Individualized Facial Growth Prediction Models Based on Artificial Intelligence

2024 Orthodontic Faculty Development Fellowships (OFDFA)

Dr Heeyeon Suh

hsuh1@pacific.edu
O: 443-488-4007

FollowUp Form

Award Information

In an attempt to make things a little easier for the reviewer who will read this report, please consider these two questions before this is sent for review:

- *Is this an example of your very best work, in that it provides sufficient explanation and justification, and is something otherwise worthy of publication? (We do publish the Final Report on our website, so this does need to be complete and polished.)*

- *Does this Final Report provide the level of detail, etc. that you would expect, if you were the reviewer?*

Title of Project*

Individualized Facial Growth Prediction Models Based on Artificial Intelligence

Award Type

Orthodontic Faculty Development Fellowship Award (OFDFA)

Period of AAOF Support

July 1, 2024 through June 30, 2025

Institution

University of the Pacific, Arthur A. Dugoni School of Dentistry

Names of principal advisor(s) / mentor(s), co-investigator(s) and consultant(s)

Heesoo Oh, Shin-Jae Lee, Jun-Ho Moon, Dongyub Ko

Amount of Funding

$29,934.00

Abstract

(add specific directions for each type here)

Abstract uploaded.

Respond to the following questions:

Detailed results and inferences:*

If the work has been published, please attach a pdf of manuscript below by clicking "Upload a file".
<u>OR</u>
Use the text box below to describe in detail the results of your study. The intent is to share the knowledge you have generated with the AAOF and orthodontic community specifically and other who may benefit from your study. Table, Figures, Statistical Analysis, and interpretation of results should also be attached by clicking "Upload a file".

Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection.pdf
Our study, "Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection" has been accepted for publication. This study used 1,257 pairs of longitudinal lateral cephalograms from 33 subjects in the Mathews Growth Collection to develop growth prediction models. Manual landmark location included 46 hard and 32 soft tissue landmarks. Two prediction methods were compared: TabNet-based deep learning (AI) and Partial Least Squares (PLS). Results showed that AI outperformed PLS, with an average of 0.61 mm less prediction error. AI was more accurate in 60 out of 77 landmarks, especially with increased training epochs. Both methods performed better on hard tissue and maxillary landmarks than on soft tissue and mandibular ones, but AI showed smaller error increases in high-variability areas. AI demonstrated strong potential for growth prediction, achieving clinically acceptable accuracy (1.49 mm for hard tissue, 1.71 mm for soft tissue), and generally surpassed PLS, particularly in regions with greater anatomical variation.

Were the original, specific aims of the proposal realized?*

Yes.
Aim 1: Build a growth prediction model, including hard and soft tissue landmarks, using the PLS method.
- We successfully built a prediction model using the Partial Least Squares (PLS) method.
Aim 2: Build a growth prediction model, including hard and soft tissue landmarks, using AI based on a deep learning algorithm.
- We developed a deep learning-based AI prediction model incorporating both hard and soft tissue landmarks.
Aim 3: Evaluate and validate the models through cross-validation.
- Both the PLS and AI models were evaluated and validated using cross-validation methods.
Aim 4: Compare the accuracy of the PLS and AI prediction models.
- We compared prediction accuracy of the PLS and AI models.

Were the results published?*

Yes

Have the results of this proposal been presented?*

Yes

To what extent have you used, or how do you intend to use, AAOF funding to further your career?*

I have used $19,834 of AAOF funding to supplement faculty salary or stipend, which allowed me to dedicate necessary time to research.

I have used $6,728 of AAOF funding to support tuition for my data science master's program, which has provided me with a strong foundation in data management, data sharing, and machine learning methods. This training has been instrumental in enhancing my ability to conduct advanced data analysis and apply artificial intelligence techniques in orthodontic research.

I have used $1,429 for travel to academic conferences, where I had the opportunity to present my project findings and engage with the broader orthodontic research community. These experiences have significantly contributed to my professional development.

Additionally, I have used $1,843 for research equipment and supplies.

The remaining funds ($100) will be used to finalize the publication process for this project.

Accounting: Were there any leftover funds?

$0.00

Published

Citations*

You indicated results have been published. Please list the cited reference/s for publication/s including titles, dates, author or co-authors, journal, issue and page numbers

Roseth J, Kim JH, Moon JH, Ko DY, Oh H, Lee SJ, Suh H. Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection. Angle Orthod. 2025 May 1;95(3):249-258. doi: 10.2319/082124-687.1. PMID: 39884314; PMCID: PMC12017544.

Was AAOF support acknowledged?

If so, please describe:

AAOF support was acknowledged in this publication: This study was supported by the AAOF Orthodontic Faculty Development Fellowship Award (OFDFA 2024) and National Institute of Dental and Craniofacial Research (NIDCR) Loan Repayment Program Award.

Presented

Please list titles, author or co-authors of these presentation/s, year and locations:*

1. Facial growth prediction models based on the artificial intelligence and partial least squares – with longitudinal growth data from Mathews growth collection, Heeyeon Suh, Jeff Roseth, Jun-Ho Moon, Shin-Jae Lee, Heesoo Oh, 2024, Consortium for Orthodontic Advances in Science and Technology (COAST), Lake Arrowhead, CA, US

2. Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection, Jeff Roseth, Heeyeon Suh, Shin-Jae Lee, Heesoo Oh, 2024, Western Orthodontic Conference (WOCON), Palm Springs, CA, US

Was AAOF support acknowledged?

If so, please describe:

Yes, for the presentation titled 'Facial Growth Prediction Models Based on Artificial Intelligence and Partial Least Squares – Using Longitudinal Growth Data from the Mathews Growth Collection,' AAOF support was acknowledged: This study was supported by the American Association of Orthodontists Foundation Orthodontic Faculty Development Fellowship Award (AAOF OFDFA 2024) and National Institute of Dental & Craniofacial Research (NIDCR) Loan Repayment Program Award.

