Caspian Journal of Dental Research
e-ISSN: 2322-2395 p-ISSN: 2251-9890
Babol University Of Medical Sciences

Fully automated orthodontic photograph analysis by machine learning

Mahdi Soleiman Mezerji [1], Sedigheh Sheikhzadeh [2*], Maysam Mirzaie [3], Hemmat Gholinia [4]

1.Dental Student, Student Research Committee, Babol University of Medical Sciences, Babol, Iran.
2.Assistant Professor, Oral Health Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran.
3.Assistant Professor, Dental Materials Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran.
4.Msc in Statistics, Social Determinants of Health Research Center, Health Research Institute, Babol University of Medical Sciences, Babol, Iran.

Article type
Research Paper

Received: 28 Apr 2023
Revised: 13 Jun 2023
Accepted: 17 Jun 2023
Pub. online: 3 Sept 2023

ABSTRACT

Introduction: The craniofacial anthropometric ratios are very useful in sciences such as dentistry, maxillofacial surgery, developmental studies and plastic surgery. The manual method of analyzing facial photographs requires a lot of time and precision. The aim of this study was to introduce an application tool that fully automates the analysis of facial photographs and compare it with the manual method.

Materials & Methods: In this cross-sectional study, the database consisted of 395 profile photographs, 271 frontal photographs in smile and 346 frontal photographs at rest. A two-stage fully convolutional network architecture was used for landmark detection. Two methods of manual and automatic analysis were compared in the measurement of 8 variables, including buccal corridor space, ratio of the height of the middle to the lower third of the face, total facial convexity angle, facial convexity angle, nasofacial angle, mentolabial angle, and nasofrontal angle. The agreement between the two methods was evaluated using the paired T-test and intraclass correlation coefficient (ICC). A value of $p<0.05$ was considered significant.

Results: For total facial convexity ($P=0.005$), nasofacial ($P=0.001$), and nasolabial ($p=0.02$) angles, the difference between the two methods was significant. However, no significant difference was found between the two methods for facial convexity, mentolabial, nasofrontal, buccal corridor space, and the ratio of the height of the middle to the lower third of the face no significant difference was observed between the two methods. The ICC for all variables was found to be greater than 0.69 except for the nasolabial angle. For most of the measured variables, the accuracy of the automatic method was similar to that of the manual method.

Conclusion: Machine learning has the potential to be used in clinical soft tissue analysis. It offers the ability to perform reliable and repeatable analyses on large image datasets.

Keywords: Orthodontics, Face, Photography, Machine Learning

Cite this article: Soleiman Mezerji M, Sheikhzadeh S, Mirzaie M, Gholinia H. Fully automated orthodontic photograph analysis by machine learning. Caspian J Dent Res 2023; 12: 70-81.

© The Author(s).
Publisher: Babol University of Medical Sciences

* **Corresponding Author:** Sedigheh Sheikhzadeh, Department of Orthodontics, Faculty of Dentistry, Babol University of Medical Sciences, Babol, Iran.
Tel: +981132297951
E-mail: elfsh@yahoo.com

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

[Downloaded from cjdr.ir on 2026-11-29]

[DOR: 20.1001.1.22519890.2023.12.2.3.8]

[DOI: 10.22088/cjdr.12.2.70]

Introduction

Facial anthropometry is a term used to describe the measurement of facial features and the method used to capture these features. [1] Taking facial photographs is relatively simple. The equipment is inexpensive and does not require a very high level of technical expertise. Measurement of facial ratios requires a large amount of time by an expert. The following factors cause problems in facial soft tissue analysis:

1. It is relatively difficult to take photographs in NHP position,
2. The variation of soft tissue morphology is very patient dependent,
3. High accuracy is required
4. Landmark detection is subjective and different observers may have different views

Artificial intelligence can automate these processes. In recent years, there has been a lot of interest in deep learning. The best known algorithm among the various deep learning models is the Convolutional Neural Network (CNN), a class of artificial neural networks that is a dominant method in computer vision tasks. CNNs have achieved expert-level performance in several domains. Facial landmark detection has been studied for decades. Several neural network-based approaches have been proposed for landmark detection, especially CNN-based approaches. [2] They are widely used in medical and non-medical sciences. [3]

Although efforts have been made to speed up soft tissue analysis in dentistry, landmark localization in most softwares still requires human interaction and is time-consuming. Most of the softwares developed so far are not able to determine all variables for photographic analysis. For example, none of the applications is able to measure the buccal corridor space. [4, 5] Some studies have been conducted to the measure aspect ratio using non-dental software [4] also, studies have been done to automatically identify landmarks in 3D images.[5-8] But so far, no study has been conducted to fully automatically analyze soft tissue in 2D photographs used in dentistry.

