



OPEN Explainable deep learning for age and gender estimation in dental CBCT scans using attention mechanisms and multi task learning

Najmeh Pishghadam¹, Rasool Esmaeilyfard^{1✉} & Maryam Paknahad²

Accurate and interpretable age estimation and gender classification are essential in forensic and clinical diagnostics, particularly when using high-dimensional medical imaging data such as Cone Beam Computed Tomography (CBCT). Traditional CBCT-based approaches often suffer from high computational costs and limited interpretability, reducing their applicability in forensic investigations. This study aims to develop a multi-task deep learning framework that enhances both accuracy and explainability in CBCT-based age estimation and gender classification using attention mechanisms. We propose a multi-task learning (MTL) model that simultaneously estimates age and classifies gender using panoramic slices extracted from CBCT scans. To improve interpretability, we integrate Convolutional Block Attention Module (CBAM) and Grad-CAM visualization, highlighting relevant craniofacial regions. The dataset includes 2,426 CBCT images from individuals aged 7 to 23 years, and performance is assessed using Mean Absolute Error (MAE) for age estimation and accuracy for gender classification. The proposed model achieves a MAE of 1.08 years for age estimation and 95.3% accuracy in gender classification, significantly outperforming conventional CBCT-based methods. CBAM enhances the model's ability to focus on clinically relevant anatomical features, while Grad-CAM provides visual explanations, improving interpretability. Additionally, using panoramic slices instead of full 3D CBCT volumes reduces computational costs without sacrificing accuracy. Our framework improves both accuracy and interpretability in forensic age estimation and gender classification from CBCT images. By incorporating explainable AI techniques, this model provides a computationally efficient and clinically interpretable tool for forensic and medical applications.

Keywords CBCT, Multi-task learning, Attention mechanisms, Age Estimation, Gender classification, Explainable AI (XAI), Forensic medicine

In recent years, the demand for precise and automated analysis of medical images has grown significantly, particularly in fields like forensic medicine and dentistry, where tasks such as age estimation and gender classification play crucial roles^{1–3}. Cone Beam Computed Tomography (CBCT), which provides detailed three-dimensional imaging of craniofacial structures, has become a valuable tool in these fields due to its ability to capture volumetric data that traditional 2D methods cannot^{4–6}. However, the high dimensionality and complexity of CBCT data introduce considerable computational challenges, and designing models that are both efficient and accurate, especially for tasks requiring nuanced interpretation of structural features like age and gender, remains a key obstacle^{7,8}.

Convolutional Neural Networks (CNNs) have transformed medical image analysis, particularly with 2D images, due to their capacity to learn relevant features automatically and outperform traditional manual methods in speed and accuracy^{9,10}.

However, applying CNNs to 3D CBCT images poses additional challenges, including significantly higher computational demands and increased risk of overfitting due to the greater number of parameters and limited

¹Computer Engineering and Information Technology Department, Shiraz University of Technology, Shiraz, Iran.

²Oral and Dental Disease Research Center, Oral and Maxillofacial Radiology Department, Dental School, Shiraz University of Medical Sciences, Shiraz, Iran. ✉email: esmaeily@sutech.ac.ir

size of annotated 3D datasets in medical imaging^{11,12}. These factors can reduce model generalizability, especially when training data lack diversity or are imbalanced. Although several studies have explored 3D CNNs and transformer-based architectures for volumetric image analysis^{13,14}, such models often require extensive computational resources and regularization techniques to avoid overfitting, which may not be practical in forensic or clinical workflows.

To address these challenges, multi-task learning (MTL) frameworks have emerged as a promising solution¹⁵. MTL allows a single model to handle multiple related tasks, such as age estimation and gender classification, by sharing feature extraction layers across tasks, reducing computational overhead and enhancing model generalization¹⁶. CBCT scans capture overlapping anatomical features—such as mandibular development and sinus morphology—relevant to both age and gender, making multi-task learning (MTL) particularly effective. In this study, age estimation is modeled as a regression task and gender classification as a binary task. MTL allows shared feature representations, improving performance and generalization while reducing overfitting, which is especially beneficial in data-limited forensic CBCT contexts¹⁷.

While most prior studies rely on 2D images lacking anatomical depth^{18,19}, our approach uses panoramic slices reconstructed from CBCT scans to balance spatial detail and computational efficiency^{4,12}. This method preserves key 3D features in a 2D format compatible with CNNs, enabling accurate age and gender estimation with reduced resource demands. CBCT offers valuable forensic cues—such as tooth eruption, pulp size, and mandibular morphology—supporting both tasks more effectively than traditional radiographs.

In forensic applications, accurate age estimation is critical for identifying individuals and determining the legal status of minors versus adults, while gender classification is essential for both medical diagnostics and forensic investigations^{2,9,20}. In CBCT imaging, several dental and craniofacial features provide biologically grounded cues for age and gender estimation. These include tooth development stages, pulp chamber volume, eruption patterns, and mandibular morphology. Such features exhibit clear age-related trends and sexual dimorphism, making them ideal for forensic analysis. Our model effectively captures these patterns through attention-guided learning and visual explanations²¹. Traditional methods are often time-consuming, subjective, and rely heavily on manual expertise²². By incorporating CNNs and multi-task learning, our framework offers a more efficient, objective, and scalable solution. Processing panoramic slices from CBCT scans also enables a streamlined workflow that retains the interpretability and accuracy of 3D data while reducing computational costs²³.

A major challenge in adopting machine learning models for medical and forensic applications is the need for interpretability, as clinicians and forensic experts must trust the model's decisions²⁴. Explainable AI (XAI) techniques, such as Grad-CAM (Gradient-weighted Class Activation Mapping), generate visual explanations by highlighting regions in an image that contribute most to the model's predictions²⁵. In contexts like forensic medicine, this transparency is essential to align the model's predictions with expert knowledge, ensuring that decisions are trustworthy and verifiable²⁶.

To further improve interpretability, we integrate attention mechanisms within our architecture. Attention enables the model to focus on the most relevant areas of the CBCT image, such as the mandible and maxilla, which are crucial for both age estimation and gender classification²⁷. This prioritization of key regions aligns the model's decision-making process with anatomical importance, making its predictions more reliable. Combining Attention with Grad-CAM yields precise visual explanations, enhancing performance and providing deeper insights into the model's reasoning for forensic experts and clinicians²⁸.

