 had an impacted third molar. Each scan contributed bilateral MC segmentations, resulting in a total of 300 MCs for evaluation. Inter- and intra-rater reliability assessments were conducted to quantify the consistency of semi-automated MC segmentation (Table 1). Inter-rater reliability, calculated as the concordance between 2 independent operators, demonstrated substantial agreement, averaging 84.5%. Given the close inter-rater values between day 0 and day 10, the results of the repeated semi-automated segmentation on day 10 were considered for evaluation. Intra-rater reliability, reflecting the reproducibility of segmentations by the same operator across repeated trials, achieved near-perfect agreement, averaging 95.5%. This underscores the precision and stability of the semi-automated protocol under controlled conditions.

Three-dimensional spatial deviations between AI-based and semi-automated MC segmentations were quantified using surface-to-surface distance metrics. The Shapiro-Wilk test confirmed a non-normal distribution

Table 1. Inter- and intra-rater reliability analysis

	Category	ICC (%)
Inter-rater reliability	Day 0	84
	Day 10	85
Intra-rater reliability	OMFR	94
	Trainee in OMFR	97

ICC – intraclass correlation coefficient, OMFR – oral and maxillofacial radiologist

($p < 0.05$), and deviations were summarised using the median. The overall median deviation was 0.29 mm, with a standard deviation ranging from 0.25 to 0.37 mm. The per-scan average distance varied from 0.19 mm to 4.72 mm (Supplementary material).

The box plot analysis (Figure 3) reveals distinct trends in segmentation accuracy across third molar categories. The absent category has the lowest deviation (0.27 mm), while the erupted category shows a minor increase (0.28 mm). The impacted category exhibits the highest deviations (0.32 mm). In terms of laterality, deviations are slightly higher on the left side than on the right.

Discussion

The aim of this study was to evaluate the accuracy and variability of a commercial AI-based tool (Diagnocat) for MC segmentation, with a particular focus on factors influencing segmentation precision, including the status of third molars. Building upon the pilot study [21], this investigation provides further insights into AI-driven segmentation accuracy by incorporating a more diverse cohort of patients and optimising CBCT image viewer settings [20]. While previous studies have explored AI performance in mandibular MC segmentation [17,18], the current study offers a detailed assessment of anatomical factors, such as the presence of impacted third molars, that may affect AI accuracy. Notably, we did not use DentalComs [22] for this study. Instead, our focus was on evaluating the overall volumetric conformity of overlapping structures by calculating the average distance. This metric, which quantifies the deviation between corresponding points on AI-generated and semi-automated segmentations, is particularly useful for assessing complex, non-regular anatomical structures like the MC [23-25].

Our results revealed a median deviation of 0.29 mm between AI and semi-automated expert segmentations, indicating strong overall agreement. This performance surpasses the 0.555 mm and 0.39 mm mean curve distance (MCD) reported by Jaskari *et al.* [26] and Järnstedt *et al.* [27], respectively. Unlike MCD, which quantifies contour-level discrepancies, our average distance metric evaluates deviations across the entire 3D surface, offering a more holistic assessment of volumetric alignment critical for complex structures like the MC. The observed variability of the standard deviation ranging from 0.25 to

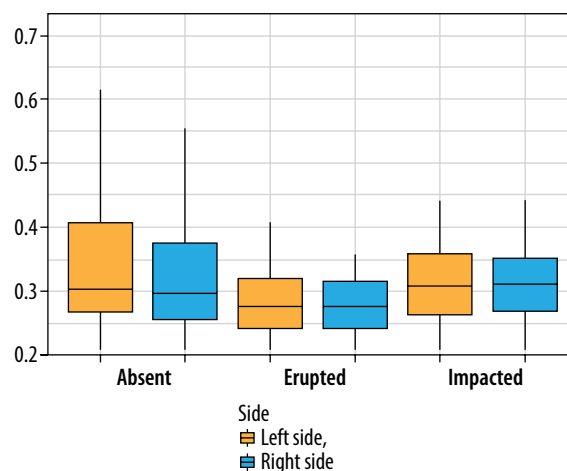


Figure 3. Box plot showing the average distance between experts and AI, stratified by third molar status (absent, erupted, impacted) and location (left side, right side). The central horizontal line represents the median, the box denotes the interquartile range (IQR), and the whiskers extend to 1.5 times the IQR

0.37 may reflect anatomical diversity in our cohort, a factor underrepresented in prior studies. This underscores the importance of training AI models on large heterogeneous datasets to improve robustness in clinically challenging scenarios.

