



OPEN ACCESS

EDITED BY

Jinxi Wang,
University of Kansas Medical Center,
United States

REVIEWED BY

Jan Christian Danz,
University of Bern, Switzerland
Sabrina Kathrin Schulze,
University of Potsdam, Germany

*CORRESPONDENCE

Sultan Uzun,
✉ dtsultanuzun@gmail.com

RECEIVED 09 August 2025

REVISED 14 October 2025

ACCEPTED 17 October 2025

PUBLISHED 12 November 2025

CITATION

Uzun S, Magat G and Evli C (2025) Detection of spheno-occipital synchondrosis fusion stages using artificial intelligence. *Front. Physiol.* 16:1682917. doi: 10.3389/fphys.2025.1682917

COPYRIGHT

© 2025 Uzun, Magat and Evli. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Detection of spheno-occipital synchondrosis fusion stages using artificial intelligence

Sultan Uzun^{1*}, Guldane Magat¹ and Cengiz Evli²

¹Department of Dentomaxillofacial Radiology, Necmettin Erbakan University Faculty of Dentistry, Konya, Türkiye, ²Department of Dentomaxillofacial Radiology, Ankara University Faculty of Dentistry, Ankara, Türkiye

Introduction: Accurate evaluation of the spheno-occipital synchondrosis (SOS) is important for growth assessment, early detection of craniofacial anomalies, and reliable forensic age estimation.

Methods: This study applied three deep learning models—YOLOv5, YOLOv8, and YOLOv11—to cone-beam computed tomography (CBCT) sagittal images from 1,661 individuals aged 6–25 years, aiming to automate SOS fusion stage classification. Model performance was compared in terms of detection accuracy and processing speed.

Results: All models achieved high accuracy, with a mean average precision (mAP) of 0.995 in complete fusion (Stage 3). YOLOv8 demonstrated the most consistent balance of precision and recall, while YOLOv11 achieved the fastest inference time (27 ms). YOLOv5 excelled in specific stages with perfect F1-scores.

Discussion: These findings demonstrate that YOLO-based AI models can provide precise, rapid, and reproducible SOS assessments, offering valuable support for both clinical decision-making and forensic investigations.

KEYWORDS

growth and development, craniofacial anomaly, spheno-occipital synchondrosis, YOLO, deep learning, artificial intelligence

1 Introduction

Craniofacial growth is a complex biological process influenced by genetic, environmental, and functional factors. The three synchondroses of the cranial base—spheno-ethmoidal, inter-sphenoid, and spheno-occipital synchondroses (SOS) are significant growth centers, with the SOS being the last to close (Krishan and Kanchan, 2013; Singh et al., 2025). The spheno-occipital synchondrosis (SOS), a cartilaginous joint between the sphenoid and occipital bones, plays a crucial role in this process. As a key growth center, SOS facilitates the extension of the cranial base axis, enabling the development of teeth and alveoli and thus contributing to craniofacial formation (Dalili Kajan et al., 2021; Funato et al., 2020).

The cranial base represents the first stable structure in craniofacial growth, rather than the occipital bone. The maxilla usually grows downward and forward as described by classic implant studies (Björk, 1955; Melsen, 1974; Solow, 1980), and its positional changes are largely influenced by remodeling and displacement relative to the cranial base rather than direct growth at the spheno-occipital synchondrosis (SOS). Growth at the SOS primarily contributes to the elongation of the posterior cranial base and the flexure angle between the anterior and posterior cranial base, which in turn may affect the

spatial relationship of the mandible and temporal bone relative to the cranial base (Halpern, 2014; Demirturk Kocasarac et al., 2016; Sinanoglu et al., 2016; Cendekiawan et al., 2010; Nie, 2005).

In skeletal Class III malocclusion, early fusion or reduced growth potential at the SOS has been linked to a shorter posterior cranial base and a more acute cranial base angle, leading to a retruded maxilla and a forward-positioned mandible (Singh et al., 2025; Halpern, 2014; Cendekiawan et al., 2010). Clinical and imaging studies have consistently demonstrated distinctive cranial base features in Class III patients, such as shorter posterior cranial base length during the prepubertal period and altered mandibular positioning (Yang et al., 2016; Singh et al., 2025). However, Yang et al. reported that the timing and fusion patterns of the SOS were not significantly different between Class I and Class III groups, suggesting that while SOS maturation contributes to cranial base morphology, it is unlikely to act as the sole determinant of sagittal jaw discrepancies.

Thus, the importance of SOS fusion and growth direction in Class III malocclusion lies in their potential to influence cranial base angulation and mandibular displacement, rather than directly altering maxillary growth direction. This perspective highlights the cranial base as a key mediator of skeletal pattern, whereas the maxilla continues its downward and forward growth trajectory relative to this stable base.

Fusion of the SOS occurs later than other midline synchondroses. While the speno-ethmoidal synchondrosis (SES) typically fuses by age six and the intersphenoid synchondrosis (ISS) at birth, SOS generally fuses between ages 12 and 15 (Evli et al., 2025). However, its exact timing is debated, with reports ranging from 8 to 21 years (Nie, 2005; Hoshino et al., 2022). Because SOS is the last cranial base synchondrosis to close, it provides a unique window for growth assessment. Its maturation stages have been linked to skeletal maturity and can offer valuable insight for treatment timing, particularly in orthodontic and orthopedic interventions such as headgear therapy and rapid maxillary expansion (Halpern, 2014). If performed before SOS fusion, these techniques can temporarily open or influence the synchondrosis, facilitating maxillary displacement and alveolar adaptation (Halpern, 2014).

In craniofacial syndromes, midface hypoplasia is believed to stem from early or atypical SOS ossification and disrupted cranial base growth (Sinanoglu et al., 2016). Aberrant fusion may indicate underlying defects; in syndromes like Apert, Crouzon, and Muenke, premature SOS closure correlates with midfacial underdevelopment (Goldstein et al., 2014; McGrath et al., 2012; Tahiri et al., 2014).

Advances in artificial intelligence (AI) have enabled the use of machine learning in craniofacial research to predict growth patterns. AI aids in diagnosing craniofacial anomalies and evaluating interventions like rapid maxillary expansion (Bazargani et al., 2013; Geisler et al., 2021). Leveraging large datasets and imaging, AI improves diagnostic accuracy and treatment planning in orthodontics and craniofacial surgery.