Loveday et al. developed a software for the analysis of facial photographs. First, the images taken with the digital camera are entered into the software. Then, the position of the landmarks is determined manually and with human interaction. The software automatically measures angles and lengths. [4] The value of this study was that the landmark detection was fully automatic with no human intervention. Further analysis was also performed automatically.

Materials & Methods

This study was approved by the Ethics Committee of Babol University of Medical Sciences (ethical number: IR.MUBABOL.REC.1400.203). For this cross-sectional study, orthodontic extra oral photographs of Babol University of Medical Science patients were collected, including profile view, frontal view at rest, and frontal view during posed social smiling (Figure1). The minimum sample size was calculated using the following formula and considering r=0.5, n=30 samples.

$$n \geq \frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2}{\left(\frac{1}{2} ln \frac{1+r}{1-r}\right)^2} + 3$$

$$n \geq \frac{\left(Z_{1-\frac{0.05}{2}} + Z_{1-0.2}\right)^2}{\left(\frac{1}{2}\ ln\frac{1+0.5}{1-0.5}\right)^2} + 3 \geq \frac{(1.96+0.84)^2}{0.3} + 3 \geq 29.1 \Rightarrow n \geq 30$$

$$\alpha = 0/05\ , \beta = 0/20\ , r = 0/5$$

All photographs were taken in the Natural Head Position (NHP). Images were taken with different sensors, such as smartphone cameras or digital cameras, were used to capture the images, ensuring a minimum resolution of 256 x 256 pixels. Any low-resolution images were excluded from the study. Datasets were created without restrictions on image background, color, size, age, gender, or whether craniofacial or dental surgery was performed. This was done to assess the ability of the system to identify images with different characteristics. Two separate datasets were created for training the model and evaluating its performance. The "Training" dataset consisted of 344 profile photographs, 299 frontal photographs taken with the lips at rest, and 239 frontal photographs captured in a posed social smile. The "Test" dataset, on the other hand, included 51 profile photographs, 47 frontal photographs taken in a resting pose, and 32 frontal photographs taken in a social smile pose. These datasets were carefully selected to ensure that all images had standardized features. It is important to note that the conditions were the same for both methods (Figure1). We made every effort to create all images with standardized features. Rest assured that the conditions remain the same for both methods. It is worth noting that minor deviations from the NHP standards will not affect the results.



Figure1. Photographs of profile, frontal lips at rest and frontal posed smile

The ground truth is defined as the exact coordinates of the landmarks. To obtain the ground truth, Computer Vision Annotation Tool (CVAT) was used in the current study. The positions of the landmarks were manually recorded using this method (see Figure 2). Specifically, 10 landmarks were identified in profile images, 4 landmarks in images of frontal posed smile, and 3 landmarks in images of lips at rest (Table 1). In the present study, the position of each landmark was recorded by the computer as a series of pixel coordinates (x, y). These landmarks were selected based on the work of Jorgensen and anthropometric analysis variables derived from the study of Farkass. [9]

[Downloaded from cjdr.ir on 2026-11-29]

[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

Table1. Definition of soft tissue landmarks

	Image	Landmark	Definition
1	Profile	Soft tissue Glabella (G')	The most anterior midpoint on the fronto-orbital soft tissue contour
2	Profile	Soft tissue Pogonion (pg)	The most anterior midpoint of the chin
3	Profile	Subnasale (Sn)	The midpoint of the nasolabial fold is located between the columella and the upper lip septum. It can be described as the meeting point of the lower border of the nasal septum and the surface of the upper lip
4	Profile	Soft tissue nasion (n)	The midpoint on soft tissue contour of the base of the nasal root
5	Profile	Columella (Cm)	The breaking point of the line drawn from Sn to the lower part of the crest of the columella
6	Profile	Labiale superior (Ls)	The middle point of outer edge of upper lip vermilion
7	Profile	Pronasale (prn)	The most anterior point of nose tip
8	Profile	Labiale inferior (Li)	The middle point of vermilion lower lip outline
9	Profile	Sublabiale	most posterior point of labiomental fold
10	Profile	Nasal dorsum	The most prominent nasal point on bony dorsum
11	Frontal at rest	Subnasale (Sn)	The meeting place of the lower border of nasal septum and surface of upper lip
12	Frontal at rest	Menton	The most inferior point on the soft tissue chin
13	Frontal at rest	Glabella	The most prominent point in the middle between the eyebrows
14	Frontal posed smile	Left teeth	The most distal point of most posterior tooth that can be seen on the left side of mouth
15	Frontal posed smile	Right teeth	The most distal point of most posterior tooth that can be seen on the right side of the mouth
16	Frontal posed smile	Left external commissure	External canthus of the left side of lip
17	Frontal posed smile	right external commissure	External canthus of the left side of lip