Internal Review

Reviewer comments

Reviewer Status*

File Attachment Summary

Applicant File Uploads

- Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection.pdf

Comparison of individualized facial growth prediction models using artificial intelligence and partial least squares based on the Mathews growth collection

Jeffrey Roseth[a]; **Jong-Hak Kim**[b]; **Jun-Ho Moon**[c]; **Dong-Yub Ko**[d]; **Heesoo Oh**[e];
Shin-Jae Lee[f]; **Heeyeon Suh**[g]

ABSTRACT

Objectives: To develop facial growth prediction models using artificial intelligence (AI) under various conditions, and to compare performance of these models with each other as well as with the partial least squares (PLS) growth prediction model.

Materials and Methods: Longitudinal lateral cephalograms from 33 subjects in the Mathews growth collection were utilized. A total of 1257 pairs of before and after growth lateral cephalograms were included. In each image, 46 hard and 32 soft tissue landmarks were manually identified. Growth prediction models were constructed using a deep learning method based on TabNet deep neural network and partial least squares (PLS) method. Prediction accuracies of the two methods were compared.

Results: On average, artificial intelligence (AI) showed 0.61 mm less prediction error than PLS. Among the 77 predicted landmarks, AI was more accurate than PLS in 60 landmarks. When comparing AI models with varying numbers of training epochs, those with higher epochs yielded more accurate predictions. Overall, PLS and AI exhibited greater prediction errors for soft tissue and mandibular landmarks compared to hard tissue and maxillary landmarks. However, AI showed a smaller increase in prediction error in areas with greater variability.

Conclusions: AI proved to be a valuable growth prediction method, with clinically acceptable prediction errors averaging 1.49 mm for 45 hard tissue landmarks and 1.71 mm for 32 soft tissue landmarks. PLS accurately predicted landmarks with low variability. However, AI generally outperformed PLS, particularly for landmarks in the lower part of the craniofacial structure and soft tissue, where uncertainty is considerable. (*Angle Orthod*. 2025;95:249–258.)

KEY WORDS: Growth prediction; Longitudinal craniofacial growth records; Artificial intelligence; Partial least squares

[a] Resident, Department of Orthodontics, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California, USA.

[b] PhD Graduate Student, Department of Orthodontics, Graduate School, Seoul National University, Seoul, Korea.

[c] Private Practice, Cheonan, Korea.

[d] Research Scientist, AI Research Center, DDH Inc, Seoul, Korea.

[e] Professor and Chair, Department of Orthodontics, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California, USA.

[f] Professor, Department of Orthodontics and Dental Research Institute, Seoul National University School of Dentistry, Seoul, Korea.

[g] Assistant Professor, Department of Orthodontics, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California, USA.

Corresponding author: Dr Heeyeon Suh, Department of Orthodontics, Arthur A. Dugoni School of Dentistry, University of the Pacific, San Francisco, California 94103, USA
(e-mail: hsuh1@pacific.edu)

INTRODUCTION

Understanding and predicting timing, pattern, and amount of human facial growth greatly impacts the effectiveness and efficiency of orthodontic treatment.[1] Although some patients benefit from early intervention, others may miss critical windows, necessitating surgery as the only viable option. Additionally, some patients receive multiple rounds of treatment as they outgrow initial results or experience relapse. Ideally, with unlimited resources, prolonged treatments could yield optimal outcomes. Nevertheless, the best available scientific evidence should be used to determine the most effective and efficient treatment options.

Historical growth prediction methods[2–10] provided general guidelines but were not always precise for individual variations. Despite efforts to understand and predict growth and development, the subjectivity in predicting dentofacial growth remains a challenge, as highlighted in Dr. Bishara's article in 2000.[11] Although growth prediction has been an important subject in orthodontics, few publications have addressed craniofacial growth prediction in the past 20 years.[12] The precise determination of future growth magnitude, direction, and resulting facial changes continues to be uncertain, with most treatment planning relying on the subjective assessment of orthodontists.

Accurate growth prediction is challenging due to its complexity and the influence of genetic and environmental factors, which cause individual variations.[13–15] As an attempt to account for these individual variations, recent statistical methods, such as discriminant analysis,[16,17] multiple linear regression analysis,[18] Bayes' theorem,[19,20] and nonlinear growth models,[21,22] have included age and gender in growth prediction. The multivariate partial least squares regression method (PLS) has been utilized in growth prediction for its ability to manage a large number of intercorrelated individual attributes and has shown improved prediction accuracy.[12] Recently, there has been growing interest in applying artificial intelligence (AI) to solve complex problems in orthodontics. There have been attempts to use AI in cephalometric landmark detection, automatic image superimposition, orthodontic diagnosis, and growth prediction.[23–30]

Although advances in technology allow for extensive computational analysis of large datasets, developing a robust growth prediction model remains challenging, as collecting longitudinal growth records solely for research purposes is often not feasible, especially when patients are not undergoing treatment. Therefore, the American Association of Orthodontists Foundation (AAOF) Craniofacial Growth Legacy Collection serves as an invaluable resource, providing longitudinal records of growing adolescents.[31]

Given the critical role of growth in successful treatment, it is imperative to base growth predictions on the best available scientific knowledge. This study aimed to develop facial growth prediction models using the Mathews collection from the AAOF growth collections, based on AI. Another goal was to compare the performance of the AI model with a prediction model that utilized the PLS method, which is one of the most recently implemented statistical approaches.

MATERIALS AND METHODS

Subjects

The institutional review board for the protection of human subjects of University of The Pacific reviewed and approved the research protocol (#2023-28). Subjects of this study were collected from one of the AAOF Growth Collections, the Mathews collection, which includes 36 subjects, primarily of European descent. Three subjects were excluded due to missing images, resulting in a final sample of 33 subjects who received annual cephalograms resulting in at least five timepoints, yielding 1257 before and after growth pair data.

Cephalometrics

Cephalometric images of the subjects, taken annually, were processed digitally to enhance quality for reliable landmark identification. Fiducials, pixel information, and magnification factors were considered to resize the images. A total of 78 anatomical landmarks (Table 1) were identified manually by a single examiner (SJL) with 33 years of clinical orthodontic experience (Figure 1). A Cartesian coordinate system was constructed using Sella as the origin, resulting in 77 landmarks for prediction. The horizontal reference plane was established by drawing a line 7° downward from the Sella-Nasion plane.

