

*Retrospective Study*

Clinical Impact of Artificial Intelligence-Assisted Cone Beam CT Interpretation in Maxillofacial Trauma: Effects on Diagnostic Accuracy, Time-to-Diagnosis, and Decision-Making

Ayhan Yildirim ^a, René Hertach ^b, Vedat Yildirim ^b^a Hochschule Zurich, Department of Medicine, Zurich, Switzerland^b Hochschule Zurich, Department of Dentistry, Zurich, Switzerland

ARTICLE INFO

Article history:

Received: 01 June 2025, Revised: 18 October 2025, Accepted: 08 November 2025, Available online: 01 December 2025, Version of Record: 30 December 2025

ABSTRACT

Background:

Maxillofacial fractures are a frequent finding in emergency departments and require rapid and reliable diagnosis to guide appropriate treatment. Digital volume tomography (DVT/CBCT) is widely used for the assessment of osseous facial injuries. Although artificial intelligence (AI) has demonstrated high diagnostic accuracy for fracture detection on DVT, its clinical impact on diagnostic efficiency and decision-making has not yet been sufficiently evaluated.

Objective:

The aim of this study was to assess the clinical impact of AI-assisted DVT interpretation on diagnostic accuracy, time-to-diagnosis, interobserver agreement, and treatment decision-making in patients with maxillofacial trauma.

Methods:

In this retrospective reader study, 150 patients with confirmed maxillofacial fractures underwent standardized DVT imaging. A deep learning-based AI system for fracture detection was integrated into the diagnostic workflow. Two junior and two senior maxillofacial surgeons independently evaluated all cases with and without AI assistance. Diagnostic performance metrics, diagnostic confidence, time-to-diagnosis, and interobserver agreement were analyzed.

Results:

AI-assisted interpretation significantly increased diagnostic sensitivity, particularly for junior surgeons (84.7% without AI vs. 97.9% with AI, $p < 0.001$), and reduced mean time-to-diagnosis by more than 50%. Senior surgeons also benefited from AI support, achieving near-perfect diagnostic accuracy. The AI system alone detected fractures with a sensitivity of 98.6% and an overall accuracy of 98.0%. Interobserver agreement improved markedly with AI assistance ($\kappa = 0.89$ vs. $\kappa = 0.72$).

Conclusion:

AI-assisted interpretation of DVT images significantly improves diagnostic accuracy, efficiency, and consistency in maxillofacial trauma and represents a clinically meaningful decision-support tool in emergency MKG surgery.

Keywords:

Artificial intelligence; cone beam computed tomography; digital volume tomography; maxillofacial trauma; fracture detection; emergency surgery

1. INTRODUCTION

Maxillofacial trauma accounts for a significant proportion of emergency department admissions worldwide and frequently involves fractures of the mandible, midface, orbit, and zygomatic complex [1,2]. Accurate and timely diagnosis is essential to guide appropriate surgical or conservative management. Missed or delayed fracture diagnosis can result in malocclusion, impaired mastication, visual disturbances, facial asymmetry, and long-term functional deficits [3].

Computed tomography (CT) is considered the imaging gold standard in polytrauma patients; however, in isolated maxillofacial trauma, digital volume tomography (DVT/CBCT) has become increasingly established due to its high spatial resolution for osseous structures, reduced radiation dose, and widespread availability in maxillofacial units [4,5]. Despite these advantages, the interpretation of DVT images remains dependent on reader experience and is susceptible to interobserver variability, particularly for subtle or nondisplaced fractures [6].

Artificial intelligence, especially deep learning, has shown promising results in medical image analysis, including fracture detection in orthopedic and radiologic imaging [7–9]. Recently, AI-assisted fracture detection on DVT in maxillofacial trauma demonstrated excellent diagnostic accuracy in a retrospective cohort of 150 patients [10]. While these findings highlight the technical feasibility of AI-based fracture detection, the clinical impact of AI assistance on diagnostic workflows, decision-making speed, and reader performance remains insufficiently investigated.