In this study, we propose a novel multi-task learning framework for age estimation and gender classification using CBCT images. By integrating attention mechanisms and Attention + Grad-CAM, our approach enhances both accuracy and transparency. Using panoramic slices reconstructed from CBCT scans, we retain the spatial richness of 3D data while reducing computational costs. Our results demonstrate that this method provides a robust solution for forensic applications, balancing high accuracy with the transparency required for clinical and forensic use.

Related work

Deep learning has driven significant progress in medical imaging, particularly in forensic applications like age estimation and gender classification^{29,30}. However, despite advancements, key research gaps remain, especially in the integration of 3D CBCT images and attention mechanisms for improving interpretability. This section critically reviews recent studies, highlighting both their contributions and limitations, while identifying areas for future exploration.

Age and gender estimation in forensic applications using deep learning

Estimating age and gender accurately is crucial in forensic medicine, where deep learning has the potential to improve both efficiency and precision. Studies on 2D imaging, such as the work by Vila-Blanco et al.¹⁹, have shown that deep learning models can achieve a Mean Absolute Error (MAE) of 0.97 years for age estimation and 91.82% accuracy in gender classification using panoramic radiographs. Similarly, Büyükcakır et al.¹⁸ leveraged EfficientNet-B4 on a dataset of 3,896 orthopantomograms (OPGs), achieving an impressive MAE of 0.562 years. These studies reveal the strength of deep learning on 2D images, particularly with large datasets, but raise questions about the generalizability of these models to smaller, imbalanced, or more diverse datasets often seen in forensic applications. Similarly, Park et al.³⁰ introduced a multi-task model based on EfficientNet-B3 and CBAM, achieving promising results (MAE = 2.93 years; accuracy = 99.2%), though their framework was limited to 2D images without CBCT integration. Similarly, Ozlu Ucan et al.³¹ developed a hybrid pipeline combining 2D and 1D CNNs with a Modified Genetic-Random Forest algorithm, reporting near-perfect performance ($R^2 = 0.999$), but focused exclusively on age estimation. Kokomoto et al.³² proposed a two-stage deep learning approach (MAE = 0.261 years), yet their model also operated solely on 2D radiographs and did not address

gender classification. In contrast, our study leverages CBCT-derived panoramic slices, enabling joint age and gender estimation through a multi-task learning framework. Additionally, we incorporate interpretability tools—Attention and Grad-CAM—to enhance model transparency, which is particularly critical in forensic applications.

Age estimation using 3D CBCT remains challenging due to volumetric complexity and computational demands. Pham et al.³³ reported a high MAE of 5.15 years, while Hou et al.³⁴ reduced this to 1.64 using NAS—but at a high computational cost. Gender classification has shown strong results with 2D CNNs, as demonstrated by Atas³⁵ and Rajee and Mythili¹⁶, though Bu et al.¹² noted performance drops in younger subjects. Venema et al.⁹ showed CNN adaptability using humerus images. Still, adapting 2D models to CBCT requires handling data imbalance and complexity. Talib et al.³⁶ further highlighted the importance of dataset integrity using transfer learning to filter artificial radiographs.

In summary, while deep learning models have shown strong performance on 2D data, adapting them to effectively utilize 3D CBCT and address challenges like generalizability and data imbalance remains underexplored. Future research should prioritize scalable CBCT-based solutions for real-world forensic applications. A comparative overview of existing methods and our proposed approach is provided in Table 1, highlighting their respective strengths, limitations, and contributions.

Integration of CBCT images in multi-task learning

While CBCT images provide detailed anatomical information, their use in multi-task learning (MTL) frameworks—where both age estimation and gender classification can be performed simultaneously—remains limited. Fujimoto et al.⁷ demonstrated the potential of using 3D CBCT images for age estimation by focusing on alveolar bone features, achieving promising results in clinical settings. However, this study did not explore the application of multi-task learning, where shared features could be used to improve both age and gender predictions.

Vila-Blanco et al.¹⁹ addressed some of the challenges in processing CBCT images by developing a multi-task deep learning framework for automatic tooth and root canal segmentation. Their model, combining DentalNet and PulpNet, efficiently segmented complex anatomical structures, significantly reducing manual segmentation time and achieving superior performance on clinical datasets. This study demonstrates the efficacy of shared feature learning in dental imaging tasks but does not extend its findings to multi-task models that could handle both age estimation and gender classification.

da Silva et al.²⁹ introduced SDetNet, a CBCT-based model with anatomy-guided attention for sex classification via frontal sinus analysis. While effective, it did not include age estimation or multi-task learning. In contrast, our model addresses both tasks simultaneously. Although multi-task learning with 3D CBCT data holds promise, it requires balancing accuracy with computational efficiency.

Attention mechanisms and explainability

Attention mechanisms have become increasingly important in improving the interpretability of deep learning models, particularly in medical imaging. Joo et al.³⁷ developed a VAE-based model for age estimation, incorporating attention maps that highlighted the most relevant regions of teeth for age prediction. This approach improved the model’s explainability, which is crucial in forensic applications where the decision-making process must be transparent.

In a different approach, Hou et al.³⁸ introduced a Multi-scale Aggregation Attention Block (MAB) within their Teeth U-Net model, which was designed to enhance teeth boundary detection in panoramic X-ray images. The model achieved a segmentation accuracy of 98.53%, illustrating the power of attention mechanisms in boosting model performance and improving interpretability.