Variables

The prediction model incorporated 159 predictor variables and 154 response variables. Predictor variables included individual characteristics such as age, gender, Angle classification, growth observation interval, and the x and y coordinates of 77 anatomic landmarks from the starting timepoint. The x and y coordinates of 77 anatomic landmarks from a later timepoint were used as response variables.

AI and PLS Prediction Models

The leave-one-out cross-validation (LOOCV) was employed to calculate test errors.[32] The TabNet Deep Neural Network (DNN) by Arik and Pfister[33] was

Table 1. List of Anatomical Landmarks

Category	Landmark	Definition
Skeletal	Sella	The geometric center of the sella turcica: Used as the origin
	Nasion	The most anterior and median point along the frontonasal suture
	Nasal bone tip	The inferior tip of the nasal bone
	Porion	The most superior point located on the external auditory meatus
	Orbitale	The lowest point on the infraorbital margin
	Key ridge	Most inferior point of the zygomatic ridge
	Key ridge contour smoothing point 1[a]	A point between Orbitale and Key ridge—upper
	Key ridge contour smoothing point 2[a]	A point between Orbitale and Key ridge—middle
	Key ridge contour smoothing point 3[a]	A point between Orbitale and Key ridge—lower
	Anterior nasal spine	The most anterior point on the maxilla
	Posterior nasal spine	The most posterior point of the bony hard palate
	Point A	The deepest concavity along the anterior contour of the maxilla
	Point A contour smoothing point[a]	A point between Point A and Supradentale
	Supradentale	The lowest and most anterior point on the alveolar bone along the anterior contour of the maxilla
	Infradentale	The highest and most anterior point on the alveolar bone along the anterior contour of the mandible
	Point B contour smoothing point[a]	A point between Infradentale and Point B
	Point B	The deepest concavity on the anterior contour of the mandible between the chin and the mandibular alveolar process
	Protuberance menti	The point at which the shape of symphysis changes from convex to concave
	Pogonion	The most anterior point on the symphysis of the mandible
	Gnathion	The most anterior and inferior point on the bony chin between Pogonion and Menton
	Menton	The lowest point on the mandibular symphysis
	Gonion, constructed	A point on the bony contour of the gonial angle, determined by bisecting the tangent angle
	Mandibular body contour smoothing point 1[a]	A point on the mandibular lower border—anterior
	Mandibular body contour smoothing point 2[a]	A point on the mandibular lower border—middle
	Mandibular body contour smoothing point 3[a]	A point on the mandibular lower border—posterior
	Gonion, anatomic	The midpoint of the mandibular angle between the ramus and the mandibular corpus
	Gonion notch contour smoothing point[a]	A point on the mandibular lower border at the gonial angle
	Gonion ramus contour smoothing point[a]	A point on the posterior ramus border at the gonial angle
	Articulare	The point at the junction of the posterior border of the ramus and the inferior border of the posterior cranial base
	Ramus contour smoothing point 1[a]	A point on the posterior ramus border between Articulare and Gonion—upper
	Ramus contour smoothing point 2[a]	A point on the posterior ramus border between Articulare and Gonion—lower
	Condylion	The most superior point on the condyle of the mandible
	Ramus tip	The tip of the coronoid process
	Pterygomaxillary fissure	The most inferior point on the outline of Pterygomaxillary fissure
	Pterygoid point	The posterior-most and superior-most point in the upper contour of the Pterygomaxillary fissure
	Basion	The most inferior point on the anterior border of the foramen magnum
Dental	Upper incisal root	The apex of the maxillary central incisor
	Upper incisal crown	The incisal tip of the maxillary central incisor
	Lower incisal crown	The incisal tip of the mandibular central incisor
	Lower incisal root	The apex of the mandibular central incisor
	U6 crown mesial edge	The most mesial point on the maxillary first molar crown
	U6 mesiobuccal cusp tip	The mesiobuccal cusp tip of the maxillary first molar
	U6 mesial root tip	The mesial root tip of the maxillary first molar
	L6 crown mesial edge	The most mesial point on the mandibular first molar crown
	L6 mesiobuccal cusp tip	The mesiobuccal cusp tip of the mandibular first molar
	L6 mesial root tip	The mesial root tip of the mandibular first molar
Soft tissue	glabella	The most prominent or anterior point of the forehead
	glabella-nasion contour point[a]	A point between glabella and soft tissue nasion
	soft tissue nasion	The most posterior point on the soft tissue profile between glabella and pronasale
	Inferior tip of nasal bone[a]	A point between soft-tissue nasion and supranasal tip—upper
	deepest point of the nose[a]	A point between soft-tissue nasion and supranasal tip—lower

Table 1. Continued

Category	Landmark	Definition
	supranasal tip	The most anterosuperior point on the outline of the soft tissue nose, located with a tangent from nasion
	pronasale	The most anterior point on the nose tip
	columella-lobular junction	The point on the lower surface of the nose where the columella and the lobule of the nose meet
	columella[a]	A point between columella-lobular junction and subnasale
	subnasale	A point located at the junction between the lower border of the nose and the beginning of the upper lip
	cheekpoint	The most convex point on the soft tissue cheek
	soft tissue point A	The most posterior point on the upper lip between subnasale and labiale superius
	superior labial sulcus	The most concave point on the upper lip
	labiale superius	The point where the curve of the upper lip intersects with the tangential line from the subnasale to the upper lip
	upper lip	Most prominent point of the upper lip
	upper lip adjunct contour point[a]	A point between upper lip and stomion superius
	stomion superius	The most inferior point located on the upper lip
	stomion inferius	The most superior point located on the lower lip
	lower lip adjunct contour point[a]	A point between stomion inferius and lower lip
	lower lip	The most prominent point of the lower lip
	labiale inferius	The point where the curve of the lower lip intersects with the tangential line from Point B to the lower lip curve.
	inferior labial sulcus	The most concave point on the lower lip
	soft tissue point B	The most posterior point on concavity between labiale inferius and soft tissue pogonion
	soft-tissue protuberance menti	A point between soft tissue point B and pogonion at bony Protuberance menti level
	soft tissue pogonion	The most anterior point on the soft tissue chin
	soft tissue gnathion	The most inferior and anterior point on the soft tissue chin
	soft tissue menton	The most inferior point on the soft tissue chin
	menton adjunct contour point[a]	A point between soft-tissue menton and cervical point—anterior
	antero-cervical contour point[a]	A point between soft-tissue menton and cervical point—posterior
	cervical point	The innermost point between the submental area and the neck. Located at the intersection of the lines drawn tangent to the neck and submental area
	postero-cervical contour point[a]	A point between cervical point and soft-tissue terminal point
	soft tissue terminal point	The most inferior point of the image on the soft tissue neck