The anthropometric measures given in Table 1 were calculated for each of the images and considered as ground truth.

[Downloaded from cjdr.ir on 2026-11-29]

[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]



Figure 2. Marked landmarks in each of the profile images, frontal at rest and frontal posed smile

Development of a fully automatic algorithm for landmark identification

Since the resolution of the images is too high, they were all converted to square images with a resolution of 256 x 256 pixels. A model was developed for each landmark using the TensorFlow framework in Python .A two-stage Fully Convolutional Network (FCN) was trained using a Training dataset for landmark detection. Based on the Training dataset completed with Cartesian coordinates (x, y) for each landmark, a deep network was built to predict a different heatmap for each key point to be found. Direct mapping of images to Cartesian coordinates is very complex, and thus any model with sufficient accuracy is also likely to overfit and not generalize well to new data. For this reason, the coordinates of each landmark were first converted to a Gaussian heatmap (Figure 3) so that the model would predict the heatmap instead of the Cartesian coordinates (Figure 4).



Figure 3. Converting Sublabiale landmark to heatmap

[Downloaded from cjdr.ir on 2026-11-29]

[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

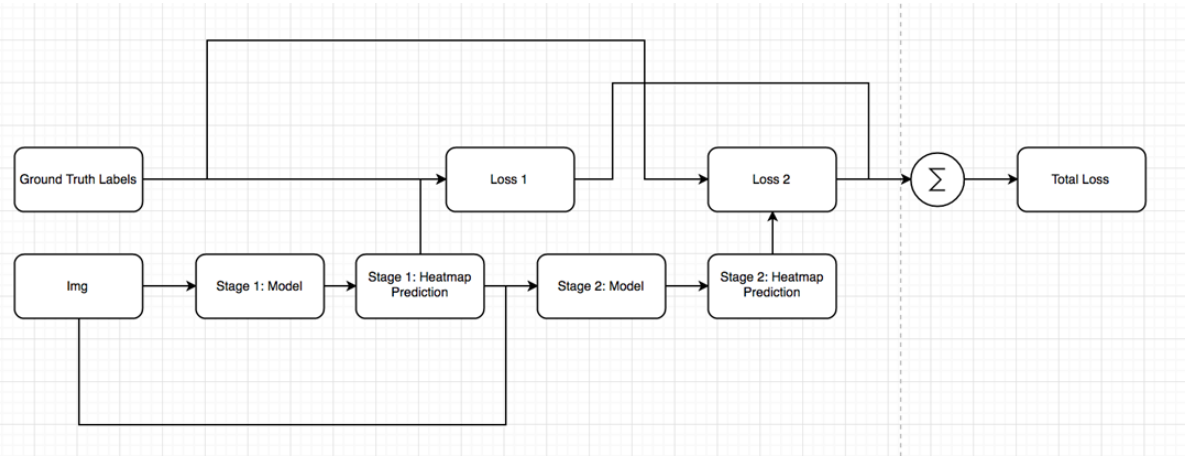


Figure 4. Flowchart of artificial intelligence model prediction steps

Photograph analysis

An algorithm was developed to calculate the required ratios and angles from the coordinates of the landmarks. The angles and ratios are described in Table 2 and in Figures 5-7. The measurements were selected based on Anicy's research. [10]