Given the clinical importance of SOS fusion, accurately identifying its stages is key for tracking development and detecting anomalies early. Cone-beam computed tomography (CBCT) offers high-resolution imaging to assess SOS fusion precisely. This study aims to apply AI algorithms to classify SOS fusion on CBCT images (CBCTIs) and examine its correlation with growth periods. We

hypothesize that AI-based evaluation will support early anomaly detection, leading to better diagnosis and treatment planning.

2 Materials and methods

2.1 Ethics approval and sample size determination

This thesis study received ethical approval from the Local Non-Drug and Non-Medical Device Research Ethics Committee on 25 January 2024, with decision number 2024/363. All procedures adhered to the principles of the Declaration of Helsinki.

Based on a one-sided independent sample t-test with a 95% confidence level, 95% test power, and an effect size of $d = 0.518$, it was determined that a minimum of 85 participants per group was needed (Geng et al., 2024).

2.2 Image collection and inclusion criteria

This retrospective study evaluated CBCT images (CBCTIs) acquired between 2020 and 2024 from the Departments of Dentomaxillofacial Radiology at Necmettin Erbakan University Faculty of Dentistry and Ankara University Faculty of Dentistry. Included were sagittal section CBCTIs from individual aged 6–25 that clearly showed the speno-occipital synchondrosis (SOS) with high diagnostic quality. Images were excluded if they:

- Were from individuals over age 25,
- Showed congenital/acquired maxillofacial deformities,
- Had large pathological lesions or trauma history,
- Showed evidence of head and neck surgery, radiotherapy, or chemotherapy,
- Came from syndromic cases impacting the craniofacial region,
- Or had insufficient resolution, artifacts, or incomplete SOS depiction.

To ensure image quality and consistency, all CBCTIs were standardized. Technical settings and imaging protocols were selected to minimize variables that could affect SOS visibility.

2.3 Radiographic processing, data labeling and preparation

Images were acquired using three CBCT devices: J. Morita 3D Accuitomo 170, Newtom Go, and Newtom Giano HR, all operating at 90 kVp, 5 mA, 17.5 s, with a 0.25 mm voxel size. DICOM files (.dcm) were viewed on a 27-inch UltraSharp LED TFT screen (2560 × 1440, 3.7 MP). Sagittal slices showing SOS were saved as 2D JPEG images (600 dpi, 1024 × 640 pixels) after contrast and brightness adjustments for optimal AI input standardization.

From 262 CBCT datasets, 1,661 sagittal 2D cross-sectional images were extracted. These were classified into four SOS fusion stages based on the Fernández-Pérez et al. (2016) system:

- Stage 0: No fusion,
- Stage 1: Endocranial fusion visible,

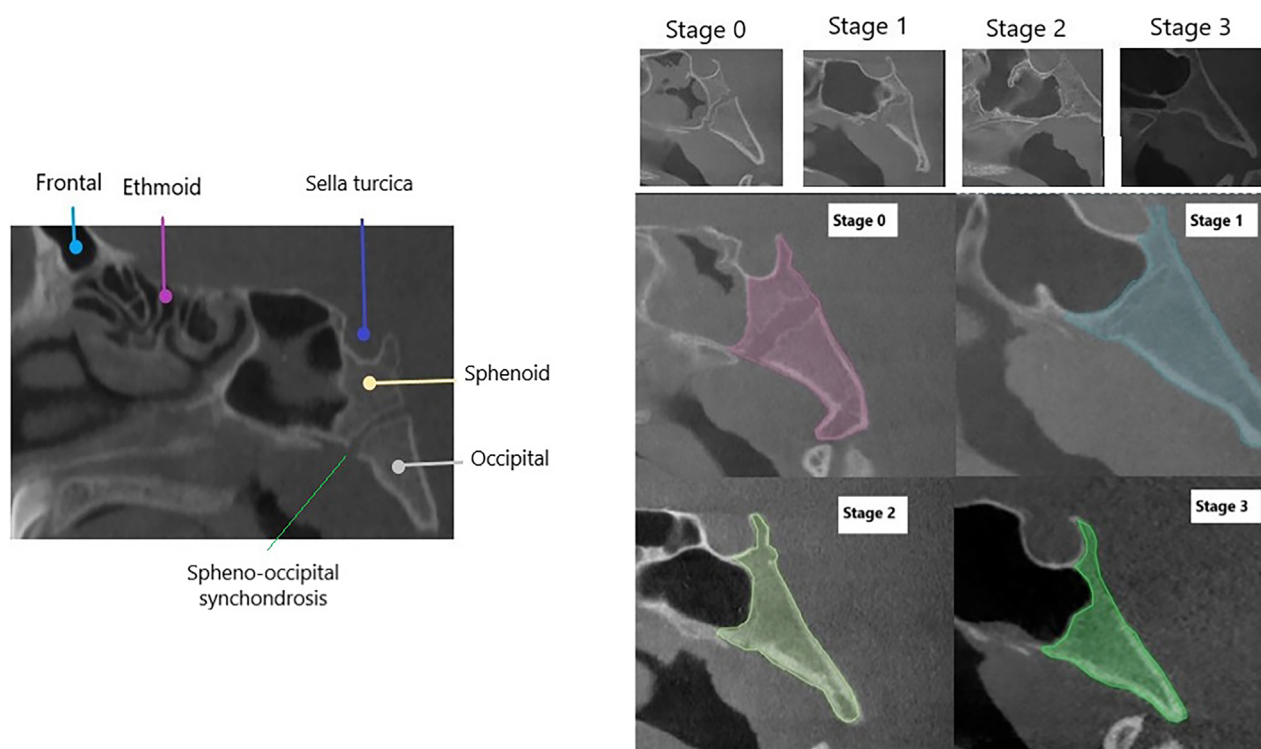


FIGURE 1
SOS fusion stages and polygonal segmentations of stages.

- Stage 2: Ectocranial fusion observed,
- Stage 3: Complete fusion with no gap.

The distribution was as follows: 379 images (Stage 0), 725 (Stage 1), 144 (Stage 2), and 413 (Stage 3).