[a] Arbitrary landmarks to render a smooth line drawing of anatomic structures.

chosen as the base model. The original TabNet DNN architecture was modified using Python programming (Python Software Foundation, Wilmington, Delaware, USA). Different numbers of training epochs of 100 and 1000 were used to compare performance while exploring options, to save computational resources.[34] The PLS prediction model[12] was implemented using the open-source programming language, R.

Statistical Analysis

Prediction errors were calculated using Euclidean distances between actual growth and predicted outcomes for specific landmarks. To compare the prediction accuracy of PLS and AI, *t*-tests adjusted for multiple comparisons using the Bonferroni correction were used. Scatterplots with 95% confidence ellipses were created to represent the pattern of prediction errors visually.[35]

RESULTS

Table 2 presents the characteristics of the subjects at the time of growth observation. The average observation period was 8.5 years, with a mean starting age of 7.4. Ninety-four percent of subjects had radiographs taken more than five times, whereas about 27.3% had radiographs taken more than 10 times. When the proportion of malocclusion was considered, 51.5% of the subjects had Class I, 48.5% of patients had Class II malocclusions, and no patient presented a Class III molar relationship at initial examination.

The performance of the developed models was evaluated based on prediction errors. Among AI models developed under various conditions, the model with an early stopping condition at 1000 training epochs was chosen to calculate the AI prediction errors. On average, AI presented more accurate prediction with 0.61 mm smaller error than that of the PLS model. The average error for 45 hard tissue landmarks with the PLS prediction model

Figure 1. Longitudinal serial growth data source: the University of the Pacific Mathews Growth Study on the AAOF Craniofacial Growth Legacy Collection website, https://www.aaoflegacycollection.org/aaof_collection.html?id = UOPMathews (A). Four fiducial points and 78 cephalometric landmarks that were manually identified using a computer vision annotation tool (B).

was 1.87 mm, whereas the AI prediction error averaged 1.49 mm. For 32 soft tissue landmarks, the errors averaged 2.63 mm for PLS and 1.71 mm for AI. Among the 77 predicted landmarks, the AI-based prediction model showed better prediction accuracy for 60 landmarks (Table 3). The PLS-based prediction model was more accurate for 13 landmarks (Nasion, Porion, Orbitale, Basion, Articulare, Condylion, Ramus tip, Pterygomaxillary fissure, Pterygoid point, PNS, glabella, glabella-nasion

contour point, and cheekpoint). There was no statistical difference in four landmarks (Nasal bone tip, Key ridge contour smoothing point 1, soft-tissue nasion, inferior tip of nasal bone). Overall, both methods showed greater prediction errors for soft tissue and mandibular landmarks compared to hard tissue and maxillary landmarks. However, the AI method demonstrated a smaller increase in error for areas with more variability.

Table 2. Subject Characteristics (n = 33)

Variable	N (%)		Mean	SD	Min	Max
All subjects	33					
Female	21	(63.6%)				
Male	12	(36.4%)				
Angle's classification at the initial examination						
Class I	17	(51.5%)				
Class II	16	(48.5%)				
Class III	0	(0%)				
Age at initial examination (y)			7.37	1.45	3.58	11.17
Age after growth observation (y)			15.86	2.27	10.33	19.83
Growth observation period (y)			8.49	2.07	3.16	11.09
Number of serial cephalograms taken						
Five	2	(6.0%)				
Six	3	(9.1%)				
Seven	3	(9.1%)				
Eight	7	(21.2%)				
Nine	4	(12.1%)				
Ten	5	(15.2%)				
Eleven	5	(15.2%)				
Twelve	3	(9.1%)				
Thirteen	0	(0%)				
Fourteen	1	(3.0%)				

Table 3. Growth Prediction Errors (mm) Between the Partial Least Squares Regression (PLS) and Artificial Intelligence (AI). The Errors are Radial Errors or the Euclidean Distance (mm) Between the Actual Growth Position of a Given Landmark in the Lateral Cephalogram and its Predicted Position