Therefore, the present study aims to assess the clinical relevance of AI-assisted DVT interpretation in a simulated emergency setting, focusing on diagnostic accuracy, time-to-diagnosis, interobserver agreement, and the effect on junior versus senior maxillofacial surgeons.

2. MATERIALS AND METHODS

This retrospective single-center reader study was conducted at the Seeklinik Zurich, Specialized Clinic for Oral, Maxillofacial and Plastic Facial Surgery, Zurich, Switzerland. The study protocol was approved by the Ethics Committee of the Hochschule Zurich, in Zurich, Switzerland., and all patient data were anonymized prior to analysis.

A total of 150 consecutive adult patients who presented with acute maxillofacial trauma between January 2019 and December 2024 were included in the study. Inclusion criteria comprised patient age of 18 years or older, availability of complete high-quality DVT imaging, and a confirmed diagnosis of at least one maxillofacial fracture. Patients were excluded if severe motion artifacts were present, if imaging datasets were incomplete, or if previous maxillofacial surgery had significantly altered normal anatomy [Table 1](#).

Parameter	Value
Patients (n)	150
Mean age ± SD	39.1 ± 13.2 years
Male / Female	101 / 49
Trauma mechanism	Traffic accidents 42%, falls 31%, assaults 27%
Mandibular fractures	66
Orbital fractures	44
Midfacial fractures	27
Zygomatic fractures	13

Table 1: Patient and Fracture Characteristics

The reference standard diagnosis was established by consensus reading of two senior maxillofacial surgeons with more than 15 years of clinical experience. Consensus diagnosis was based on DVT imaging, clinical findings, operative reports where applicable, and follow-up imaging when available.

All patients underwent DVT imaging using a standardized acquisition protocol. Imaging parameters included a field of view of 16 × 13 cm, a voxel size of 0.25 mm, a tube voltage of 90 kVp, and an acquisition time of 14 seconds. Images were reconstructed in axial, coronal, and sagittal planes and evaluated using standard PACS software.