Despite these advancements, attention mechanisms in 3D CBCT data remain underexplored. Applying attention layers to 3D data could not only improve model transparency but also enhance performance in tasks

Study	Data type	Model architecture	Accuracy/Error (MAE)	Advantages	Disadvantages
Vila-Blanco et al. [19]	Panoramic Images (2D)	Basic CNN	MAE = 0.97 years	High precision in dental predictions	Needs large datasets
Büyükcakır et al. [18]	OPG Images (2D)	EfficientNet-B4	MAE = 0.562 years	High accuracy on large datasets	Poor generalizability
Pham et al. [34]	CBCT Images (3D)	Basic CNN	MAE = 5.15 years	Rich structural info from 3D	High error; needs further optimization
Hou et al. [35]	OPG Images (2D)	NAS-optimized model	MAE = 1.64 years	Accurate via NAS	High computational cost; requires strong resources
Atas [36]	Panoramic Images (2D)	DenseNet121	Accuracy = 97.25%	Strong gender prediction	Limited to 2D data; requires more data
Rajee and Mythili [16]	Digital Dental X-rays (2D)	ResNet50	Accuracy = 98.27%	Strong performance on dental images	Reduced performance across age groups
Bu et al. [12]	OPG Images (2D)	Basic CNN	Adults: 90.97%, Minors: 82.64%	Age-specific insights	Drop in accuracy for minors
Venema et al. [9]	Humerus Images	Basic CNN	Accuracy = 91.03%	Adaptable to non-dental data	Not usable for CBCT
Proposed Method	Selected CBCT Slices	ResNet50, VGG16, DenseNet121, InceptionV3	MAE = 1.08 years (age), 95.3% (gender)	Multi-task, interpretable, efficient	Higher radiation dose

Table 1. Comparative analysis of related work and proposed method.

that require detailed anatomical analysis, such as age estimation and gender classification. Future work should focus on developing attention mechanisms tailored to 3D data, addressing the unique challenges posed by the spatial complexity of CBCT images.

Materials and methods

This retrospective study was approved by the institutional ethics committee (IR.SUMS.REC.1402.215). All CBCT scans were originally acquired for diagnostic purposes unrelated to this research. Informed consent was obtained from all patients, including permission for the anonymized use of imaging data in future research. No additional imaging was performed specifically for this study, and all methods complied with institutional and international guidelines.

This study employs an advanced deep learning framework for age estimation and gender classification using CBCT imaging, leveraging multi-task learning and attention mechanisms. The methodology follows a structured workflow (as depicted in Fig. 1) that includes data preprocessing, model development, optimization, and evaluation. A complete breakdown of the methodology is provided in Appendix A1.

Data acquisition and preprocessing

CBCT scans were retrospectively collected and preprocessed to ensure standardization and enhance interpretability. The preprocessing pipeline involved panoramic slice reconstruction, region-of-interest (ROI) extraction, and contrast enhancement to highlight critical anatomical features. Advanced data augmentation strategies, including geometric transformations (rotation, scaling), noise injection, and histogram equalization, were applied to improve model generalizability.

Model architecture and training

A multi-task deep learning framework was implemented, integrating state-of-the-art architectures (ResNet50, DenseNet121, InceptionV3) optimized for forensic radiology. Feature extraction was enhanced using Convolutional Block Attention Module (CBAM) to prioritize clinically relevant structures, while Grad-CAM visualization ensured model interpretability. The training process was optimized using Adam optimizer, learning rate decay, early stopping, and cross-validation across multiple CBCT datasets. The loss functions included Mean Squared Error (MSE) for age estimation and Binary Cross-Entropy (BCE) for gender classification.

Evaluation and explainability

Model performance was rigorously evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), accuracy, precision, recall, and AUC. Advanced explainability techniques, including Grad-CAM, and attention heatmaps, were used to validate the model's decision-making process by highlighting key craniofacial features. Additionally, clinical validation was conducted by forensic experts to assess real-world applicability and reliability.

For a comprehensive breakdown of data preprocessing techniques, hyperparameter tuning, model architecture variations, and evaluation metrics, refer to Appendix A1, where detailed methodological explanations are provided.

Results

The proposed multi-task deep learning model demonstrated high accuracy and robustness in age estimation and gender classification tasks using CBCT images. The model's performance was evaluated across various architectures, data augmentation strategies, and attention mechanisms, highlighting the impact of interpretability techniques on forensic radiology.

Key findings

- **Age estimation performance** The model achieved strong predictive accuracy, as shown in Table 2, with significant improvements when incorporating attention mechanisms. A detailed comparison of actual and predicted age values across the dataset is illustrated in Fig. 2.

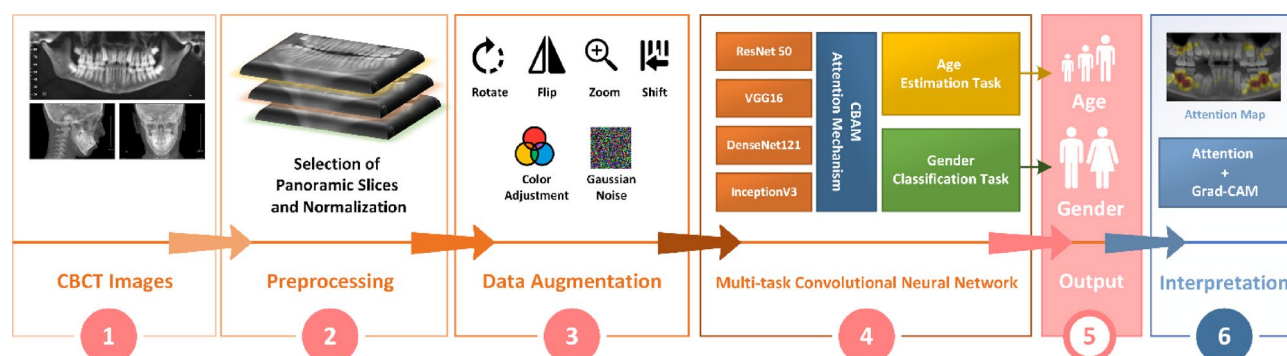


Fig. 1. Overview of the proposed method for age estimation and gender classification from CBCT images.

Architecture	MAE (years)	RMSE (years)	R^2 Score
ResNet50	1.56	2.01	0.89
VGG16	1.73	2.2	0.87
DenseNet121	1.64	2.1	0.88
ViT	1.12	1.66	0.91
InceptionV3	1.08	1.44	0.93

Table 2. Performance of the model for age Estimation across architectures.