| | PLS | | AI | | |
Growth Prediction Model	Mean	SD	Mean	SD	P value[b]
Hard tissue landmarks[a]					
Porion	1.39	0.82	1.99	1.11	<.0001
Orbitale	1.09	0.70	1.46	1.00	<.0001
Anterior nasal spine	1.63	0.96	1.53	0.80	<.0001
Posterior nasal spine	1.39	0.89	1.62	1.11	.0699
Point A	1.63	1.00	1.21	0.76	<.0001
Upper incisor root tip	1.78	1.07	1.41	0.89	<.0001
Upper incisal edge	2.15	1.23	1.54	0.99	<.0001
Lower incisal edge	2.16	1.14	1.48	0.86	<.0001
Lower incisor root tip	2.31	1.25	1.33	0.76	<.0001
Point B	2.44	1.28	1.22	0.74	<.0001
Protuberance menti	2.54	1.36	1.26	0.81	<.0001
Pogonion	2.68	1.44	1.32	0.87	<.0001
Gnathion	2.78	1.45	1.32	0.87	<.0001
Menton	2.78	1.43	1.34	0.83	<.0001
Gonion	2.44	1.40	1.69	0.97	<.0001
Articulare	1.20	0.65	1.50	0.88	<.0001
Condylion	1.22	0.66	1.68	0.92	<.0001
Pterygoid	0.93	0.58	1.65	0.94	<.0001
Basion	1.28	0.71	1.92	1.09	<.0001
Soft tissue landmarks					
glabella	2.80	2.06	3.45	2.15	<.0001
soft tissue nasion	1.64	0.91	1.65	0.89	1.0000
supranasal tip	1.95	1.12	1.52	0.90	<.0001
pronasale	2.01	1.16	1.47	0.85	<.0001
columella	1.95	1.14	1.22	0.74	<.0001
subnasale	1.88	1.11	1.22	0.72	<.0001
soft tissue point A	1.87	1.09	1.19	0.68	<.0001
superior labial sulcus	1.97	1.14	1.16	0.65	<.0001
labiale superius	2.09	1.16	1.22	0.73	<.0001
upper lip	2.21	1.20	1.28	0.77	<.0001
stomion superius	2.21	1.20	1.14	0.72	<.0001
stomion inferius	2.15	1.19	1.28	0.84	<.0001
lower lip	2.39	1.30	1.29	0.91	<.0001
labiale inferius	2.39	1.33	1.31	0.93	<.0001
soft tissue point B	2.54	1.45	1.51	0.97	<.0001
soft tissue protuberance menti	2.64	1.47	1.45	0.94	<.0001
soft tissue pogonion	2.81	1.53	1.70	1.07	<.0001
soft tissue gnathion	3.00	1.59	1.53	0.92	<.0001
soft tissue menton	3.03	1.55	1.64	0.93	<.0001

[a] Soft tissue landmarks are indicated by small case letters and hard tissue landmarks by capital letters. Among the 77 landmarks predicted, to succinctly provide the results, the prediction results of only 19 among 35 hard tissue landmarks and 19 among 32 soft tissue landmarks are listed in this table.

[b] Results from paired *t*-tests with Bonferroni correction.

The pattern of growth prediction errors for representative landmarks are shown in Figure 2. For hard tissue landmarks, PLS demonstrated better prediction accuracy in 13 landmarks including Nasion, Porion, Orbitale, Basion, and Condylion (Figure 2A). Generally, AI demonstrated significantly more accurate results than PLS, with AI model accuracy improving with more training epochs (Figure 2B). Among the soft

tissue landmarks, only glabella, glabella-nasion contour point, and cheek point were better predicted by PLS, whereas all other landmarks were more accurately predicted by AI (Figure 2C). Overall, PLS exhibited better or comparable prediction performance than the AI method in the upper part of the craniofacial structure, whereas AI outperformed PLS in the lower part of the craniofacial structure and in soft tissue.

Comparisons of actual growth and prediction outcomes based on the AI and PLS methods for real case examples are shown in Figure 3. The soft tissue landmarks from glabella to the terminal point of the lower neck were connected by applying the natural cubic spline function so that those landmarks could represent a smooth curve. Although both prediction results deviated from the actual profile after growth, AI-based predictions generally appeared to be closer to the actual profile.

DISCUSSION

Research on growth prediction has not been actively conducted for about two decades.[12] The complexity of predicting craniofacial growth with significant individual variation might have contributed to the lack of active research in this area. However, recent advancements in high-performance computing capable of handling the large computational demands of sophisticated algorithms have enabled the inclusion of large number of variables to develop more customized growth prediction models. This study used TabNet, one of the DNN algorithms,[33] to address challenges in growth prediction. The results indicated that growth prediction remains challenging, as larger error values were observed for specific landmarks in some subjects. Nevertheless, this study utilized the best available growth data and current technologic advancements to identify methods that were more effective in predicting individual growth patterns.