Table 2. Defenition of variables

	Image	Variable	Definition	*Unit of measurement*
1	Frontal at rest	The height ratio of the of the middle third to the lower third of the face	The height ratio of the of the middle third to the lower third of the face	-
2	Frontal at smile	Buccal corridor space	The ratio of the width of the empty spaces between the most distal point of last tooth and the external commissure to the distance between two external commissures	-
3	Profile	Total facial convexity angle	Angle between Glabella, Pronasal and Pogonion	Degree
4	Profile	Facial convexity angle	Angle between Glabella, subnasal and Pogonion	Degree
5	Profile	Nasofacial angle	The angle between the Glabella-Pogonion and nasion-pronasal lines	Degree
6	Profile	Nasolabial angle	Angle between columella, subnasal and superior labial	Degree
7	Profile	Mentolabial angle	Angle between Pogonion, supramental and inferior labial	Degree
8	Profile	Nasofrontal angle	Angle between nasal dorsum, nasion and Glabella	Degree

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

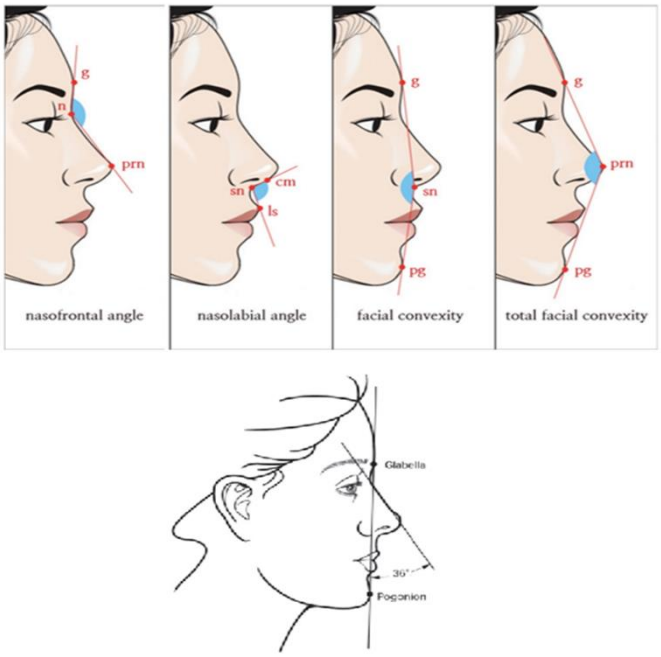


Figure 5. Angles used in soft tissue analysis of profile photography

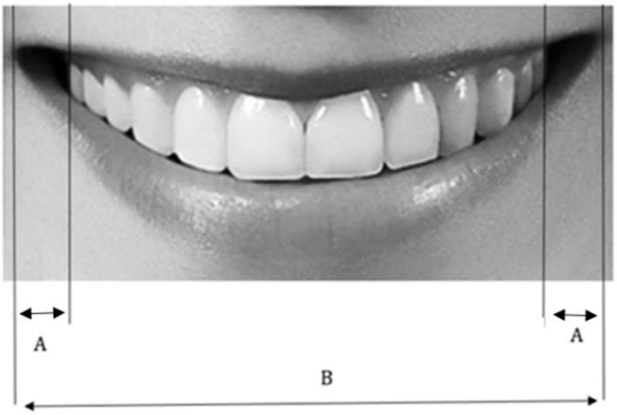


Figure 6. Buccal corridor space measurement

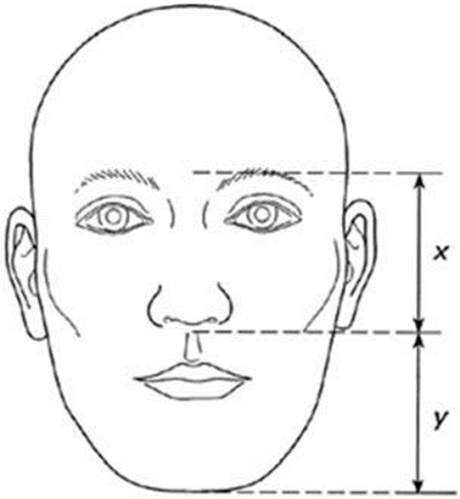


Figure 7. Ratio of the middle to the lower third of the face

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

Evaluation

The former measurements for the test dataset were performed manually by an expert; then the images were analyzed automatically by the developed application. Descriptive statistics of the analyzed parameters were presented as mean and standard deviation. The paired T-test was used to determine the significance of the difference between the two methods for each of the mentioned variables. The agreement between the measurements performed with the manual and automatic methods was evaluated using the intraclass correlation coefficient (ICC). All analyses were performed using SPSS 26 and at significance level α=0.05.