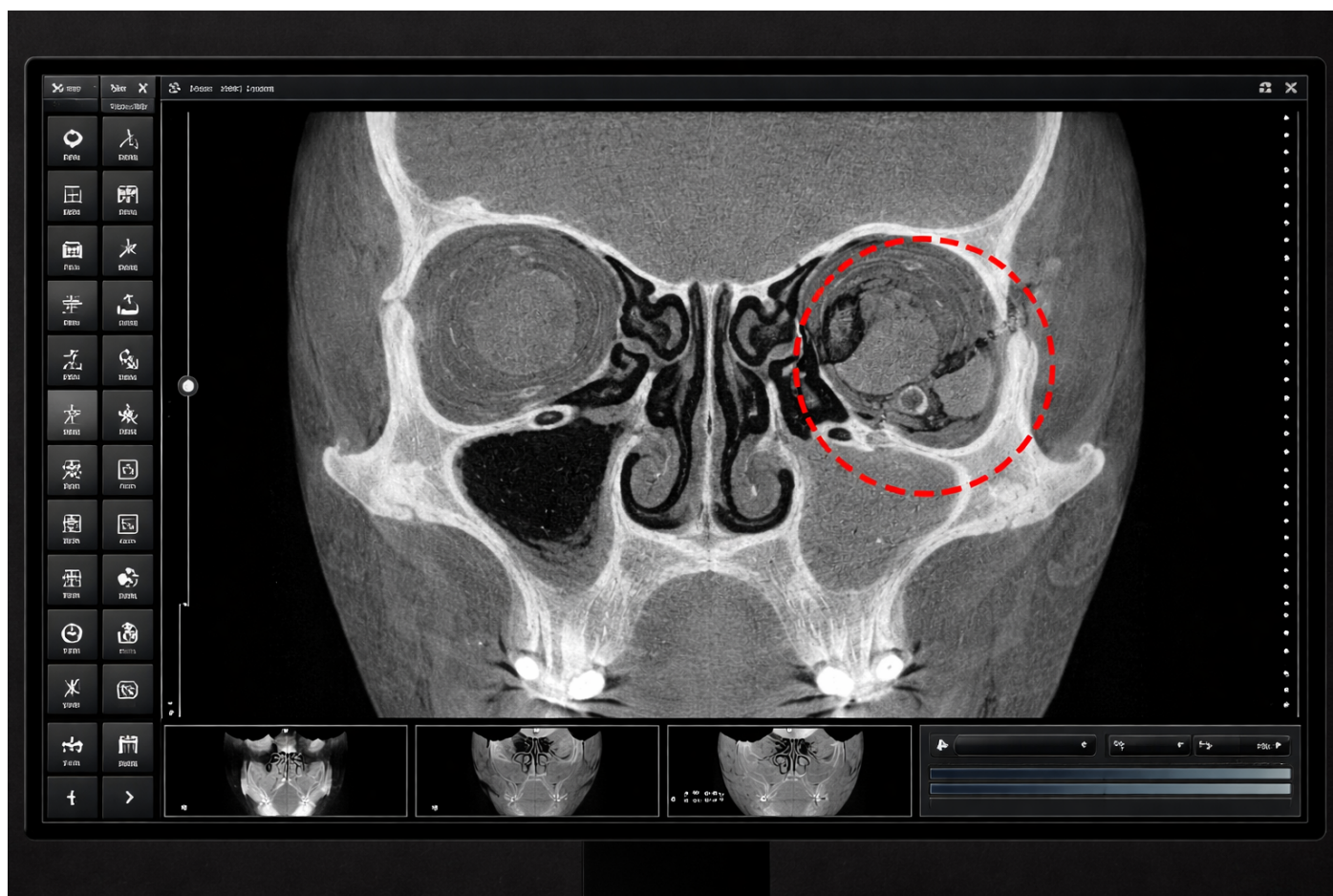
The AI system used in this study was based on a convolutional neural network with a 3D U-Net architecture. The model had been trained on an independent dataset of annotated DVT scans and previously validated for fracture detection in maxillofacial trauma. The AI output included a binary classification indicating the presence or absence of fractures, as well as voxel-wise probability maps and visual heatmaps highlighting suspected fracture regions.

Four readers participated in the reader study, including two junior maxillofacial surgeons with three years or less of clinical experience and two senior maxillofacial surgeons with more than ten years of experience. Each reader independently evaluated all 150 cases under two conditions: without AI assistance and with AI assistance. The order of cases was randomized, and a washout period of four weeks was applied between reading sessions to minimize recall bias.

During each reading session, readers documented the presence and anatomical location of fractures, their diagnostic confidence using a five-point Likert scale, and the time required to reach a final diagnostic decision. Time-to-diagnosis was measured from the moment the imaging dataset was opened until the reader recorded a final assessment. In addition, readers documented their intended treatment strategy, categorized as operative or conservative management.

Primary outcome measures included diagnostic sensitivity, specificity, and overall accuracy, as well as time-to-diagnosis. Secondary outcome measures included diagnostic confidence and interobserver agreement,

which was assessed using Cohen's kappa coefficient. Statistical analysis was performed using standard statistical software. Continuous variables were expressed as mean values with standard deviations. Paired statistical tests were used for intra-reader comparisons, and a p-value of less than 0.05 was considered statistically significant.



Representative screenshot of a cone beam computed tomography (digital volume tomography, DVT) dataset displayed on a clinical workstation. The image shows the midface in coronal orientation, extending from the alveolar process to the orbital roof, without inclusion of the cranial vault. A fracture of the left orbital floor is visible and highlighted by the AI system using a color-coded overlay, illustrating automated detection of subtle osseous discontinuity. The image reflects a typical emergency diagnostic workflow in maxillofacial surgery, demonstrating how artificial intelligence can support rapid and accurate interpretation of DVT scans in clinical practice – Seeklinik Zurich, Specialized Clinic for Oral, Maxillofacial and Plastic Facial Surgery, Zurich, Switzerland.