Overall, AI predicted growth more accurately than PLS. However, accuracy varied according to the landmarks being predicted. Regarding performance in predicting various landmarks, the PLS was comparable to AI in landmarks with minor growth variations, whereas AI was more accurate in areas with significant variability, consistent with findings from a previous study.[29] Among the 77 cephalometric landmarks, the PLS-based prediction demonstrated higher accuracy in 13 landmarks, primarily cranial base landmarks such as Nasion, Porion, and Basion. Additionally, PLS was more accurate in predicting Articulare and Condylion, aligning with expected outcomes since the positions of these mandibular landmarks are determined by the cranial base. Although statistically significant, the difference in errors between the two methods for these 13 landmarks averaged 0.45 mm, all being less than
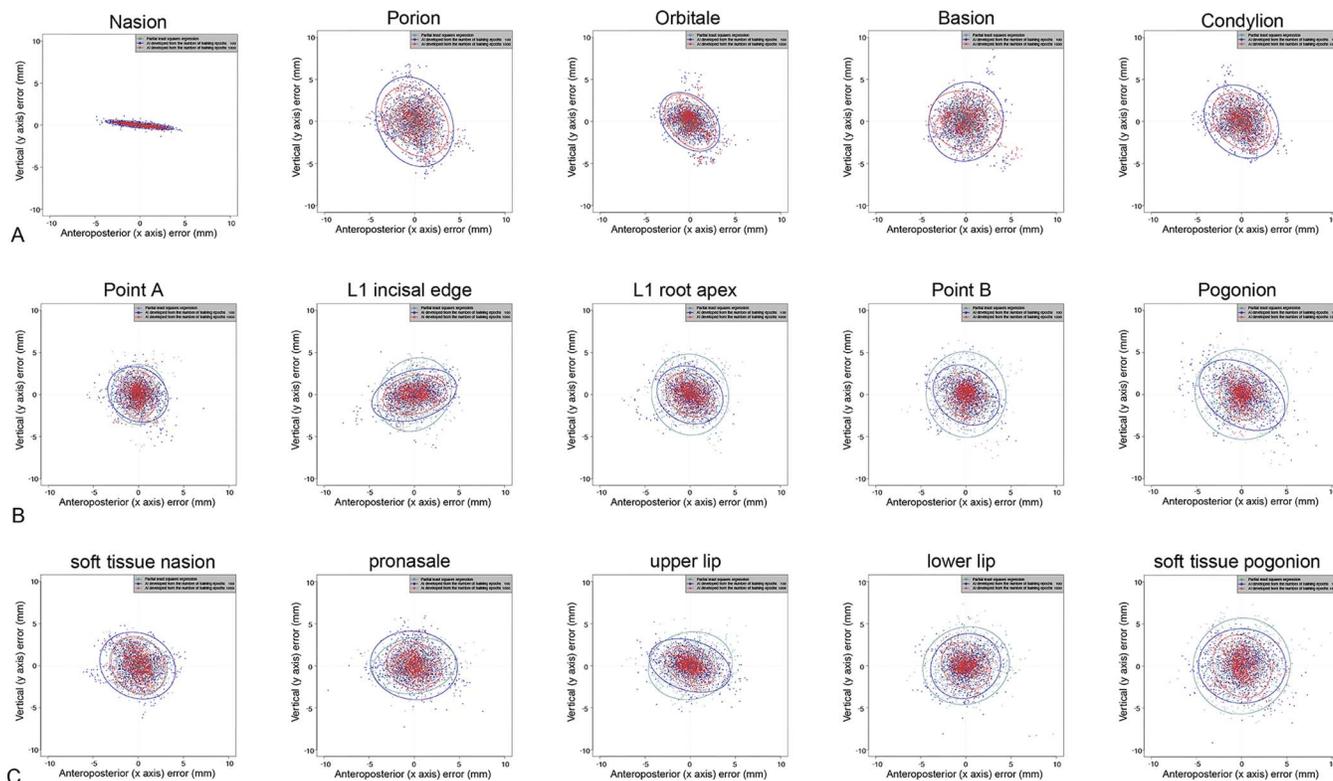
Figure 2. Scatter plots presenting errors and 95% confidence ellipses for the three prediction models. Green, PLS; Blue, AI developed from the number of training epochs 100; Red, AI developed from the number of training epochs 1000. (A) Hard tissue landmarks better predicted by the PLS model; (B) Hard tissue landmarks better predicted by the AI models; (C) Soft tissue landmarks.

1 mm. On the other hand, AI demonstrated greater prediction accuracy for 78% of the landmarks, including most of the landmarks in the lower part of the craniofacial structure and soft tissue. Soft tissue growth is more challenging to predict than skeletal growth, due to the influence of unpredictable factors such as posture and tonicity. This trend suggests that the development of AI growth prediction models can be beneficial

for areas with greater variability in which predictions are more challenging.

In terms of model training, higher training epochs led to more accurate predictions with AI. This study was set to 100 and 1000 epochs, although previous publications reported up to 10,000 sessions for AI training.[34] However, if 10,000 sessions had been included, a single computation could have taken several months.[34]
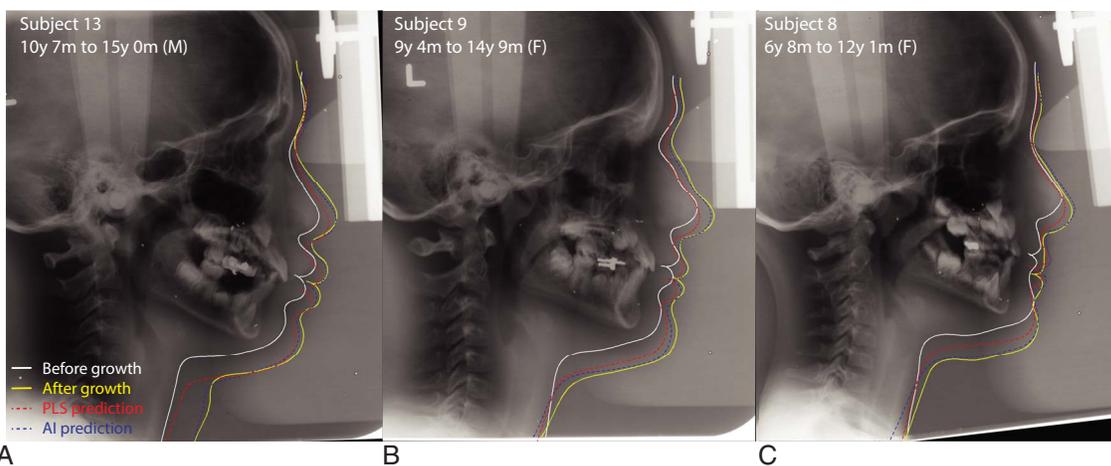