Clinically, deviations ≤ 0.50 mm are considered acceptable for dental implant planning [28], and in our study, 88% of AI segmentations met this threshold, demonstrating strong potential for integration into preoperative workflows. However, the presence of outliers highlights the necessity of expert evaluation, particularly in anatomically complex regions, to improve segmentation accuracy and compensate for errors caused by under-segmentation and partial volume effects. Notably, scans 87 and 95 exhibited significant deviations, with average distances of $1.93 \text{ mm} \pm 3.47$ and $4.72 \text{ mm} \pm 3.87$, respectively.

In scan 87 (Figure 4), AI-based segmentation of the MC was disrupted, probably due to the partial volume effect, which occurs when a voxel contains a mixture of different tissue types, leading to blurred boundaries and reduced contrast [29]. This effect may have impaired the ability of Diagnocat to distinguish the canal from surrounding bone, particularly in areas where the canal is narrow or obscured by anatomical complexity near the impacted third molar. Additionally, other factors, such as beam hardening artifacts or limited training data for complex cases, may have further contributed to the segmentation failure. Conversely, under-segmentation occurs when critical anatomical details are omitted from the AI-based segmentation [30]. A notable example is seen in scan 95 (Figure 5), where Diagnocat failed to segment the portion of the MC extending from the molar root region to the mental foramen. This limitation may stem from insufficient anatomical diversity in the AI's training dataset, reducing its accuracy when encountering unfamiliar or complex anatomical patterns. In such cases, manual inter-

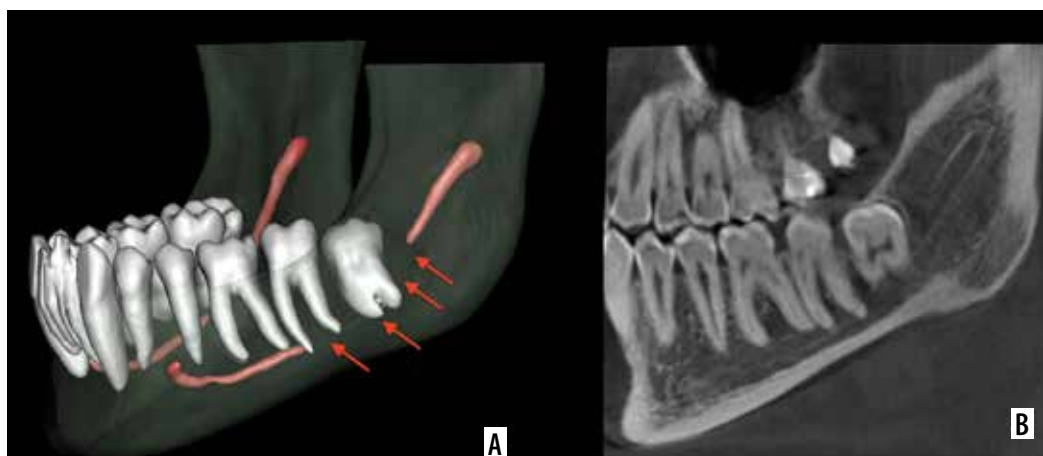


Figure 4. Mandibular canal of scan 87 with partial volume effect. A) AI-based segmentation. B) Sagittal view on CBCT image