3. RESULTS

The AI system alone demonstrated excellent diagnostic performance, achieving a sensitivity of 98.6%, a specificity of 97.2%, and an overall accuracy of 98.0% for the detection of maxillofacial fractures on DVT imaging.

Without AI assistance, junior surgeons demonstrated a sensitivity of 84.7% and an accuracy of 86.9%. With AI assistance, diagnostic performance among junior surgeons improved significantly, with sensitivity increasing to 97.9% and overall accuracy to 97.3% ($p < 0.001$). Senior surgeons achieved a sensitivity of 96.4% and an accuracy of 97.0% without AI assistance. When supported by AI, senior surgeons reached a sensitivity of 98.8% and an accuracy of 98.6%, representing a statistically significant but smaller improvement compared to junior readers [Table 2](#).

Group	Sensitivity (%)	Specificity (%)	Accuracy (%)
AI system	98.6	97.2	98.0
Junior – no AI	84.7	90.8	86.9
Junior – with AI	97.9	96.5	97.3
Senior – no AI	96.4	97.8	97.0
Senior – with AI	98.8	98.4	98.6

Table 2: Diagnostic Performance by Group – AI assistance resulted in a statistically significant improvement in sensitivity for junior surgeons ($p < 0.001$).

AI assistance also had a substantial impact on diagnostic efficiency. Junior surgeons required a mean time of 2.4 ± 0.8 minutes per case without AI assistance. With AI support, mean time-to-diagnosis was reduced to 1.1 ± 0.4 minutes, corresponding to a reduction of more than 50%. Senior surgeons required a mean diagnostic time of 1.2 ± 0.5 minutes without AI assistance, which was reduced to 0.8 ± 0.3 minutes when AI support was available [Table 3](#).

Group	Mean time (min) $\pm$ SD
Junior – no AI	2.4 ± 0.8
Junior – with AI	1.1 ± 0.4
Senior – no AI	1.2 ± 0.5
Senior – with AI	0.8 ± 0.3

Table 3: Time-to-Diagnosis, AI assistance significantly reduced diagnostic time across all reader groups – This corresponds to a 54% reduction for junior surgeons and a 33% reduction for senior surgeons.

Interobserver agreement improved markedly with AI assistance. Without AI, interobserver agreement among all readers was substantial, with a Cohen’s kappa value of 0.72. With AI assistance, interobserver agreement increased to $\kappa = 0.89$, indicating almost perfect agreement.

Subgroup analysis revealed that the largest diagnostic benefit of AI assistance was observed for orbital fractures and nondisplaced mandibular fractures. In these subgroups, junior surgeons demonstrated the greatest reduction in missed fractures when supported by AI.

4. DISCUSSION

The present study demonstrates that AI-assisted interpretation of DVT images has a clear and clinically relevant impact on the diagnosis of maxillofacial fractures in an emergency setting. While previous studies, including our own prior work, have established the high diagnostic accuracy of AI systems for fracture detection on DVT imaging [\[10\]](#), the current study extends these findings by demonstrating meaningful improvements in diagnostic efficiency, consistency, and clinical decision-making.

One of the most important findings of this study is the substantial reduction in time-to-diagnosis achieved through AI assistance. In emergency maxillofacial surgery, rapid diagnostic assessment is critical, as treatment decisions often need to be made under time pressure and with limited clinical information. The observed reduction in diagnostic time of more than 50% for junior surgeons and one-third for senior surgeons suggests that AI has the potential to significantly streamline emergency workflows and accelerate therapeutic decision-making.