Figure 3. Example of profile predictions for patients included in the study. White, initial; Yellow, actual profile after growth; Red, PLS prediction; Blue AI. (A) a male subject from 10 y 7 mo to 15 y 0 mo; (B) a female subject from 9 y 4 mo to 14 y 9 mo; (C) a female subject from 6 y 8 mo to 12 y 1 mo.
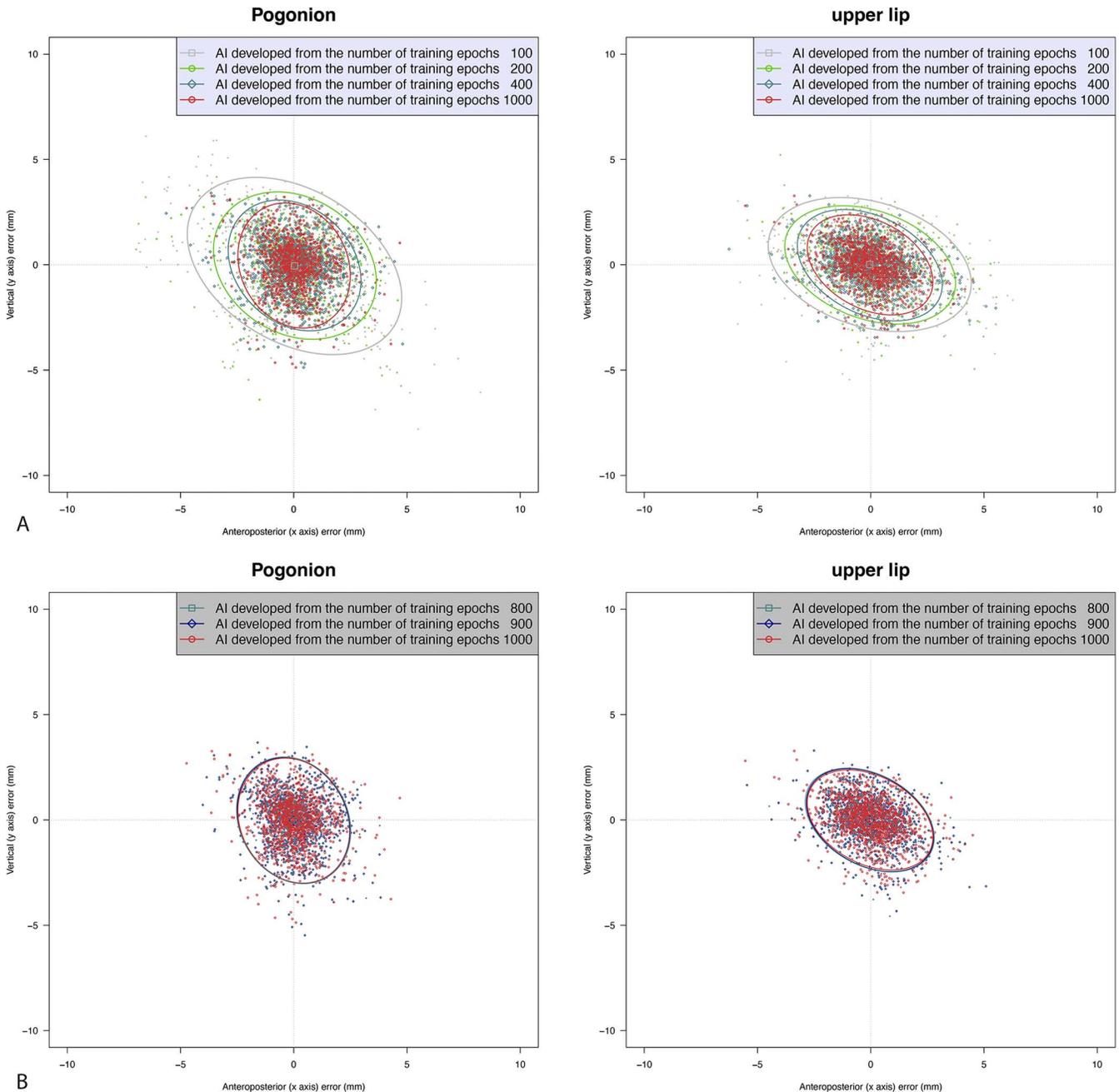
Figure 4. Scatter plots presenting errors and 95% confidence ellipses for prediction models from different training epochs. (A) prediction error decreases as the number of sessions increases from 100, 200, 400 to 1000; (B) prediction error does not decrease significantly beyond 800 sessions.

Figure 4 shows that, while the prediction error decreases as the number of sessions increases from 100 to 1000 (Figure 4A), the decrease in error was not significant beyond 800 to 1000 sessions (Figure 4B). As no substantial improvement in prediction performance was expected beyond 1000 sessions, this study chose 1000 epochs to achieve acceptable accuracy with a reasonable input of resources.

The errors from the models developed in this study were smaller than the errors of previously developed models.[29] In a prior study using longitudinal data from 410 subjects, yielding 679 pairs of before and after growth data, PLS exhibited errors 2.11 mm greater than those of the AI method, which showed an average error of 2.78 mm.[29] In this study, 33 subjects from longitudinal craniofacial growth records, resulting in 1257 pairs of before and after growth data, were included. PLS showed errors 0.61 mm greater than those of AI, which had an average error of 1.58 mm. Given that the inter-examiner error of cephalometric

tracing was reported to be 1.5 ± 1.5 mm,[24] the errors from this study were considered clinically acceptable. However, in some subjects or landmarks, the predicted landmarks still showed larger errors (Figure 2), partly due to the inherent nature of landmark location. These deviations may not be significant as long as the predicted landmark positions fall within the traced lines. In Figure 3A, despite a larger AI prediction error of 5.84 mm in soft tissue menton, the predicted lower mandibular soft tissue profile remained close to the actual profile.

Currently, collecting longitudinal growth data is challenging due to ethical concerns. Meanwhile, the AAOF Craniofacial Growth Legacy Collection compiles nine of the 11 recognized longitudinal collections of craniofacial growth records in the United States and Canada.[31] Presently, approximately 20,000 digital images from 842 subjects are available on the AAOF website, which could facilitate further development of growth prediction methods using AI. The use of growth collections in this paper involved a significantly larger number of pairs of growth data than previous studies, offering a better representation of the general population. However, this study only included Class I and Class II subjects, whereas growth of Class III subjects is expected to be more challenging to predict due to increased mandibular growth. This limitation can be addressed by incorporating additional growth collections. Additionally, the AAOF growth collections predominantly consist of subjects of European descent, and the effects of different ethnicities on prediction accuracy still needs to be explored.

CONCLUSIONS

- AI has shown to be an effective growth prediction method, with clinically acceptable prediction errors averaging 1.49 mm for 45 hard tissue landmarks, and 1.71 mm for 32 soft tissue landmarks.
- Among AI prediction models, those with increased training epochs showed improved prediction performance, but there was no significant improvement beyond 1000 epochs.
- AI generally outperformed PLS, particularly for landmarks in the lower part of the craniofacial structure and soft tissue, where uncertainty is considerable.