In our study, all dataset labeling and model development were performed through the CranioCatch artificial intelligence platform (accessible at <https://dentalai.ogu.edu.tr/>), which is a web-based system designed for medical and dental imaging analysis. This platform provides tools for image annotation, dataset management, and AI model training, eliminating the need for direct coding by the researchers. All data were anonymized before being uploaded to the CranioCatch platform (Eskişehir, Türkiye).

Images were labeled using polygonal segmentation in CranioCatch. Structures like the sphenoid body, SOS, and occipital bone were outlined, including cortical boundaries. Labeling was done in four classes (Stage 0–3) (Figure 1). All segmentations were reviewed by two observers—one with 7 years and another with 15 years of experience. Intra- and inter-observer agreement values were excellent (0.995 and 0.983, respectively) (Table 1).

Preprocessing included sharpening unclear images and resizing for model training. Finalized data were split into training (1,329), validation (166), and testing (166) subsets. Specific allocations were:

- Stage 0: 300 train, 46 validation, 33 test
- Stage 1: 585 train, 67 validation, 73 test
- Stage 2: 119 train, 11 validation, 14 test
- Stage 3: 325 train, 42 validation, 46 test

2.4 Segmentation model training

Images were resized to 1024×640 pixels for training with convolutional neural network (CNN) models on the PyTorch platform. YOLOv5, YOLOv8, and YOLOv11—modern, single-stage object detection algorithms—were used due to their speed and accuracy.

Each model underwent 600 training steps, using Stochastic Gradient Descent (SGD) with a batch size of 4. The most successful training step was saved as “best.pt” (124.9 MB). During the test phase, IoU and stability threshold values were set to 0.5.

2.5 Model performance evaluation

Model success was assessed with a confusion matrix, comparing AI predictions to expert-labeled data. Key evaluation metrics included:

- True Positive (TP): Correct identification of fusion stages.
- False Positive (FP): Incorrect classification of non-fusion regions.
- False Negative (FN): Missed detections of actual fusion areas.

From these, the following performance metrics were calculated:

- Sensitivity (Recall) = $TP / (TP + FN)$: Indicates the model's ability to correctly detect SOS fusion.
- Precision = $TP / (TP + FP)$: Reflects how many identified regions were truly SOS fusion stages.

TABLE 1 Performance metrics calculated using 50% IoU threshold of YOLOv5, YOLOv8, and YOLOv11 models, with comparable values from experts and AI models.

AI models							Experts						
YOLOv5	TP	FP	FN	Sensitivity	Accuracy	F1 Score	Intra-observer	TP	FP	FN	Sensitivity	Accuracy	F1 Score
General	154	12	0	1	0.92771084	0.9625							
Stage 0	27	6	6	0.81818182	0.81818182	0.81818182							
Stage 1	68	5	5	0.93150685	0.93150685	0.93150685							
Stage 2	13	1	1	0.92857143	0.92857143	0.92857143	General	1,652	4	5	0.994	0.995	0.995
Stage 3	46	0	0	1	1	1	Stage 0	300	0	0	1.000	1.000	1.000
YOLOv8	TP	FP	FN	Sensitivity	Accuracy	F1 Score	Stage 1	580	2	3	0.991	0.991	0.991
General	153	13	13	0.92168675	0.92168675	0.9216867	Stage 2	118	1	0	0.992	0.992	0.992
Stage 0	29	4	4	0.87878788	0.87878788	0.8787878	Stage 3	325	0	0	1.000	1.000	1.000
Stage 1	68	5	5	0.93150685	0.93150685	0.9315068	Inter-observer	TP	FP	FN	Sensitivity	Accuracy	F1 Score
Stage 2	13	1	1	0.92857143	0.92857143	0.9285714							
Stage 3	43	3	3	0.93478261	0.93478261	0.9347826							
YOLOv11	TP	FP	FN	Sensitivity	Accuracy	F1 Score							
General	157	9	9	0.94578313	0.94578313	0.9457831	General	1,633	13	15	0.983	0.983	0.983
Stage 0	31	2	2	0.93939394	0.93939394	0.9393939	Stage 0	300	0	0	1.000	1.000	1.000
Stage 1	69	4	4	0.94520548	0.94520548	0.9452055	Stage 1	567	9	9	0.969	0.969	0.969
Stage 2	12	2	2	0.85714286	0.85714286	0.8571429	Stage 2	115	3	1	0.966	0.966	0.966
Stage 3	45	1	1	0.97826087	0.97826087	0.9782609	Stage 3	325	0	0	1.000	1.000	1.000

IoU, intersection over union; TP, true positive; FP, false positive; FN, false negative.

- F1 Score = $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$: Balances precision and recall for overall accuracy.
- Mean Average Precision (mAP): A benchmark metric that summarizes model performance across multiple thresholds and is widely used in object detection tasks.

These metrics collectively ensured a thorough validation of model reliability and diagnostic utility.

3 Results

3.1 YOLOv5 labeling model training and test results

In the YOLOv5 model, training was conducted using 1,329 images, and performance was tested on 166 images. Key training metrics showed that the train/box_loss and val/box_loss values were 0.01084 and 0.00726, respectively, indicating high segmentation accuracy. Similarly, train/cls_loss and val/cls_loss were 0.00735 and 0.00436, suggesting effective object classification. The gradual