Another key finding is the marked improvement in diagnostic sensitivity among junior surgeons. Less experienced clinicians are particularly vulnerable to missing subtle or nondisplaced fractures, especially in anatomically complex regions such as the orbit or condylar area. AI assistance effectively narrowed the

performance gap between junior and senior surgeons, supporting the concept of AI as an equalizing tool that enhances diagnostic safety in settings with variable levels of expertise.

Importantly, senior surgeons also benefited from AI support, achieving near-perfect diagnostic accuracy and reduced diagnostic time. This finding underscores that AI should not be viewed as a replacement for expert judgment, but rather as a complementary decision-support tool that enhances performance even among highly experienced clinicians.

The observed improvement in interobserver agreement further highlights the potential of AI to standardize fracture interpretation and reduce diagnostic variability. Consistent interpretation is particularly important in multidisciplinary trauma care and in institutions where imaging is interpreted by clinicians with different backgrounds and levels of experience.

From a clinical perspective, the integration of AI-assisted DVT interpretation may contribute to improved patient safety by reducing the risk of missed fractures and facilitating earlier and more confident treatment decisions. Given the lower radiation dose and widespread availability of DVT in maxillofacial units, AI-supported DVT imaging may represent a viable alternative to conventional CT in selected trauma cases, particularly when polytrauma is not suspected.

This study has several limitations. Its retrospective single-center design may limit generalizability, and only osseous injuries were evaluated. Soft tissue injuries and complex polytrauma cases were not included. Additionally, while the reader study design provides valuable insights into diagnostic performance, prospective studies assessing real-time clinical integration and patient outcomes are warranted.

Future research should focus on prospective multicenter validation, integration of AI systems into routine PACS environments, and evaluation of their impact on clinical outcomes such as treatment delays, complication rates, and healthcare costs.

5. CONCLUSION

AI-assisted interpretation of DVT images significantly improves diagnostic accuracy, reduces time-to-diagnosis, and enhances interobserver agreement in maxillofacial trauma. These findings support the integration of AI systems into emergency MKG workflows as a clinically meaningful decision-support tool.

6. ETHICS STATEMENT

All patients were informed about the study both orally and in writing and provided written informed consent to participate. The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of the Hochschule Zurich, in Zurich, Switzerland.

7. CONFLICTS OF INTEREST

The authors have no financial conflicts of interest.

References

- [1] Ellis E 3rd, Moos KF, el-Attar A. Ten years of mandibular fractures: an analysis of cases. *J Oral Maxillofac Surg.* 1985;43:238–243.
 - [2] Haug RH, et al. Epidemiology of facial fractures. *J Oral Maxillofac Surg.* 1990;48:926–932.
 - [3] van den Bergh B, et al. Complications of delayed fracture treatment. *Int J Oral Maxillofac Surg.* 2012;41:123–129.
 - [4] Scarfe WC, Farman AG. Cone-beam computed tomography principles and clinical applications. *Dentomaxillofac Radiol.* 2008;37:4–25.
 - [5] Algerban A, et al. The role of cone-beam CT in trauma diagnostics. *Oral Radiol.* 2016;32:1–9.
 - [6] Mazzoni A, et al. Interobserver variability in facial fracture diagnosis. *Radiology.* 2014;273:534–541.
 - [7] Chung SW, et al. Artificial intelligence for fracture detection in medical imaging. *Radiology.* 2018;289:901–910.
 - [8] Lindsey R, et al. Deep learning in orthopedic imaging. *Clin Orthop Relat Res.* 2018;476:237–248.
 - [9] Liu X, et al. Artificial intelligence versus health-care professionals in medical imaging. *Lancet Digit Health.* 2019;1:e271–e297.
 - [10] Yildirim A, Hertach R, Yildirim V. Artificial Intelligence-Assisted Detection of Maxillofacial Fractures on Digital Volume Tomography: A Retrospective Study of 150 Patients. *J Med Dent (JMDNT).* 2025;1(2):44–46. doi:10.64951/jmdnt.2025.2.5.
-