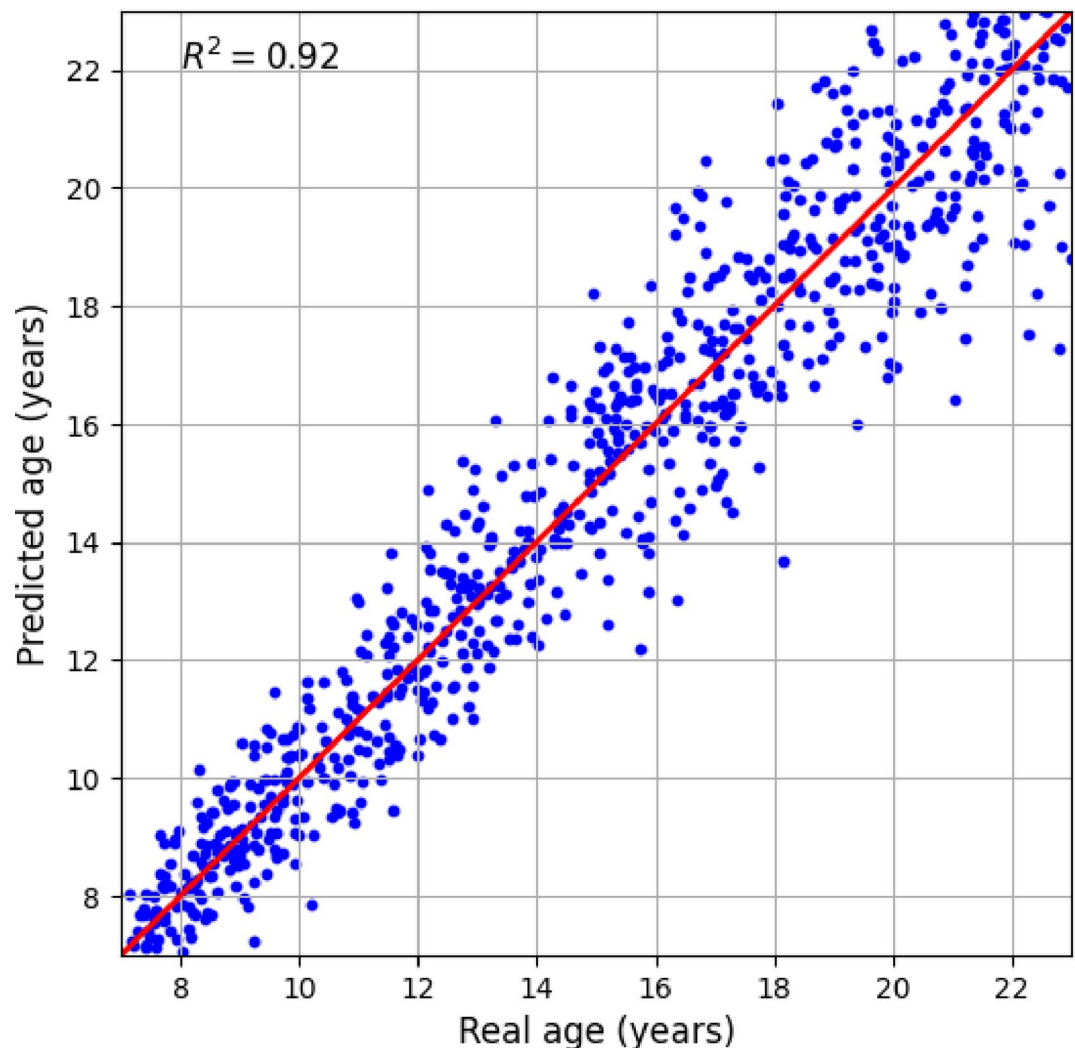


Fig. 2. Comparison of actual and predicted age across the entire dataset.

- *Performance Across Age Subgroups and Gender:* The model's accuracy varied based on age groups and gender, with performance trends depicted in Figs. 3 and 4.
- *Impact of attention mechanisms* Tables 3 and 4 present a comparative analysis of InceptionV3 with and without attention mechanisms, demonstrating the enhanced feature extraction capabilities of the model when attention mechanisms were included.
- *Model Interpretability* The integration of Grad-CAM and CBAM attention mechanisms significantly enhanced visual interpretability, as shown in Figs. 5 and 6, highlighting key craniofacial and dental regions.

For a detailed breakdown of experimental results, statistical comparisons, and additional performance metrics, please refer to Appendix A2, where comprehensive evaluations of different architectures, training configurations, and hyperparameter optimizations are provided.