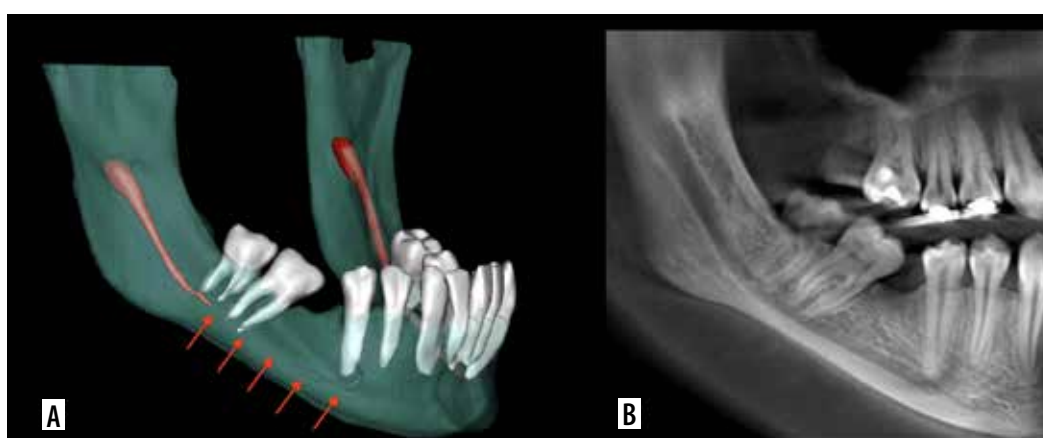


Figure 5. Mandibular canal of scan 95. A) AI-based under-segmentation. B) Sagittal view on CBCT image

vention by clinicians remains essential to ensure accurate segmentation and prevent potential surgical complications.

An important finding in our study was the significant impact of third molar status on AI segmentation accuracy. The box plot analysis demonstrated that AI segmentation was most accurate when the third molar was absent (median: 0.27 mm) or erupted (median: 0.28 mm), whereas the accuracy declined in cases with impacted third molars (median: 0.32 mm). This reduction in accuracy can be attributed to the segmentation approach used by Diagnocat, which, like many AI-based tools, relies on a straight-line pattern based on the radiolucent trajectory of the MC [19]. While effective when the canal is clearly distinguishable, impacted third molars introduce additional radiodense structures that obscure the canal's path, leading to segmentation errors. This limitation is particularly critical in clinical settings where the roots of impacted third molars overlap with the mandibular canal, because AI performs worst in these cases, precisely where accurate segmentation is most crucial for surgical planning. To address this limitation, AI models should be trained on more diverse datasets that include a wide range of third-molar impaction scenarios.

While there is limited literature directly assessing the influence of third molar status on MC segmentation accu-

racy, several studies have evaluated the capability of AI in detecting third molars and their relationship with the MC. For instance, Orhan *et al.* [31] assessed the diagnostic performance of Diagnocat for third molar evaluation on CBCT, reporting high accuracy. Additionally, while multiple studies demonstrate AI's promising ability to assess third molar and MC spatial relationships on panoramic radiographs [32-34], Kazimierczak *et al.* [35] highlighted its limitations in evaluating root apex and MC proximity on computed tomography images, reporting low diagnostic accuracy. These findings suggest that while AI performs well in detecting third molars, its ability to accurately segment the MC in cases of impaction remains an area for further investigation.

Despite the strengths of our study, including a standardised methodology, large sample size ($n = 150$), and robust comparison to semi-automated expert segmentations, there are notable limitations. The retrospective design and single-centre dataset may limit the generalisability of our results. Additionally, the proprietary nature of Diagnocat restricts transparency regarding the decision-making process. Open-source AI models could offer better insight into segmentation mechanisms, enabling improvements in algorithmic performance. Furthermore, our dataset excluded cases with artifacts and pathologies,

which may lead to an overestimation of AI performance in real-world scenarios. Future research should focus on incorporating multi-centre cohorts, expanding datasets to include more anatomically diverse cases (particularly those with impacted third molars), and integrating open-source AI models and advanced image-processing techniques to enhance segmentation accuracy, generalisability, and reliability in complex anatomical regions.

Conclusions

AI-driven segmentation demonstrated high accuracy with low median deviation (0.29 mm), with 88% of cases within clinically acceptable limits. However, segmentation errors increased in cases with impacted third molars, emphasising the need for diverse AI training datasets. Future research should focus on multi-centre validation and advanced image-processing techniques to improve robustness in complex cases.

Disclosures

1. Institutional review board statement: The study was conducted in accordance with the Declaration of Helsinki and exempted from ethical approval on the grounds of its observational nature. The Ethics Committee of Poznan University of Medical Sciences granted a waiver of patient consent for the use of anonymised radiographs in this retrospective study.
2. Assistance with the article: None.
3. Financial support and sponsorship: Julien Issa is a participant of the STER program – Internationalisation of Doctoral Schools from the Polish National Agency for Academic Exchange (NAWA) – and received grant No. PPI/STE/2020/1/00014/DEC/02.
4. Conflicts of interest: None.