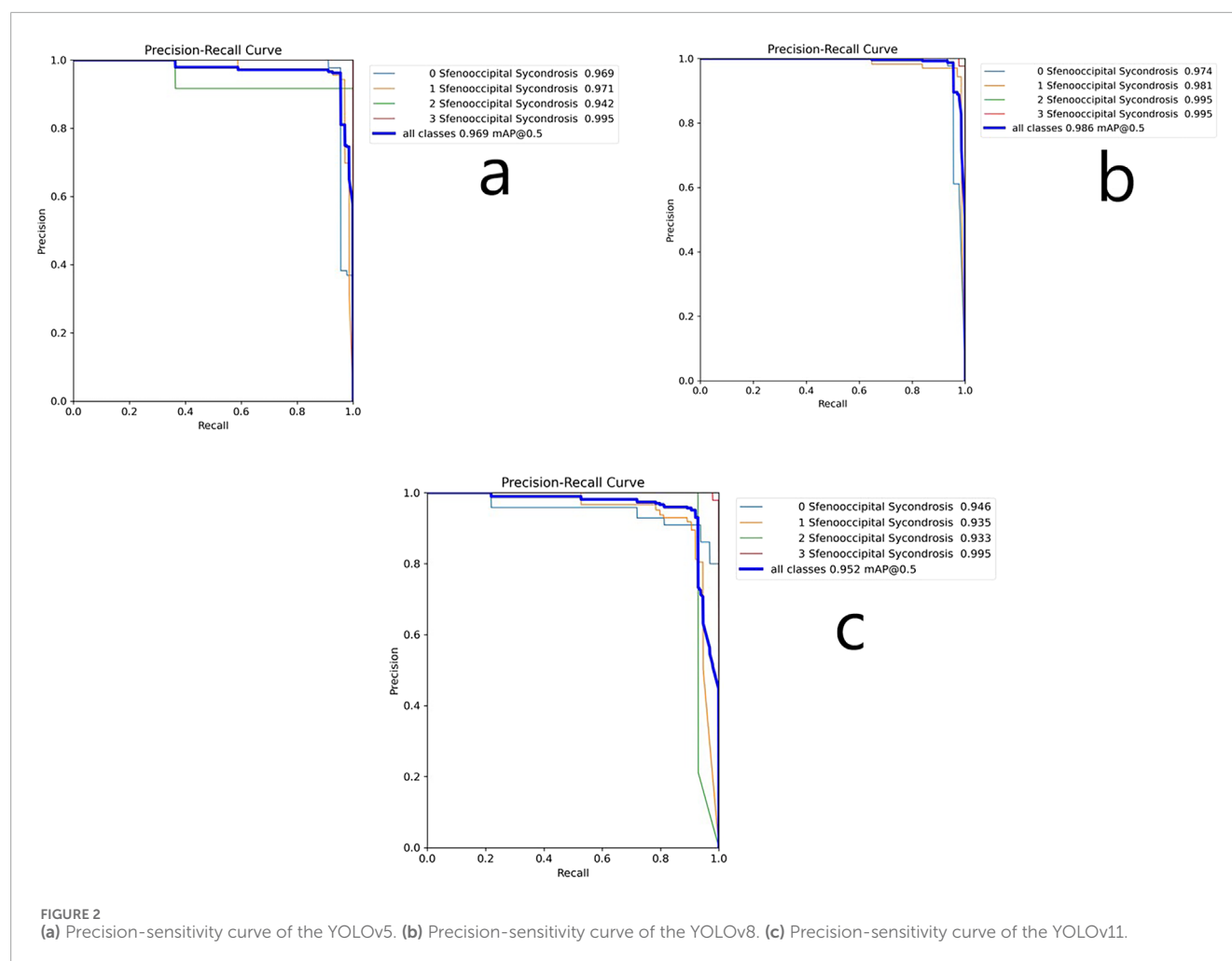
decrease in loss values across epochs reflected a successful learning process.

Figure 2 presents the precision-sensitivity curve at the 0.5 IoU threshold. The largest area under the curve was observed in Stage 3, followed by Stage 1, Stage 0, and Stage 2. The average mAP value was 0.969, showing strong model performance. High precision reflects a low FP rate, and high sensitivity indicates a low FN rate.

Model outputs were evaluated as True Positive (TP), False Positive (FP), or False Negative (FN) depending on their correspondence with expert annotations (Table 1). Confusion matrix analysis revealed an overall F1-score of 0.9625. The highest performance was observed in Stage 3, with 100% recognition. Stage 1 and Stage 2 followed, while Stage 0 had the lowest performance (F1 = 0.81).

3.2 YOLOv8 labeling model training and test results

YOLOv8 was also trained with 1329 images and tested on 166. Training metrics showed train/box_loss = 0.31148 and val/box_loss = 0.46831, with train/cls_loss = 0.31394 and val/cls_loss =



0.29065. These values, though higher than YOLOv5's, remained low and indicated stable training. A unique metric in YOLOv8, dfl_loss , was 0.86602 (train) and 1.0124 (val), suggesting the model effectively adapted to detecting features at varying shapes and sizes.

Figure 2 shows that YOLOv8 reached its highest mAP (0.995) in Stages 2 and 3, followed by Stage 1 (0.981) and Stage 0 (0.974). The average mAP across all stages was 0.986, indicating excellent detection ability.

Evaluation metrics classified model predictions as TP, FP, or FN, compared to radiologist labels (Table 1). The overall F1-score was 0.92. Stage 3 achieved the highest F1 (0.9347), followed closely by Stage 1 (0.9315), Stage 2 (0.92), and Stage 0 (0.87), indicating consistently high performance across all fusion stages.

3.3 YOLOv11 labeling model training and test results

The YOLOv11 model was trained and validated using the same data distribution. During training, the $train/box_loss$ and val/box_loss values were 0.36598 and 0.47306, respectively, and

$train/cls_loss$ and val/cls_loss were 0.33717 and 0.35153. The dfl_loss values were 0.88558 (train) and 0.99697 (val), reflecting high adaptability with minimal error in handling complex image structures.

As shown in Figure 2, YOLOv11 achieved the highest mAP in Stage 3 (0.995), followed by Stage 0 (0.946), Stage 1 (0.935), and Stage 2 (0.933). The overall average mAP was 0.952.

According to model evaluation (Table 1), YOLOv11 achieved an overall F1-score of 0.94. The stage-wise breakdown revealed Stage 3 had the highest F1 (0.97), followed by Stage 1 (0.94), Stage 0 (0.93), and Stage 2 (0.85). Although its accuracy was slightly lower than YOLOv8, it maintained strong consistency across most categories.

3.4 Comparison of performance times of YOLO models

Table 2 summarizes the processing times of the three models. YOLOv5 recorded the longest average time at 40 ms per image, followed by YOLOv8 at 30 ms, and YOLOv11 with the shortest time of 27 ms. Although YOLOv11 offered faster inference,

TABLE 2 Comparison of the results of YOLO models according to SOS fusion stages.

Stages	YOLOv5	YOLOv8	YOLOv11
Stage 0	mAP: 0.969	mAP: 0.974	mAP: 0.946
	F1: 0.81	F1: 0.878	F1: 0.939
Stage 1	mAP: 0.971	mAP: 0.981	mAP: 0.935
	F1: 0.93	F1: 0.931	F1: 0.945
Stage 2	mAP: 0.942	mAP: 0.995	mAP: 0.933
	F1: 0.92	F1: 0.928	F1: 0.85
Stage 3	mAP: 0.995	mAP: 0.995	mAP: 0.995
	F1: 1	F1: 0.934	F1: 0.9782

YOLOv8 showed better overall balance between speed and accuracy.