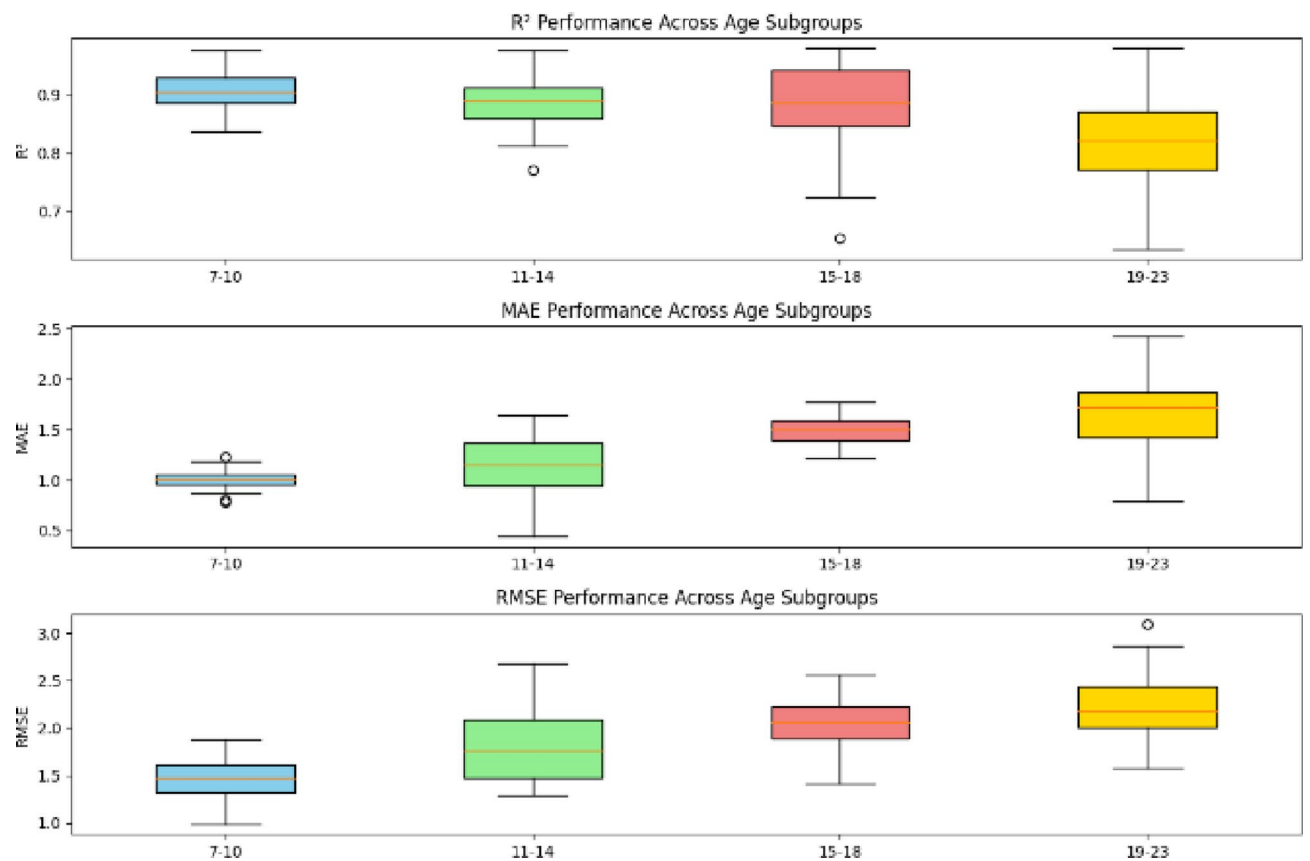


Fig. 3. Model performance analysis across different age subgroups.

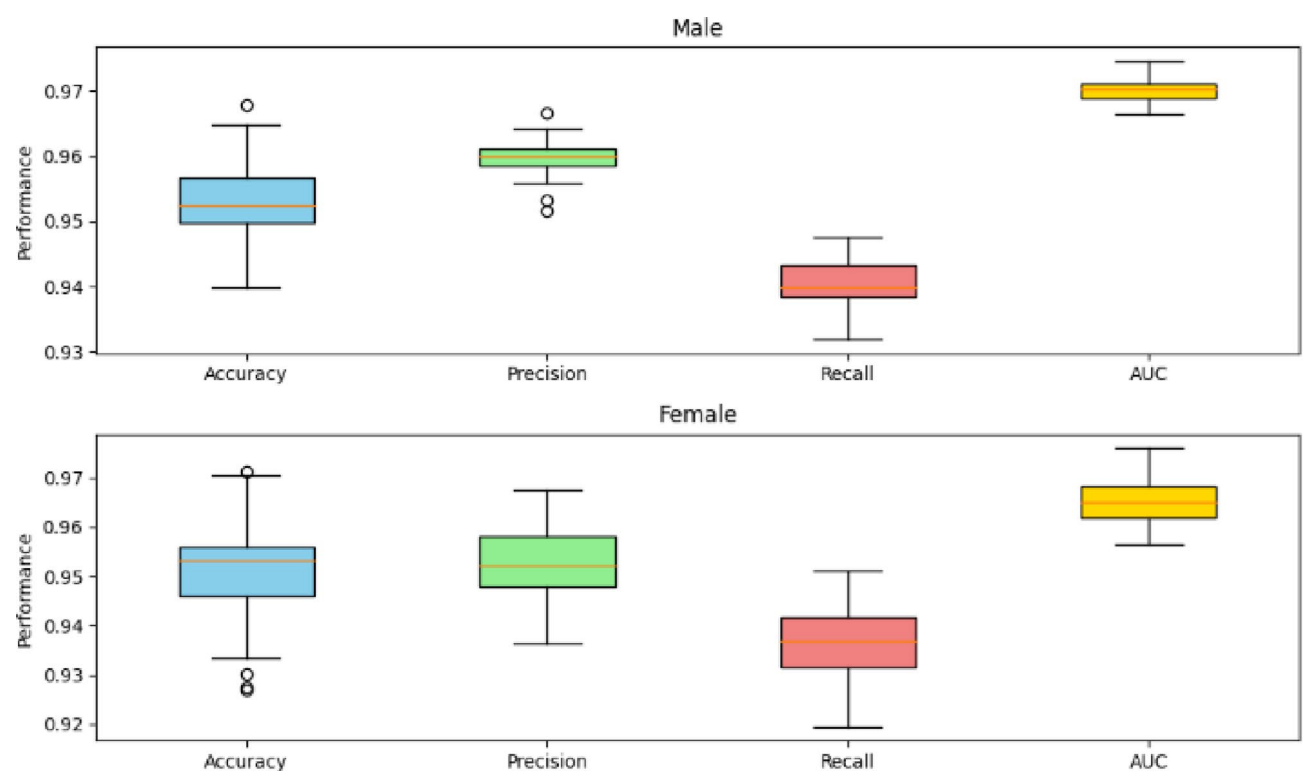


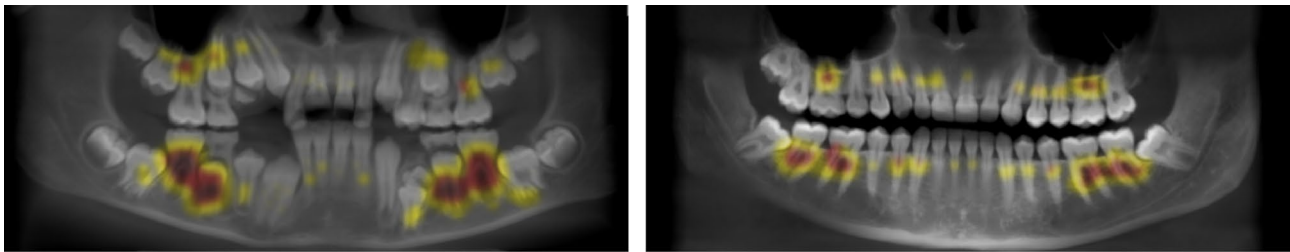
Fig. 4. Performance across gender.