References

1. Carter JB, Stone JD, Clark RS, Mercer JE. Applications of cone-beam computed tomography in oral and maxillofacial surgery: an overview of published indications and clinical usage in United States academic centers and oral and maxillofacial surgery practices. *J Oral Maxillofac Surg* 2016; 74: 668-679.
2. Schulze RKW, Drage NA. Cone-beam computed tomography and its applications in dental and maxillofacial radiology. *Clin Radiol* 2020; 75: 647-657.
3. Kaasalainen T, Ekholm M, Siiskonen T, Korttinen M. Dental cone beam CT: an updated review. *Phys Med* 2021; 88: 193-217.
4. Agbaje JO, de Castele EV, Salem AS, Anumendem D, Lambrichts I, Politis C. Tracking of the inferior alveolar nerve: its implication in surgical planning. *Clin Oral Investig* 2017; 21: 2213-2220.
5. Shankland WE 2nd. The trigeminal nerve. Part IV: the mandibular division. *Cranio* 2001; 19: 153-161.
6. Jacobs R, Quirynen M, Bornstein MM. Neurovascular disturbances after implant surgery. *Periodontol* 2000 2014; 66: 188-202.
7. Sarikov R, Juodzbalys G. Inferior alveolar nerve injury after mandibular third molar extraction: a literature review. *J Oral Maxillofac Res* 2014; 5: e1. DOI: 10.5037/jomr.2014.5401.
8. Li Y, Ling Z, Zhang H, Xie H, Zhang P, Jiang H, Fu Y. Association of the inferior alveolar nerve position and nerve injury: a systematic review and meta-analysis. *Healthcare (Basel)* 2022; 10: 1782. DOI: 10.3390/healthcare10091782.
9. Poort LJ, van Neck JW, van der Wal KG. Sensory testing of inferior alveolar nerve injuries: a review of methods used in prospective studies. *J Oral Maxillofac Surg* 2009; 67: 292-300.
10. Li Y, Ling Z, Zhang H, Xie H, Zhang P, Jiang H, Fu Y. Association of the inferior alveolar nerve position and nerve injury: a systematic review and meta-analysis. *Healthcare (Basel)* 2022; 10: 1782. DOI: 10.3390/healthcare10091782.
11. organ N, Meeus J, Shujaat S, Cortellini S, Bornstein MM, Jacobs R. CBCT for diagnostics, treatment planning and monitoring of sinus floor elevation procedures. *Diagnostics (Basel)* 2023; 13: 1684. DOI: 10.3390/diagnostics13101684.
12. Xiang B, Lu J, Yu J. Evaluating tooth segmentation accuracy and time efficiency in CBCT images using artificial intelligence: a systematic review and meta-analysis. *J Dent* 2024; 146: 105064. DOI: 10.1016/j.jdent.2024.105064.
13. Zheng Q, Gao Y, Zhou M, Li H, Lin J, Zhang W, Chen X. Semi or fully automatic tooth segmentation in CBCT images: a review. *PeerJ Comput Sci* 2024; 10: e1994. DOI: 10.7717/peerj-cs.1994.
14. Najjar R. Redefining radiology: a review of artificial intelligence integration in medical imaging. *Diagnostics (Basel)* 2023; 13: 2760. DOI: 10.3390/diagnostics13172760.
15. Waller J, O'Connor A, Razaat E, Amireh A, Dempsey J, Martin C, Umair M. Applications and challenges of artificial intelligence in diagnostic and interventional radiology. *Pol J Radiol* 2022; 87: e113-e117. DOI: 10.5114/pjr.2022.113531.
16. Abesi F, Jamali AS, Zamani M. Accuracy of artificial intelligence in the detection and segmentation of oral and maxillofacial structures using cone-beam computed tomography images: a systematic review and meta-analysis. *Pol J Radiol* 2023; 88: e256-e263. DOI: 10.5114/pjr.2023.127624.
17. Naufal MF, Fatichah C, Astuti ER, Putra RH. Deep learning for mandibular canal segmentation in digital dental radiographs: a systematic literature review. *IEEE Access* 2024; 12: 76794-76815. DOI: 10.1109/ACCESS.2024.3406342.
18. Issa J, Olszewski R, Dyszkiewicz-Konwinska M. The effectiveness of semi-automated and fully automatic segmentation for inferior alveolar canal localization on CBCT scans: a systematic review. *Int J Environ Res Public Health* 2022; 19: 560. DOI: 10.3390/ijerph19010560.
19. Assiri HA, Hameed MS, Alqarni A, Dawasaz AA, Arem SA, Assiri KI. Artificial intelligence application in a case of mandibular third molar impaction: a systematic review of the literature. *J Clin Med* 2024; 13: 4431. DOI: 10.3390/jcm13154431.