Results

The descriptive and statistical results related to the comparison of the manual and automatic measurements are presented in Table 3. For the frontal images with lips at rest, there was no significant difference between the two methods for the height ratio between the middle and lower parts of the face (P=0.24). The average differences were 0.01 and high agreement was found between the two methods (ICC = 0.9). For frontal posed smile images, there was no significant difference between the two methods for the variable buccal corridor space (P=0.25). The average of the differences was 0.013. Relatively, good agreement was obtained between two automatic and manual methods (ICC = 0.73).

No significant differences were found in the profile images between the facial convexity, mentolabial, and nasofrontal angles (P=0.98, P=0.64, and P=0.54, respectively). The ICCs for these angles were 0.71, 0.69, and 0.72, respectively. For the measurement of the nasofacial angle, although the difference between the two methods was statistically significant, the average of the differences was 1.84, and the ICC was 0.72. For the measurement of the nasolabial angle (ICC=0.29), little agreement was found two methods. The correlation between the two methods in measuring each variable can be seen in Figure 8.

Table3. Descriptive data of parameters measured by manual and automatic methods

	Image	Variable	Sample size	Analysis method	Average	Standard deviation	ICC	Average of differences	**P value
1	Frontal at rest	Height ratio of middle part to lower part	47	Manual	0.99	0.13	0.9	0.01	0.24
				Automated	1	0.14			
2	Frontal posed smile	Buccal corridor space	32	Manual	0.28	0.08	0.73	0.013	0.25
				Automated	0.27	0.09			
3	Profile	Total facial convexity angle	51	Manual	140.1	5.04	0.91	0.99	0.005*
				Automated	141.1	4.99			
4	Profile	Facial convexity angle	51	Manual	169.88	7.72	0.71	0.52	0.54
				Automated	170.4	6.87			
5	Profile	Nasofacial angle	51	Manual	31.05	3.85	0.72	1.84	0.001*
				Automated	21.29	4.55			

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

	Image	Variable	Sample size	Analysis method	Average	Standard deviation	ICC	Average of differences	**P value
6	Profile	Nasolabial angle	51	Manual	104.97	11.13	0.29	5.5	0.024*
				Automated	110.47	14			
7	Profile	Mentolabial angle	51	Manual	128/89	14.23	0.69	0.75	0.64
				Automated	129/64	11.37			
8	Profile	Nasofrontal angle	51	Manual	141.32	6.97	0.72	0.078	0.98
				Automated	141.34	9.45			