3.5 Comparison of YOLO models by SOS fusion stage

As detailed in Table 2, YOLOv8 consistently outperformed the others in accuracy and stability across SOS fusion stages. For Stage 0, YOLOv8 had the highest mAP (0.974), followed by YOLOv5 (0.969) and YOLOv11 (0.946). In Stage 1, YOLOv8 again led with a mAP of 0.981. For Stage 2, YOLOv8 achieved a peak mAP of 0.995, outperforming both YOLOv5 and YOLOv11. In Stage 3, all models reached a shared maximum mAP of 0.995, though YOLOv5 achieved a perfect F1 score of 1, showing its strength in this stage.

4 Discussion

Since the beginning of civilization, technological progress has significantly eased workloads. Innovations like electronics, automobiles, computers, and the Internet have transformed numerous sectors, including healthcare, education, and media. Dentistry, which increasingly utilizes digital workflows, has also embraced artificial intelligence (AI) to improve diagnosis, treatment planning, image interpretation, patient management, and automation, thus enhancing oral healthcare quality (Kaya and Koc, 2024; Rahim et al., 2024).

Accurate diagnosis and planning are crucial in clinical decision-making. Studies show that AI-assisted cephalometric analysis (CA) offers more consistent results than manual methods (Lin et al., 2021; Nishimoto et al., 2020; Rahim et al., 2024). In patients undergoing skeletal maturation, AI also helps estimate growth rate and development by analyzing skeletal age, cervical vertebrae, skeletal class, and surgical outcomes (Amasya et al., 2020; Lin et al., 2021; Yu et al., 2020). These capabilities are also valuable in forensic dentistry, especially for age estimation (De Tobel et al., 2017).

Dental age and vertebral development are key in orthodontic planning, particularly since the spheno-occipital synchondrosis

(SOS) is the last cranial synchondrosis to fuse. Conventional skeletal maturation indicators such as the hand-wrist (HW) and cervical vertebrae maturation (CVM) methods have been widely applied, but both present significant limitations. The HW method requires expert knowledge, is time-consuming, has moderate accuracy, and exposes patients to additional radiation. The CVM method, while more convenient, suffers from poor reproducibility, heavy reliance on clinician experience, and limited ability to predict craniofacial growth, especially in female patients with Class II malocclusion. Consequently, neither method guarantees a fully reliable tool for skeletal age assessment, and the orthodontic community recognizes the need for more accurate alternatives (Al-Gumaei et al., 2023).

Clinically, the accurate evaluation of craniofacial growth and treatment response requires stable reference structures for superimposition. Traditional cephalometric superimposition techniques rely on landmarks such as sella, nasion, or basion, but these are subject to growth-related positional changes, which reduces precision and introduces systematic error. Recent advances such as Digital Image Correlation (DIC) applied to cephalometric imaging enable superimposition on growth-stable cranial base structures without reliance on landmarks. DIC with Walker's Point Line Combination (WPLC) has shown the highest precision, surpassing manual and conventional methods. This suggests that AI-driven approaches based on cranial base maturation can reduce observer bias, improve reproducibility, and allow more accurate longitudinal monitoring of growth and treatment outcomes. Looking ahead, combining AI-based SOS classification with advanced digital superimposition methods like DIC may create a comprehensive growth analysis framework that integrates the strengths of CBCT-based SOS staging with stable cranial base references, ultimately providing a reproducible tool for orthodontic and surgical applications (Danz et al., 2024).

In this context, SOS evaluation with CBCT represents a promising approach, as it provides high-resolution three-dimensional imaging of cranial base maturation and offers a valid and reliable indicator of skeletal maturity compared with HW, CVM, and chronological age (Al-Gumaei et al., 2023). Beyond its diagnostic accuracy, integrating AI models to classify SOS fusion stages on CBCT images may enhance orthodontic assessment, improve the prediction of craniofacial syndromes, and support more precise evaluation of developmental completion. Moreover, this approach holds potential value in forensic applications, where accurate skeletal maturity assessment is essential.

The SOS is a cartilaginous joint between the sphenoid and occipital bones and serves as a critical cranial base growth center (Alhazmi et al., 2017). Its timely fusion shapes cranial base morphology and impacts midfacial development. Premature fusion has been linked to midface hypoplasia (Tahiri et al., 2014), and its timing is crucial for adolescent age estimation in forensic science (Sinanoglu et al., 2016). However, studies vary in SOS fusion timelines due to differing methodologies like autopsy, histology, and imaging, including CT and CBCT (Kahana et al., 2003). Among these, 3D imaging modalities, especially CBCT—offer greater accuracy due to high-resolution capabilities (Alhazmi et al., 2017).

Despite the importance of SOS fusion assessment, no standard staging system is universally accepted. Different studies use varied classifications (Bassed et al., 2010; Franklin and Flavel, 2014; Shirley

and Jantz, 2011). Our study adopts Franklin and Flavel (2014) four-stage classification system. Fusion generally occurs between ages 11–14 in females and 13–16 in males (Alhazmi et al., 2017), yet collecting sufficient data remains difficult due to CBCT's limited use in children because of radiation exposure risks.

Comparable findings were reported by Al-Gumaei et al. (2023), who also applied the Franklin and Flavel classification system but labeled the stages from 1 to 4 instead of 0–3. When aligning their Stage 1 with our Stage 0, their results demonstrated that SOS maturation stages represent valid and reliable indicators of maxillary skeletal growth in both genders. Notably, they observed greater increases in maxillary length and height between stages 2 and 3 than between earlier or later stages, whereas changes in maxillary width were more pronounced between stages 1 (our Stage 0) and 2. Growth activity appeared to peak while the SOS was still fusing (particularly stages 2 and 3), with reduced increments after complete fusion (stage 4). Moreover, female patients exhibited earlier acceleration of growth compared with males when assessed by chronological age, although this sex difference was not observed when staging was based directly on SOS maturation. These findings reinforce the clinical relevance of SOS staging as a practical indicator of skeletal maturity, highlighting its potential to optimize treatment timing in orthodontic and orthopedic interventions. In addition, Geng et al. (2024), using the Lottering SOS classification, provided further insight into maxillomandibular growth dynamics across fusion stages. They found that in girls, sagittal maxillary growth remained active until SOS stage 3, slowed at stages 4–5, and continued to decline at stages 5–6. In boys, sagittal maxillary growth was stable until stage 4, with deceleration beginning from stages 5–6. Mandibular growth in both genders followed a pattern of increasing, accelerating, and then decelerating relative growth rates (RGRs) across SOS stages 2–6, with peak mandibular length observed between stages 3–4 and 4–5. These results highlight that SOS maturation reflects not only maxillary but also mandibular growth potential, further underscoring its clinical significance in timing interventions.