20. Issa J, Riad A, Olszewski R, Dyszkiewicz-Konwińska M. The influence of slice thickness, sharpness, and contrast adjustments on inferior alveolar canal segmentation on cone-beam computed tomography scans: a retrospective study. *J Pers Med* 2023; 13: 1518. DOI: 10.3390/jpm13101518.
21. Issa J, Kulczyk T, Rychlik M, Czajka-Jakubowska A, Olszewski R, Dyszkiewicz-Konwińska M. Artificial intelligence versus semi-automatic segmentation of the inferior alveolar canal on cone-beam computed tomography scans: a pilot study. *Dent Med Probl* 2024; 61: 893-899. DOI: 10.17219/dmp/175968.
22. Büttner M, Rokhsad R, Brinz J, et al. Core outcome measures in dental computer vision studies (DentalCOMS). *J Dent* 2024;150: 105318. DOI: 10.1016/j.jdent.2024.105318.
23. Fenster A, Chiu B. Evaluation of segmentation algorithms for medical imaging. *Conf Proc IEEE Eng Med Biol Soc* 2005; 7186-7189. DOI: 10.1109/IEMBS.2005.1616166.
24. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 2015; 15: 29. DOI: 10.1186/s12880-015-0068-x.
25. Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group no. 132. *Med Phys* 2017; 44: e43-e76. DOI: 10.1002/mp.12256.
26. Jaskari J, Sahlsten J, Järnstedt J, et al. Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes. *Sci Rep* 2020; 10: 5842. DOI: 10.1038/s41598-020-62321-3.
27. Järnstedt J, Sahlsten J, Jaskari J, et al. Comparison of deep learning segmentation and multigrader-annotated mandibular canals of multicenter CBCT scans. *Sci Rep* 2022; 12: 18598. DOI: 10.1038/s41598-022-20605-w.
28. Lysenko AV, Yaremenko AI, Ivanov VM, et al. Comparison of dental implant placement accuracy using a static surgical guide, a virtual guide and a manual placement method – an in-vitro study. *Ann Maxillofac Surg* 2023; 13: 158-162. DOI: 10.4103/ams.ams_100_23.
29. Tohka J. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: a review. *World J Radiol* 2014; 6: 855-864. DOI: 10.4329/wjr.v6.i11.855.
30. Juneja M, Aggarwal N, Saini S, et al. A comprehensive review on artificial intelligence-driven preprocessing, segmentation, and classification techniques for precision furcation analysis in radiographic images. *Multimed Tools Appl* 2024. DOI: 10.1007/s11042-024-19920-3.
31. Orhan K, Bilgir E, Bayrakdar IS, Ezhov M, Gusarev M, Shumilov E. Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans. *J Stomatol Oral Maxillofac Surg* 2021; 122: 333-337. DOI: 10.1016/j.jomas.2020.12.006.
32. Kempers S, van Lierop P, Hsu TH, Moin DA, Bergé S, Ghaemina H, Xi T, Vinayahalingam S. Positional assessment of lower third molar and mandibular canal using explainable artificial intelligence. *J Dent* 2023; 133: 104519. DOI: 10.1016/j.jdent.2023.104519.
33. Joo Y, Moon SY, Choi C. Classification of the relationship between mandibular third molar and inferior alveolar nerve based on generated mask images. *IEEE Access* 2023; 11: 81777-81786. DOI: 10.1109/ACCESS.2023.3302271.
34. Fukuda M, Aiji Y, Kise Y, Nozawa M, Kuwada C, Funakoshi T, et al. 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* 2020; 130: 336-343. DOI: 10.1016/j.oooo.2020.04.005.
35. Kazimierczak W, Kazimierczak N, Kędziora K, Szcześniak M, Serafin Z. Reliability of the AI-assisted assessment of the proximity of the root apices to mandibular canal. *J Clin Med* 2024; 13: 3605. DOI: 10.3390/jcm13123605.