Dataset	Age Regression			Sex Classification			
	E	AE	R ²	Accuracy	Precision	Recall	AUC
All Data	1.5	μ±δ: 1.80, Med: 1.60	0.85	90.0	0.91	0.89	0.92
Age 7–10	1.4	μ±δ: 1.70, Med: 1.50	0.88	91.0	0.92	0.90	0.93
Age 11–14	1.5	μ±δ: 1.90, Med: 1.60	0.87	90.5	0.91	0.88	0.92
Age 15–18	1.6	μ±δ: 2.00, Med: 1.70	0.85	90.0	0.90	0.87	0.91
Age 19–23	1.7	μ±δ: 2.10, Med: 1.80	0.82	89.5	0.89	0.86	0.90

Table 3. Results of InceptionV3 without attention mechanism.

Dataset	Age regression			Sex classification			
	E	AE	R ²	Accuracy	Precision	Recall	AUC
All Data	1.08	μ±δ: 1.44, Med: 1.20	0.93	95.3	0.96	0.94	0.97
Age 7–10	0.9	μ±δ: 1.20, Med: 1.00	0.95	95.5	0.97	0.96	0.98
Age 11–14	1	μ±δ: 1.30, Med: 1.10	0.94	95.3	0.96	0.94	0.97
Age 15–18	1.1	μ±δ: 1.40, Med: 1.15	0.93	95.1	0.95	0.93	0.97
Age 19–23	1.2	μ±δ: 1.50, Med: 1.25	0.92	94.8	0.94	0.92	0.96

Table 4. Results of InceptionV3 with attention mechanism.



(a) A boy of 9 years and 2 months (b) A girl of 17 years and 11 months

Fig. 5. Attention + Grad-CAM Maps highlighting key dental regions for age and gender estimation.

Discussion

This study advances previous research by implementing a multi-task learning framework on CBCT-derived panoramic reconstructions instead of conventional 2D radiographs. It introduces a hybrid loss-balancing strategy tailored for imbalanced forensic datasets and integrates CBAM with Grad-CAM to enhance model interpretability. These innovations contribute both practical value and methodological novelty to deep learning in forensic radiology. The proposed model demonstrated robust performance, achieving a low MAE of 1.08 years and an R² of 0.93 in age estimation, particularly excelling in younger age groups. In gender classification, it achieved an accuracy of 95.3% and an AUC of 0.97. These results indicate that the model effectively leverages CBCT data, especially with the inclusion of multi-scale CNN features and attention mechanisms that focus on clinically relevant regions, enhancing interpretability and generalization.

The attention mechanism, combined with Grad-CAM, significantly enhances the interpretability of the model by providing visual insights into the model’s focus areas, such as the mandible and maxillary regions. The model’s capacity to produce interpretable heatmaps allows practitioners to see which anatomical regions are prioritized, making the model more reliable in clinical and forensic settings.

Our model demonstrates strong performance when compared to recent studies in age and gender prediction. Vila-Blanco et al.¹⁹ and Büyükcakır et al.¹⁸ reported slightly lower MAEs using 2D panoramic radiographs; however, their approaches lack the spatial depth and contextual detail provided by CBCT imaging. Our method leverages CBCT-derived panoramic slices, which preserve key anatomical features while reducing computational complexity. In contrast to Pham et al.³³, who used full 3D CBCT volumes and observed a high error rate (MAE = 5.15), our slice-based strategy enables more accurate predictions (MAE = 1.08). Furthermore, while Joo et al.³⁷ and Park et al.²⁹ incorporated attention mechanisms, they lacked external interpretability techniques like Grad-CAM, which our model integrates to enhance clinical transparency. Finally, unlike Ozlu Ucan et al.³⁰, whose hybrid method was limited to age estimation on 2D data, our model supports both age and gender classification in a unified, interpretable multi-task framework.

While 3D CBCT scans provide rich anatomical detail, processing full volumes is computationally demanding and prone to overfitting with limited data. By using reconstructed 2D panoramic slices, our method retains

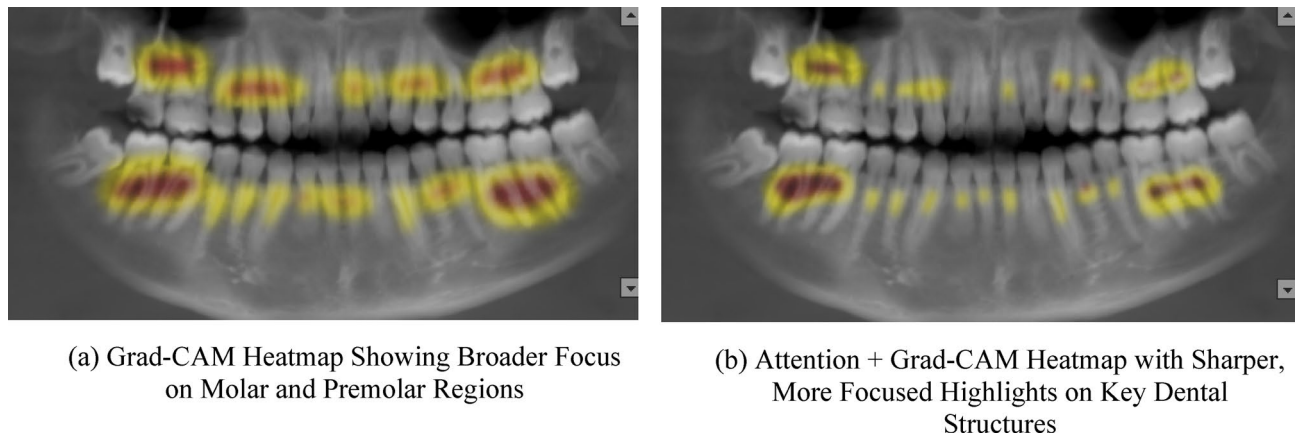


Fig. 6. Comparison of Grad-CAM and Attention + Grad-CAM heatmaps for dental X-ray interpretation.

essential spatial features while enabling efficient, interpretable modeling. Compared to 3D approaches like Pham et al.³³, which showed high error rates, our 2D strategy offers a practical balance between accuracy and feasibility for forensic use.

The integration of attention mechanisms, particularly Attention + Grad-CAM, has proven effective in enhancing both performance and interpretability. By guiding the model to focus on specific craniofacial structures, attention mechanisms contribute to more accurate age and gender predictions. This focus is particularly beneficial for complex anatomical variations in older age groups, where the model's ability to prioritize key features like the mandible and molars becomes essential.