As object detection technologies have evolved, convolutional neural networks (CNNs) have replaced earlier algorithms. CNNs offer higher accuracy, particularly with large datasets and adequate computing power (Zhang and Hong, 2019). Object detection models are grouped into single-stage (e.g., YOLO, SSD) and two-stage (e.g., RCNN, Faster RCNN) approaches. Single-stage models prioritize speed with acceptable accuracy, while two-stage models are more precise but slower (Jegham et al., 2024; Vijayakumar and Vairavasundaram, 2024).

This study utilized three Ultralytics-supported single-stage models—YOLOv5, YOLOv8, and YOLOv11. The original YOLO (You Only Look Once) introduced by Redmon (2016) revolutionized object detection by predicting bounding boxes and class probabilities simultaneously (Hussain, 2023). To maintain comparability, unsupported versions (e.g., YOLOv1, v2, v4, v6, v7) were excluded due to architectural differences (Jegham et al., 2024).

YOLOv5, launched in 2020 by Glen Jocher, introduced CSPDarknet as a backbone, improving computational efficiency (Ultralytics, 2021; Jocher et al., 2020). YOLOv8 (2023) added the C2f module and advanced context fusion for enhanced object detection (Jegham et al., 2024). YOLOv11 (2024) further incorporated the C2PSA module—combining partial structures and self-attention

for better detection of small or obscured features (Jocher and Qiu, 2024).

Mean average precision (mAP) is the preferred evaluation metric in object detection due to class imbalance challenges (Vijayakumar and Vairavasundaram, 2024). In our results, YOLOv5 yielded mAP 0.969 and F1-score 0.9625; YOLOv8 achieved mAP 0.986 and F1-score 0.9216; YOLOv11 reached mAP 0.952 with F1-score 0.945. YOLOv8 performed most consistently and accurately, aligning with previous findings (Fitria et al., 2024; Deeptho et al., 2024; Bonfanti-Gris et al., 2024).

Though YOLOv11 had the fastest inference time (27 ms), it showed greater accuracy and variability, raising stability concerns. Özcan et al. (2024) similarly observed YOLOv8 outperforming YOLOv11 in dental landmark detection. Despite YOLOv11's efficient C3k2 architecture, YOLOv8 maintained superior reliability.

All models showed peak performance in Stage 3 detection (mAP: 0.995), likely due to the distinct radiographic signs of complete fusion. While results varied in other stages, YOLOv8 outperformed others, and YOLOv11 had the lowest sensitivity.

In this study, experts achieved slightly higher sensitivity and accuracy than the AI models, particularly in Stage 0 and Stage 3, where their performance was perfect. These differences are expected, as the AI models were trained on expert-labeled data, thereby validating the reliability of the ground truth used for training. Most FN and FP results produced by the AI corresponded to borderline cases or image artifacts, which are typically recognizable by experienced observers. This suggests that AI errors are not arbitrary but remain visually interpretable, supporting the complementary role of expert review. For this reason, the most effective diagnostic workflow would involve AI providing a preliminary classification subsequently reviewed and confirmed by experts, combining the reproducibility and efficiency of AI with the diagnostic assurance of human expertise. It should also be noted that AI performance was calculated on the test dataset, whereas expert sensitivity and specificity were derived from the entire dataset, limiting direct comparability. Slightly higher values in expert evaluation should therefore be seen not as a shortcoming of AI but as confirmation of the reliability of expert annotations. The high concordance between experts and AI highlights the reproducibility of the system and its potential to replicate expert-level staging in a rapid and automated manner.

The main study limitation was the difficulty of assembling a large, balanced dataset due to age restrictions and radiation concerns. Pediatric images also showed motion artifacts and anatomical variation, affecting generalizability. Still, the models performed robustly. Future work should involve larger, multi-center datasets to validate these findings.

5 Conclusion

Ultralytics' YOLO models (YOLOv5, YOLOv8, and YOLOv11) accurately detect SOS fusion stages in CBCT images, with mAP scores above 95% and F1-scores over 90%. These AI-based approaches enhance growth monitoring and early diagnosis of craniofacial anomalies. YOLOv8's superior performance highlights the importance of model selection in improving treatment

outcomes. The study demonstrates the potential of deep learning in medical imaging and suggests that future research with larger datasets and broader clinical applications could lead to widespread clinical adoption of these models.

Data availability statement

The datasets generated and analyzed during the current study are not publicly available due to privacy and ethical restrictions but are available from the corresponding author on reasonable request.

Ethics statement

The studies involving humans were approved by Necmettin Erbakan University Non-Drug and Non-Medical Device Research Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

SU: Validation, Project administration, Visualization, Writing – original draft, Conceptualization, Data curation, Writing – review and editing, Investigation, Methodology. GM: Project administration, Methodology, Conceptualization, Supervision, Formal Analysis, Writing – review and editing. CE: Writing – review and editing, Investigation, Visualization, Data curation.