** Paired T-Test
*Significant at 0.05 level

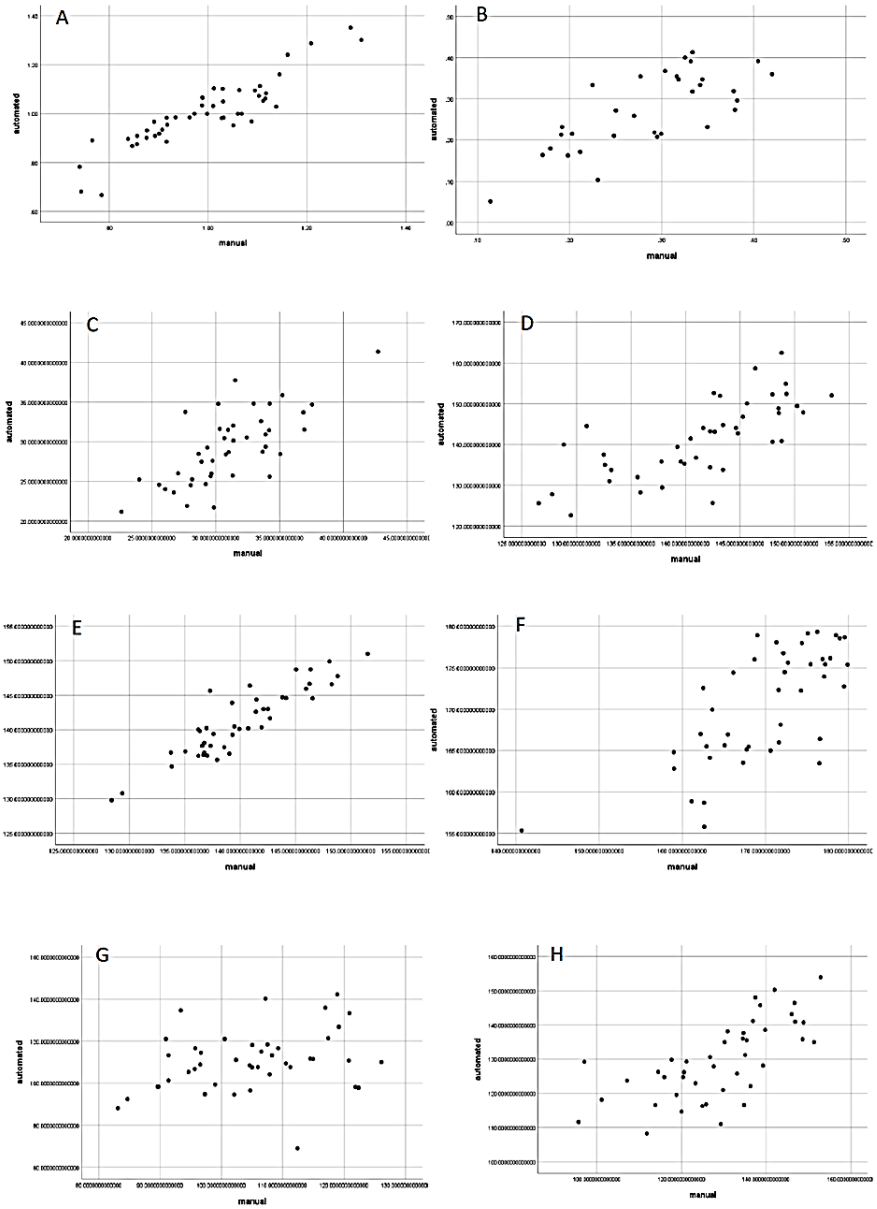


Figure 8. Comparison of two methods of measuring (A) height ratio of middle part to lower part of the face ,(B) buccal corridor space, (C) nasofacial angle, (D) nasofrontal angle, (E) Total facial convexity angle, (F) facial convexity angle, (G) nasolabial angle and (H) mentolabial angle

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

Discussion

In the present study, a high correlation was found between the manual and automatic analysis methods for most of the measured variables. Although there was a statistically significant difference between the two methods in the measurement of the nasofacial angle, the average difference was only 1.84 degrees and the ICC was 0.72. Based on these results, it is safe to conclude that the ability of the automatic method is considered clinically acceptable.

It seems that the relatively poor performance of our system in predicting the nasolabial angle was due to two main factors. First, there was instability in the detection of the subnasale and columella landmarks. This inconsistency in the detection of these landmarks could affect the accuracy of the measurements. Second, the effects of imaging errors could also play a role in the performance of the system. If errors occur during imaging that affect the accurate positioning of these landmarks, this could lead to inaccurate predictions of the nasolabial angle. .

Unlike some studies by Oghenemavwe et al. [4] and Ozkul et al., [11] which used a semiautomated method, a fully automated method was used in the current study. In a study conducted by Loveday and Ozkul et al., the systems they developed follow a two-step process. First, the landmarks are defined manually and then the soft tissue analysis variables are calculated automatically. [11] In some studies, such as Hong [5] and Asi [6], although the landmark identification was performed automatically using artificial intelligence methods, the angles and aspect ratios were not calculated. In the present study, both steps, landmark identification and variable measurement, were performed automatically. This makes the app more practical for clinical use. Ozkul et al. utilized 2D images in basic, nasal, frontal, and profile views, whereas Oghenemavwe et al. focused on profile and frontal images. [11, 4]

In the present study, three types of extra oral photographs were used, namely frontal images in the resting state, frontal images in the smiling state, and profile images. This approach allows for a comparison of analysis precision between the different types of images, which can provide valuable insight into the effects of facial expression on the variables being measured. The current study found an average ICC value of 0.67 for profile image variables, while it was 0.81 for frontal images. The results showed that the machine was better able to analyze frontal images than profile images. This difference can be attributed to the low capability of the two-stage FCN architecture in localizing boundary landmarks.

In the study by Salvarzi et al., facial dimensions were measured using Digimizer software. They conducted linear measurements such as the width of the face, the width of the nose, the length of the lips, and the length of the face, to name a few. They measured the variables once with a meter and caliper directly on the face and once in the Digimizer software environment with a digital ruler. The ICC of the measured variables ranged from 0.56 to 0.94. [12] In the present study, the obtained ICC ranged from 0.29 to 0.91. Unlike the aforementioned study, angular and proportional variables were used instead of linear criteria due to the different scale of the images.