As presented in Tables 3 and 4 and visualized in Figs. 3 and 4, the model exhibited highest accuracy for age estimation in the 7–10 and 15–18 age groups, with slightly higher errors in older individuals (19–23), likely due to greater anatomical variability. Gender classification was consistently accurate across sexes, though male subjects showed marginally higher precision and recall. These findings highlight the model's robustness and generalizability across age and gender subgroups, reinforcing its applicability in diverse forensic contexts. Attention mechanisms also provide visual explanations of the model's focus areas, allowing practitioners to better understand its decision-making process. This transparency is essential in clinical and forensic applications, where interpretability builds trust and facilitates model validation. By visually highlighting relevant features, the model aids practitioners in understanding the basis for predictions, making it a valuable tool for medical and legal assessments.

The findings have practical implications for both clinical and forensic applications:

InceptionV3 performs better due to its multi-scale architecture, which captures both fine and broad features simultaneously. This is especially useful in CBCT images where anatomical cues vary in size and location. Its parallel convolutional paths give it an advantage over linear models like ResNet50 and VGG16 in handling such structural complexity. *Forensic Applications:* The model's accuracy in age estimation and gender classification could aid forensic experts in identification processes, expediting age and gender determination in legal cases. The integration of attention mechanisms further enhances model trustworthiness, allowing experts to visually validate the model's focus areas and ensuring alignment with established anatomical markers.

Clinical applications: In clinical settings, the model can help automate diagnostic workflows, especially in scenarios where dental and craniofacial analysis is required for treatment planning or patient evaluation. This is particularly beneficial in pediatric cases for growth assessment and orthodontic planning, as well as in reconstructive surgery. By providing accurate, interpretable outputs, the model can reduce diagnostic errors, speed up analysis, and ultimately contribute to improved patient outcomes.

Although the dataset used in this study is relatively large compared to previous CBCT-based research, it remains modest in size when considered against typical deep learning requirements. It lacks sufficient representation of older age groups beyond 23 years. This limited age range, may restrict the model's ability to generalize across broader forensic populations with greater anatomical variability. Furthermore, all CBCT scans were obtained from a single center, using a specific scanner and imaging protocol. While this standardization minimizes internal variability, it introduces potential institutional bias, meaning that the model's performance may decline when applied to images from different centers, scanners, ethnic groups, or clinical contexts. Finally, the lack of external validation on independent datasets remains a major limitation; despite robust internal cross-validation, independent testing is essential to truly confirm the model's stability, generalizability, and suitability for real-world forensic or clinical deployment. Future research should focus on expanding the dataset to include a broader range of age groups and diverse ethnic backgrounds, improving the model's generalizability. Future work should involve collaboration with forensic institutions to collect case-specific CBCT data and evaluate the model's robustness across broader, more varied populations. Additional improvements in attention mechanisms and interpretability techniques, such as layer-wise relevance propagation (LRP) or integrated gradients, could enhance the model's focus on clinically significant areas.

Moreover, adapting the model for additional clinical and forensic applications, including the detection of dental anomalies, disease classification, or facial reconstruction, could expand its utility. Future models could

also consider incorporating non-binary or diverse gender categories, making the system more inclusive and suitable for a wider range of clinical scenarios.

Although Grad-CAM visualizations were primarily performed on the InceptionV3 architecture due to its superior overall performance, future work could explore comparative interpretability across multiple architectures to better understand how different models attend to forensic-relevant anatomical regions.

Conclusion

This study introduced an explainable multi-task deep learning model for simultaneous age estimation and gender classification using CBCT-derived panoramic slices. By leveraging the InceptionV3 architecture with integrated attention mechanisms and Grad-CAM visualizations, the model achieved strong predictive performance (MAE = 1.08 years for age; 95.3% accuracy for gender) along with high interpretability. Compared to traditional 2D radiograph-based models, our approach benefits from CBCT's anatomical richness while maintaining computational efficiency through 2D reconstruction. These findings highlight the model's potential in forensic and clinical contexts where both accuracy and transparency are essential. Future work should focus on expanding the dataset, validating across external populations, and exploring transformer-based architectures for enhanced feature learning.

Data availability

The relevant processed data are available from the corresponding author upon reasonable request for research purposes.