References

- Al-Gumaei, W. S., Long, H., Al-Attar, R., Elayah, S. A., Alhammadi, M. S., Almagrabi, I., et al. (2023). Comparison of three-dimensional maxillary growth across sphenoid-occipital synchondrosis maturation stages. *BMC Oral Health* 23 (1), 100. doi:10.1186/s12903-023-02774-w
- Alhazmi, A., Vargas, E., Palomo, J. M., Hans, M., Latimer, B., and Simpson, S. J. (2017). Timing and rate of sphenoid-occipital synchondrosis closure and its relationship to puberty. *PLoS ONE* 12, e0183305. doi:10.1371/journal.pone.0183305
- Amasya, H., Cesur, E., Yildirim, D., and Orhan, K. (2020). Validation of cervical vertebral maturation stages: artificial intelligence vs human observer visual analysis. *Am. J. Orthod. Dentofac. Orthop.* 158, e173–e179. doi:10.1016/j.ajodo.2020.08.014
- Bassed, R. B., Briggs, C., and Drummer, O. H. (2010). Analysis of time of closure of the sphenoid-occipital synchondrosis using computed tomography. *Forensic Sci. Int.* 200, 161–164. doi:10.1016/j.forsciint.2010.04.009
- Bazargani, F., Feldmann, I., and Bondemark, L. (2013). Three-dimensional analysis of effects of rapid maxillary expansion on facial sutures and bones: a systematic review. *Angle Orthod.* 83, 1074–1082. doi:10.2319/020413-103.1
- Björk, A. (1955). Facial growth in man studied with the aid of metallic implants. *Acta Odontol. Scand.* 13, 9–34. doi:10.3109/00016355509028170
- Bonfanti-Gris, M., Herrera, A., Paraiso-Medina, S., Alonso-Calvo, R., Martínez-Rus, F., and Pradies, G. (2024). Performance evaluation of three versions of a convolutional neural network for object detection and segmentation using a multiclass and reduced panoramic radiograph dataset. *J. Dent.* 144, 104891. doi:10.1016/j.jdent.2024.104891
- Cendekiawan, T., Wong, R. W., and Rabie, A. B. M. (2010). Relationships between cranial base synchondroses and craniofacial development: a review. *Open Anat. J.* 2, 67–75. doi:10.2174/1877609401002010067
- Dalili Kajan, Z., Hadinezhad, A., Khosravifard, N., Gholinia, F., Rafiei, E., and Ghandari, F. (2021). Fusion patterns of the sphenoid-occipital synchondrosis in the age range of 9–22: a computed tomography analysis. *Orthod. Craniofacial Res.* 24, 405–413. doi:10.1111/ocr.12455
- Danz, J. C., Stöckli, S., and Rank, C. P. (2024). Precision and accuracy of craniofacial growth and orthodontic treatment evaluation by digital image correlation: a prospective cohort study. *Front. Oral Health* 5, 1419481. doi:10.3389/froh.2024.1419481
- Deeptho, C., Khlaisuwan, V., Pengchai, C., Intarachana, W., Rakchuai, P., Kajhan, K., et al. (2024). "Toward the development of an oral-diagnosis framework: a case study of teeth segmentation and numbering in bitewing radiographs via YOLO models," in IEEE International Conference on Cybernetics and Innovations (ICCI). IEEE, 1–6.
- De Tobel, J., Radesch, P., Vandermeulen, D., and Thevissen, P. W. (2017). An automated technique to stage lower third molar development on panoramic radiographs for age estimation: a pilot study. *J. Forensic Odonto-Stomatol.* 35, 42–54.
- Demirtürk Kocasarac, H., Sinanoglu, A., Noujeim, M., Helvacioğlu Yigit, D., and Baydemir, C. (2016). Radiologic assessment of third molar tooth and sphenoid-occipital synchondrosis for age estimation: a multiple regression analysis study. *Int. J. Leg. Med.* 130, 799–808. doi:10.1007/s00414-015-1298-8
- Evli, C., Uzun, S., and Mağat, G. (2025). Evaluation of growth and development period according to sphenoid-occipital synchondrosis fusion stages in cone-beam