In the ongoing study, unlike the studies by Ozkul et al., Salvarzi et al., and Oghenemavwe et al., there were no restrictions on the camera sensor, lens, or image background. Additionally, the images used in the present study did not have to have the same resolution. [11, 12, 4] The wide variety of the training dataset leads to higher performance in clinical diagnosis.

Asi et al. developed a model using the Haar Cascade Classifier to localize facial landmarks. [6] The Haar classifier developed by Viola and Jones is capable of detecting very small facial features. [13] But

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

despite the high speed, this architecture has low accuracy. Therefore, a two-stage FCN was used in the present study. Although this method has a lower speed, its accuracy is higher. This may be the reason why people want to participate in studies to create large databases. This study may be the beginning of further studies focusing on creating a facial anthropometric database for analytical purposes. The current study assumes that the developed application will be widely used in clinics in the form of a suitable user interface. However, since conditions are not equal in terms of technology and environmental factors, it would have been better to plan ahead.

Conclusion

An artificial intelligence system that utilizes deep learning with proper training models can successfully perform orthodontic analysis of facial photographs. We expect that this fully automated cephalometric analysis algorithm can be widely used in various medical environments to save time and effort in diagnosis.

Funding

This manuscript was derived from a research project (Grant No: 140012527), supported and funded by Babol University of Medical Sciences.

Conflicts of Interest

There is no conflict of interest to declare.

Author's Contribution

Mahdi Soleiman Mezerji developed the original idea and protocol, summarized the data, drafted the manuscript, and edited the article. Hemmat Gholinia analyzed the data. The study was supervised by Sedigheh Sheikhzadeh and Maysam Mirzaei.

References

1. Douglas TS. Image processing for craniofacial landmark identification and measurement: a review of photogrammetry and cephalometry. Comput Med Imaging Graph 2004;28:401-9.
2. Hsu CF, Lin CC, Ting YC, Lei CL, Chen KT. A detailed look at CNN-based approaches in facial landmark detection. arXiv:2005.08649. Available from:https://arxiv.org/pdf/2005.08649.pdf
3. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018;9:611-29.
4. Oghenemavwe EL, Fawehinmi HB, Daenwin LT. A Software Tool for Facial Analysis. Res. J Appl Sci Eng Technol 2012; 4:551–6.
5. Hong C, Choi K, Kachroo Y, Kwon T, Nguyen A, McComb R, et al. Evaluation of the 3dMDface system as a tool for soft tissue analysis. Orthod Craniofac Res 2017;20 (Suppl1):119–24.
6. Asi SM, Ismail NH, Ahmad R, Ramlan EI, Rahman ZA. Automatic Craniofacial Anthropometry Landmarks Detection and Measurements for the Orbital Region. Procedia Comput Sci 2014; 42:372–7.
7. Baksi S, Freezer S, Matsumoto T, Dreyer C. Accuracy of an automated method of 3D soft tissue landmark detection. Eur J Orthod 2021; 43:622-30.

[Downloaded from cjdr.ir on 2026-11-29]
[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]

8. Deli R, Di Gioia E, Galantucci LM, Percoco G. Automated landmark extraction for orthodontic measurement of faces using the 3-camera photogrammetry methodology. J Craniofac Surg 2010;21:87–93.
9. Farkas LG, Deutsch CK. Anthropometric determination of craniofacial morphology. Am J Med Genet 1996;65:1–4.
10. Anicy-Milosevicy S, Lapter-Varga M, Slaj M. Analysis of the soft tissue facial profile by means of angular measurements. Eur J Orthod 2008;30:135–40.
11. Ozkul T, Ozkul MH, Akhtar R, Al-Kaabi F, Jumaia T. A Software Tool for Measurement of Facial Parameters. Open Chem Biomed Methods J 2009; 2: 69-74.
12. Salvarzi E, Choobineh A, Jahangiri M, Keshavarzi S. Application of Digimizer Image Analysis Software In Facial Anthropometry. Iran J Ergon 2020; 8: 61-71.[In Persian]
13. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001; 8-14 Dec. 2001; Cambridge, MA, USA: Mitsubishi Electric Research Laboratories, Inc. 2004; 12:511-8.

[Downloaded from cjdr.ir on 2026-11-29]

[DOR: 20.1001.1.22519890.2023.12.2.3.8]
[DOI: 10.22088/cjdr.12.2.70]