Received: 17 March 2025; Accepted: 20 May 2025

Published online: 24 May 2025

References

- Galea-Holho, L. B., Delcea, C., Siserman, C. V. & Ciocan, V. Age estimation of human remains using the dental system: A review. *Ann. Dent. Specialty* **11**(3), 14–18 (2023).
- Zainuddin, M. Z., Mohamad, N. S., Su Keng, T. & Mohd Yusof, M. Y. P. The applications of MicroCT in studying age-related tooth morphological change and dental age estimation: A scoping review. *J. Forensic Sci.* **68**(6), 2048–2056 (2023).
- Mohamed, E. G., Redondo, R. P. D., Koura, A., EL-Mofly, M. S. & Kayed, M. Dental age estimation using deep learning: A comparative survey. *Computation* **11**(2), 18 (2023).
- Lang, Y. et al. Localization of craniomaxillofacial landmarks on CBCT images using 3D mask R-CNN and local dependency learning. *IEEE Trans. Med. Imaging* **41**(10), 2856–2866 (2022).
- Merdietio Boedi, R., Shepherd, S., Oscandar, F., Mânica, S. & Franco, A. 3D segmentation of dental crown for volumetric age Estimation with CBCT imaging. *Int. J. Legal Med.* **137**(1), 123–130 (2023).
- Mackiewicz, E., Bonsmann, T., Kaczor-Wiankowska, K. & Nowicka, A. Volumetric assessment of apical periodontitis using Cone-Beam computed Tomography—A systematic review. *Int. J. Environ. Res. Public Health* **20**(4), 2940 (2023).
- Fujimoto, H., Kimura-Kataoka, K., Takeuchi, A., Yoshimiya, M. & Kawakami, R. Evaluation of age estimation using alveolar bone images. *Forensic Sci. International* **112237** (2024).
- Zheng, Z. et al. Anatomically constrained deep learning for automating dental CBCT segmentation and lesion detection. *IEEE Trans. Autom. Sci. Eng.* **18**(2), 603–614 (2020).
- Venema, J., Peula, D., Irurita, J. & Mesejo, P. Employing deep learning for sex Estimation of adult individuals using 2D images of the humerus. *Neural Comput. Applications*, **35**, 8, pp. 5987–5998, 2023/03/01 2023.
- Milošević, D., Vodačević, M., Galić, I. & Subašić, M. Automated estimation of chronological age from panoramic dental X-ray images using deep learning. *Expert Syst. Appl.* **189**, 116038 (2022).
- Wang, Y. et al. Root canal treatment planning by automatic tooth and root canal segmentation in dental CBCT with deep multi-task feature learning. *Med. Image Anal.* **85**, 102750 (2023).
- Bu, W. et al. Automatic sex estimation using deep convolutional neural network based on orthopantomogram images. *Forensic Sci. Int.* **348**, 111704 (2023).
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* 424–432 (Springer International Publishing, 2016).
- Hatamizadeh, A. et al. UNETR: Transformers for 3D medical image segmentation. in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* 1748–1758 (2022).
- Zhao, Y., Wang, X., Che, T., Bao, G. & Li, S. Multi-task deep learning for medical image computing and analysis: A review. *Comput. Biol. Med.* **153**, 106496 (2023).
- Rajee, M. V. & Mythili, C. Gender classification on digital dental x-ray images using deep convolutional neural network. *Biomed. Signal Process. Control* **69**, 102939 (2021).
- Sukegawa, S. et al. Evaluation of multi-task learning in deep learning-based positioning classification of mandibular third molars. *Sci. Rep.* **12**(1), 684 (2022).
- Büyükcakır, B. et al. OPG-based dental age Estimation using a data-technical exploration of deep learning techniques. *J. Forensic Sci.* **69**(3), 919–931 (2024).
- Vila-Blanco, N., Varas-Quintana, P., Aneiros-Ardao, Á., Tomás, I. & Carreira, M. J. XAS: Automatic yet eXplainable Age and Sex determination by combining imprecise per-tooth predictions. *Comput. Biol. Med.* **149**, 106072 (2022).
- Mualla, N., Houssein, E. H. & Hassan, M. Dental age Estimation based on X-ray images. *Comput. Mater. Continua* **62**, 2 (2020).
- Esmailyfard, R., Paknahad, M. & Dokohaki, S. Sex classification of first molar teeth in cone beam computed tomography images using data mining. *Forensic Sci. Int.* **318**, 110633 (2021).
- Santosh, K. & Pradeep, N. Machine learning techniques for gender recognition and age prediction using digital images of human dentition. *Int. Res. J. Adv. Eng. Manage. (IRJAEM)* **2**(04), 609–618 (2024).
- Zhang, J. et al. A fast automatic reconstruction method for panoramic images based on cone beam computed tomography. *Electronics* **11**(15), 2404 (2022).
- Solanke, A. A. Explainable digital forensics AI: towards mitigating distrust in AI-based digital forensics analysis using interpretable models. *Forensic Sci. International: Digit. Invest.* **42**, 301403 (2022).
- Ihongbe, I. E., Fouad, S., Mahmoud, T. F., Rajasekaran, A. & Bhatia, B. Evaluating explainable artificial intelligence (XAI) techniques in chest radiology imaging through a human-centered Lens. *PloS One* **19**(10), e0308758 (2024).
- Lefèvre, T. & Tournois, L. Artificial intelligence and diagnostics in medicine and forensic science. *Diagnostics* **13**(23), 3554 (2023).

27. Xiao, H. et al. Real-Time 4-D-Cone beam CT accurate estimation based on single-angle projection via dual-attention mechanism residual network. *IEEE Trans. Radiation Plasma Med. Sci.* **7**(6), 618–629 (2023).
28. Raghavan, K. Attention guided grad-CAM: an improved explainable artificial intelligence model for infrared breast cancer detection. *Multimed. Tools Appl.* **83**(19), 57551–57578 (2024).
29. da Silva, R. L. B. et al. Automatic segmentation and classification of frontal sinuses for sex determination from CBCT scans using a two-stage anatomy-guided attention network. *Sci. Rep.* **14**(1), 11750 (2024).
30. Park, S. J. et al. Automatic and robust Estimation of sex and chronological age from panoramic radiographs using a multi-task deep learning network: A study on a South Korean population. *Int. J. Legal Med.* **138**(4), 1741–1757 (2024).
31. Ozlu Ucan, G., Gwass, O. A. H., Apaydin, B. K. & Ucan, B. Automated age estimation from OPG images and patient records using deep feature extraction and modified genetic-random forest. *Diagnostics* **15**(3), 314 (2025).
32. Kokomoto, K. et al. Automatic dental age calculation from panoramic radiographs using deep learning: A two-stage approach with object detection and image classification. *BMC Oral Health* **24**(1), 143 (2024).
33. Pham, C. V. et al. Age Estimation based on 3D post-mortem computed tomography images of mandible and femur using convolutional neural networks. *PLoS One.* **16**(5), e0251388 (2021).
34. Hou, W. et al. Exploring effective DNN models for forensic age estimation based on panoramic radiograph images. in *International Joint Conference on Neural Networks (IJCNN)* 1–8 (2021).
35. Atas, I. Human gender prediction based on deep transfer learning from panoramic radiograph images, *arXiv preprint arXiv:2205.09850* (2022).
36. Talib, M. A. et al. Transfer learning-based classifier to automate the extraction of false X-Ray images from hospital's database. *Int. Dent. J.* **74**(6), 1471–1482 (2024).
37. Joo, S., Jung, W. & Oh, S. E. Variational autoencoder-based Estimation of chronological age and changes in morphological features of teeth. *Sci. Rep.* **13**(1), 704 (2023).
38. Hou, S. et al. Teeth U-Net: A segmentation model of dental panoramic X-ray images for context semantics and contrast enhancement, *Comput. Biol. Med.* **152**, 106296 (2023).

Author contributions

N. P. was responsible for software development, data collection, evaluation, visualization, and drafting the initial manuscript. R. E. contributed to supervision, project management, initial manuscript writing and revision, evaluation, visualization, and conceptualization. M. P. contributed to conceptualization, data collection, and final manuscript editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-03305-z>.

Correspondence and requests for materials should be addressed to R.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025