ACKNOWLEDGMENTS

REFERENCES

1. Proffit WR. The timing of early treatment: an overview. *Am J Orthod Dentofacial Orthop*. 2006;129:S47–49.
2. Moorrees CFA, Lebret L. The mesh diagram and cephalometrics. *Angle Orthod*. 1962;32:214–231.
3. Moorrees CF, uan Venrooij ME, Lebret LM, Glatky CG, Kent RL, Reed RB. New norms for the mesh diagram analysis. *Am J Orthod*. 1976;69:57–71.
4. Johnston LE. A simplified approach to prediction. *Am J Orthod*. 1975;67:253–257.
5. Harris JE, Johnston LE, Moyers RE. A cephalometric template: its construction and clinical significance. *Am J Orthod*. 1963;49:249–263.
6. Popovich F, Thompson GW. Craniofacial templates for orthodontic case analysis. *Am J Orthod*. 1977;71:406–420.
7. Broadbent BH, Golden WH. Bolton standards of dentofacial developmental growth. Mosby; 1975.
8. Ricketts RM. A principle of arcial growth of the mandible. *Angle Orthod*. 1972;42:368–386.
9. Ricketts RM. The value of cephalometrics and computerized technology. *Angle Orthod*. 1972;42:179–199.
10. Ricketts RM. Planning treatment on the basis of the facial pattern and an estimate of its growth. *Angle Orthod*. 1957;27:14–37.
11. Bishara SE. Facial and dental changes in adolescents and their clinical implications. *Angle Orthod*. 2000;70:471–483.
12. Moon JH, Kim MG, Hwang HW, Cho SJ, Donatelli RE, Lee SJ. Evaluation of an individualized facial growth prediction model based on the multivariate partial least squares method. *Angle Orthod*. 2022;92:705–713.
13. Hennessy RJ, Stringer CB. Geometric morphometric study of the regional variation of modern human craniofacial form. *Am J Phys Anthropol*. 2002;117:37–48.
14. Ursi WJ, Trotman CA, McNamara JA, Jr., Behrents RG. Sexual dimorphism in normal craniofacial growth. *Angle Orthod*. 1993;63:47–56.
15. Hersberger-Zurfluh MA, Papageorgiou SN, Motro M, Kantarci A, Will LA, Eliades T. Genetic and environmental components of vertical growth in mono- and dizygotic twins up to 15-18 years of age. *Angle Orthod*. 2021;91:384–390.
16. Abu Alhaija ES, Richardson A. Growth prediction in Class III patients using cluster and discriminant function analysis. *Eur J Orthod*. 2003;25:599–608.
17. Oh H, Knigge R, Hardin A, et al. Predicting adult facial type from mandibular landmark data at young ages. *Orthod Craniofac Res*. 2019;22 Suppl 1:154–162.
18. Suzuki A, Takahama Y. Parental data used to predict growth of craniofacial form. *Am J Orthod Dentofacial Orthop.* 1991;99:107–121.
19. Sherwood RJ, Oh HS, Valiathan M, et al. Bayesian approach to longitudinal craniofacial growth: The Craniofacial Growth Consortium Study. *Anat Rec (Hoboken)*. 2021;304:991–1019.
20. Rudolph DJ, White SE, Sinclair PM. Multivariate prediction of skeletal Class II growth. *Am J Orthod Dentofacial Orthop*. 1998;114:283–291.
21. Lee SJ, An H, Ahn SJ, Kim YH, Pak S, Lee JW. Early stature prediction method using stature growth parameters. *Ann Hum Biol*. 2008;35:509–517.
22. Lee YS, Lee SJ, An H, Donatelli RE, Kim SH. Do Class III patients have a different growth spurt than the general population? *Am J Orthod Dentofacial Orthop*. 2012;142:679–689.

23. Hwang HW, Moon JH, Kim MG, Donatelli RE, Lee SJ. Evaluation of automated cephalometric analysis based on the latest deep learning method. *Angle Orthod*. 2021; 91:329–335.

24. Hwang HW, Park JH, Moon JH, et al. Automated identification of cephalometric landmarks: Part 2-Might it be better than human? *Angle Orthod*. 2020;90:69–76.

25. Kim MG, Moon JH, Hwang HW, Cho SJ, Donatelli RE, Lee SJ. Evaluation of an automated superimposition method based on multiple landmarks for growing patients. *Angle Orthod*. 2022;92:226–232.

26. Moon JH, Hwang HW, Lee SJ. Evaluation of an automated superimposition method for computer-aided cephalometrics. *Angle Orthod*. 2020;90:390–396.

27. Moon JH, Hwang HW, Yu Y, Kim MG, Donatelli RE, Lee SJ. How much deep learning is enough for automatic identification to be reliable? *Angle Orthod*. 2020;90:823–830.

28. Park JH, Hwang HW, Moon JH, et al. Automated identification of cephalometric landmarks: Part 1-Comparisons between the latest deep-learning methods YOLOV3 and SSD. *Angle Orthod*. 2019;89:903–909.

29. Moon JH, Shin HK, Lee JM, et al. Comparison of individualized facial growth prediction models based on the partial least squares and artificial intelligence. *Angle Orthod*. 2024; 94:207–215.

30. Park JA, Moon JH, Lee JM, et al. Does artificial intelligence predict orthognathic surgical outcomes better than conventional linear regression methods? *Angle Orthod*. 2024;94(5):549–556.

31. Baumrind S, Curry S. American Association of Orthodontists Foundation Craniofacial Growth Legacy Collection: overview of a powerful tool for orthodontic research and teaching. *Am J Orthod Dentofacial Orthop*. 2015;148:217–225.

32. Donatelli RE, Lee SJ. How to test validity in orthodontic research: a mixed dentition analysis example. *Am J Orthod Dentofacial Orthop*. 2015;147:272–279.

33. Arik SÖ, Pfister T. TabNet: Attentive Interpretable Tabular Learning. *Proc Conf AAAI Artif Intell*. 2021;35:6679–6687.

34. Lee JM, Moon JH, Park JA, Kim JH, Lee SJ. Factors influencing the development of artificial intelligence in orthodontics. *Orthod Craniofac Res*. 2024.

35. Moon JH, Lee JM, Park JA, Suh H, Lee SJ. Reliability statistics every orthodontist should know. *Semin Orthod*. 2024;30:45–49.