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by Necmettin Erbakan University Scientific Research Projects Coordination Unit (BAP) under project number 24DU24002.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- computed tomography with ImageJ program. *Sci. Rep.* 15, 13821. doi:10.1038/s41598-025-92098-2
- Fernández-Pérez, M. J., Alarcón, J. A., McNamara, J. A., Velasco-Torres, M., Benavides, E., Galindo-Moreno, P., et al. (2016). Spheno-occipital synchondrosis fusion correlates with cervical vertebrae maturation. *PLoS ONE* 11, e0161104. doi:10.1371/journal.pone.0161104
- Fitria, M., Elma, Y., Oktiana, M., Saddami, K., Novita, R., Putri, R., et al. (2024). Technology I. The Deep Learning Model for Decayed-Missing-Filled Teeth Detection: A Comparison Between YOLOV5 and YOLOV8. *Jord. J. Comput. Inf. Technol.* 10 (3), 335–350.
- Franklin, D., and Flavel, A. (2014). Brief communication: timing of spheno-occipital closure in modern Western Australians. *Am. J. Phys. Anthropol.* 153, 132–138. doi:10.1002/ajpa.22399
- Funato, N., Srivastava, D., Shibata, S., and Yanagisawa, H. (2020). TBX1 regulates chondrocyte maturation in the spheno-occipital synchondrosis. *J. Dent. Res.* 99, 1182–1191. doi:10.1177/0022034520925080
- Geisler, E. L., Agarwal, S., Hallac, R. R., Daescu, O., and Kane, A. A. (2021). A role for artificial intelligence in the classification of craniofacial anomalies. *J. Craniofac. Surg.* 32, 967–969. doi:10.1097/SCS.00000000000007369
- Geng, J., Zhao, G., and Gu, Y. (2024). Feasibility of spheno-occipital synchondrosis fusion stages as an indicator for the assessment of maxillomandibular growth: a mixed longitudinal study. *Orthod. and Craniofacial Res.* 27 (4), 589–597. doi:10.1111/ocr.12774
- Goldstein, J. A., Paliga, J. T., Wink, J. D., Bartlett, S. P., Nah, H. D., and Taylor, J. A. (2014). Earlier evidence of spheno-occipital synchondrosis fusion correlates with severity of midface hypoplasia in patients with syndromic craniosynostosis. *Plast. Reconstr. Surg.* 134, 504–510. doi:10.1097/PRS.0000000000000419
- Halpern, R. M. (2014). Spheno-occipital synchondrosis maturation as related to the development of cervical vertebrae, mandibular canine and chronologic age: a cone-beam computed tomography analysis.
- Hoshino, Y., Takechi, M., Moazen, M., Steacy, M., Koyabu, D., Furutera, T., et al. (2022). Synchondrosis fusion contributes to the progression of postnatal craniofacial dysmorphology in syndromic craniosynostosis. *J. Anat.* 242, 387–401. doi:10.1111/joa.13790
- Hussain, M. (2023). YOLO-v1 to YOLO-v8: the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines* 11, 677. doi:10.3390/machines11070677
- Jegham, N., Koh, C. Y., Abdelatti, M., and Hendawi, A. (2024). Evaluating the evolution of YOLO (you only look once) models: a comprehensive benchmark study of YOLOv11 and its predecessors. Preprint arXiv:2402.12345.
- Jocher, G., and Qiu, J. (2024). Ultralytics YOLOv11. Available online at: <https://docs.ultralytics.com/tr/models/yolo11/> (Accessed June 10, 2025).
- Jocher, G., Stoken, A., Borovec, J., Changyu, L., Hogan, A., Diaconu, L., et al. (2020). ultralytics/yolov5: v3.1 – bug fixes and performance improvements. *Zenodo*. Available online at: <https://zenodo.org/records/4154370> (Accessed October 22, 2025).
- Kahana, T., Birkby, W., and Hiss, J. (2003). Estimation of age in adolescents—The basilar synchondrosis. *J. Forensic Sci.* 48, 1–5. doi:10.1520/jfs2001400
- Kaya, S., and Koc, A. (2024). Radiologic evaluation of associated symptoms and fractal analysis of unilateral dens invaginatus cases. *Oral Radiol.* 40, 484–491. doi:10.1007/s11282-024-00756-4
- Krishan, K., and Kanchan, T. (2013). Evaluation of spheno-occipital synchondrosis: a review of literature and considerations from forensic anthropologic point of view. *J. Forensic Dent. Sci.* 5, 72–76. doi:10.4103/0975-1475.119764
- Lin, H. H., Chiang, W. C., Yang, C. T., Cheng, C. T., Zhang, T., and Lo, L. J. (2021). On construction of transfer learning for facial symmetry assessment before and after orthognathic surgery. *Comput. Methods Programs Biomed.* 200, 105928. doi:10.1016/j.cmpb.2021.105928
- McGrath, J., Gerety, P. A., Derderian, C. A., Steinbacher, D. M., Vossough, A., Bartlett, S. P., et al. (2012). Differential closure of the spheno-occipital synchondrosis in syndromic craniosynostosis. *Plast. Reconstr. Surg.* 130, 681e–689e. doi:10.1097/PRS.0b013e318267d4c0
- Melsen, B. (1974). *The cranial base: the postnatal development of the cranial base studied histologically on human autopsy material.* Acta Odontologica Scandinavica. Supplementum 62. Copenhagen: Munksgaard.
- Nie, X. (2005). Cranial base in craniofacial development: developmental features, influence on facial growth, anomaly, and molecular basis. *Ann. Anat.* 187, 127–135. doi:10.1080/00016350510019847
- Nishimoto, S., Kawai, K., Fujiwara, T., Ishise, H., and Kakibuchi, M. (2020). Locating cephalometric landmarks with multi-phase deep learning.
- Özcan, T., Karayılan, R., and Yılmaz, S. (2024). Artificial intelligence-assisted automatic detection of anatomical landmarks in panoramic radiographs (Panoramik radyograflarda anatomik yer işaretlerinin yapay zeka destekli otomatik tespiti). *Erciyes University Journal of the Institute of Science and Technology*, 40, 535–558.
- Rahim, A., Khatoun, R., Khan, T. A., Syed, K., Khan, I., Khalid, T., et al. (2024). Artificial intelligence-powered dentistry: probing the potential, challenges, and ethicality of artificial intelligence in dentistry. *Digit. Health* 10, 20552076241291345. doi:10.1177/20552076241291345
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: unified, real-time object detection. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 779–788. doi:10.1109/cvpr.2016.91
- Shirley, N. R., and Jantz, R. L. (2011). Spheno-occipital synchondrosis fusion in modern Americans. *J. Forensic Sci.* 56, 580–585. doi:10.1111/j.1556-4029.2011.01705.x
- Sinanoglu, A., Kocasarac, H. D., and Noujeim, M. (2016). Age estimation by an analysis of spheno-occipital synchondrosis using cone-beam computed tomography. *Leg. Med.* 18, 13–19. doi:10.1016/j.legalmed.2015.11.004
- Singh, S., Jain, R. K., Naveed, N., and Balasubramaniam, A. (2025). Spheno-occipital synchondrosis as a reliable indicator of skeletal maturity: a systematic review and meta-analysis. *J. Dent. Res. Dent. Clin. Dent. Prospects* 19 (1), 1–8. doi:10.34172/joddd.025.41168
- Solow, B. (1980). The dentoalveolar compensatory mechanism: background and clinical implications. *Br. J. Orthod.* 7 (3), 145–161. doi:10.1179/bjo.7.3.145
- Tahiri, Y., Paliga, J. T., Vossough, A., Bartlett, S. P., and Taylor, J. A. (2014). The spheno-occipital synchondrosis fuses prematurely in patients with Crouzon syndrome and midface hypoplasia compared with age- and gender-matched controls. *J. Oral Maxillofac. Surg.* 72, 1173–1179. doi:10.1016/j.joms.2013.11.015
- Ultralytics (2021). YOLOv5: a state-of-the-art real-time object detection system. Available online at: <https://github.com/ultralytics/yolov5> (Accessed June 10, 2025).
- Vijayakumar, A., and Vairavasundaram, S. (2024). YOLO-based object detection models: a review and its applications. *Multimed. Tools Appl.* 83 (35), 83535–83574. doi:10.1007/s11042-024-18872-y
- Yang, L. (2016). Fusion pattern of the spheno-occipital synchondrosis in Class I and Class III malocclusion: a CT study. *Angle Orthod.* 86 (4), 569–577. doi:10.2319/052218-386.1
- Yu, H., Cho, S., Kim, M., Kim, W., Kim, J., and Choi, J. (2020). Automated skeletal classification with lateral cephalometry based on artificial intelligence. *J. Dent. Res.* 99, 249–256. doi:10.1177/0022034520901715
- Zhang, H., and Hong, X. (2019). Recent progresses on object detection: a brief review. *Multimed. Tools Appl.* 78, 27809–27847. doi:10.1007/s11042-019-07898-2