Al-Rafidain J Med Sci. 2025;8(2):71-75.
DOI: https://doi.org/10.54133/ajms.v8i2.1839

AJMS
Online ISSN (2789-3219)

Research Article

Assessing the Accuracy and Reliability of Fully AI-Driven Cephalometric Analysis in Comparison to Digital Manual Methods

Noor Ali Alhamdani[1]*, Brwa Mahdi Khoshnaw[2], Zana Qadir Omer[2]
[1]*Department of Orthodontics, College of Dentistry, Al-Kitab University, Kirkuk 36015, Iraq;* [2]*Department of Orthodontics, College of Dentistry, Hawler Medical University, Erbil, Iraq*
Received: 5 March 2025; Revised: 8 April 2025; Accepted: 12 April 2025

Abstract

Background: Artificial intelligence (AI) has invaded radiographic analysis in a massive way. Besides saving time and effort during decision-making and treatment planning, AI-assisted cephalometric analysis must be reliable, reproducible, accurate, and user-friendly. ***Objective***: To assess fully AI-driven cephalometric analysis. ***Methods***: 47 lateral cephalometric radiographs were used for comparing the accuracy of AI-operated and manually operated skeletal and dental cephalometric analysis. Both dental and skeletal analyses were digitally performed using web-based platforms (WebCeph™, Cephio, and Ceppro DDH Inc., Korea). SPSS was used for statistical analysis with paired t-test and intra-class correlation. ***Results***: There were statistical differences between AI landmarking and manual landmarking regarding SNA, ANB, FH to AB, U1 to FH, U1 to SN, U1 to UOP, interincisal angle, U1 to NA (mm), and U1 to NA (deg). The intraclass correlation coefficient (ICC) data showed that the two sets of measurements were very consistent for most readings. ***Conclusions***: Even though AI provides strong reliability and agreement between methods, the significant difference indicates that AI-operated and manually operated cephalometric analysis may not be interchangeable despite their consistency. AI-based analyses primarily function as assistant tools, and orthodontists need to make judgments or adjustments before making decisions.

Keywords: Artificial intelligence, Cephalometric analysis, Webceph.

تقييم دقة وموثوقية تحليل قياس الرأس القائم على الذكاء الاصطناعي بالكامل مقارنة بالطرق اليدوية الرقمية

الخلاصة
الخلفية: اقتحم الذكاء الاصطناعي (AI) مجال تحليل الصور الشعاعية بشكل واسع. بالإضافة إلى توفير الوقت والجهد أثناء اتخاذ القرارات والتخطيط للعلاج، يجب أن يكون تحليل السيفالومتري المدعوم بالذكاء الاصطناعي موثوقًا، قابلاً لإعادة الإنتاج، دقيقًا وسهل الاستخدام. **الهدف**: يهدف هذا البحث إلى تقييم التحليل السيفالومتري القائم بالكامل على الذكاء الاصطناعي. **الطرائق:** تم استخدام 47 صورة شعاعية سيفالومترية جانبية لمقارنة دقة التحليل السيفالومتري الهيكلي والسنّي بين التحليل اليدوي والمدعوم بالذكاء الاصطناعي. تم إجراء كل من التحليلين الهيكلي والسنّي رقميًا باستخدام منصة إلكترونية (WebCephTM, Cephio, و Ceppro DDH Inc, تم استخدام برنامج SPSS للتحليل الإحصائي مع اختبار t للعينات المزدوجة ومعامل الارتباط داخل الصفوف(ICC). **النتائج:** أظهرت النتائج وجود فروق ذات دلالة إحصائية بين تحديد المعالم بواسطة الذكاء الاصطناعي والتحديد اليدوي، وذلك فيما يتعلق بـSNA ، ANB، الزاوية بين FH وAB، الزاوية بين U1 وFH، الزاوية بين U1 وSN، الزاوية بين U1 وUOP، الزاوية القاطعية بين الأسنان، المسافة بين U1 و NA(مم)، والزاوية بين U1 و NA(درجة). أظهرت بيانات معامل الارتباط داخل الصفوف (ICC) أن مجموعتي القياسات كانتا متسقتين بشكل كبير في معظم القراءات.**الاستنتاج:** على الرغم من أن الذكاء الاصطناعي يوفر موثوقية قوية واتفاقًا بين الطرق المختلفة، فإن الفروقات المعنوية تشير إلى أن التحليل السيفالومتري المدعوم بالذكاء الاصطناعي والتحليل اليدوي قد لا يكونان قابلين للتبادل بالكامل على الرغم من اتساقهما. تعمل التحليلات القائمة على الذكاء الاصطناعي بشكل أساسي كأدوات مساعدة، ولا بد لأخصائيي تقويم الأسنان من اتخاذ قرارات أو إجراء تعديلات قبل اتخاذ القرار النهائي.

* ***Corresponding author***: Noor A. Alhamdani, Department of Orthodontics, College of Dentistry, Al-Kitab University, Kirkuk 36015, Iraq; Email: noor.ali@uoalkitab.edu.iq
Article citation: Alhamdani NA, Khoshnaw BM, Omer ZQ. Assessing the Accuracy and Reliability of Fully AI-Driven Cephalometric Analysis in Comparison to Digital Manual Methods. *Al-Rafidain J Med Sci*. 2025;8(2):71-75. doi: https://doi.org/10.54133/ajms.v8i2.1839

INTRODUCTION

To create a precise treatment plan, the orthodontist needs to thoroughly diagnose the orthodontic case. Typically, by collecting a variety of data, such as study models, radiographs, and pictures, which require further examination and analysis. The cephalometric analysis of the 2D lateral cephalometric radiography is crucial in skeletal malocclusion cases [1]. Cephalometric analysis is a vital part of orthodontic diagnosis and treatment planning, requiring the exact identification of anatomical landmarks. Traditionally, these landmarks have been recognized manually, which is time-consuming and sensitive to variability between operators [2]. Recent breakthroughs in artificial intelligence (AI) have introduced automated systems for cephalometric landmark detection, with increased efficiency and consistency [3]. A subfield of artificial intelligence called machine learning is concerned with developing models and algorithms that enable computers to acquire information from data and generate judgments or predictions without external guidance [4]. A form of machine learning called deep learning employs neural networks that are modeled as the human brain to automatically identify patterns and characteristics in data. This allows computers to independently make deeper predictions or choices [5]. Despite these advances, research into the reliability of AI-generated landmarks in clinical practice continues. Studies have shown that AI systems can recognize cephalometric landmarks with excellent accuracy. However, the automated cephalometric measurements were considered clinically accepted but with manual adjustment of the landmark position when compared to the conventional hand tracing [6]. This research aimed to assess the performance of AI-driven web-based cephalometric analysis by comparing the accuracy of AI-predicted landmarks with those manually located by qualified specialists by comparing some dental and skeletal cephalometric analyses.

METHODS

Study design and setting

This retrospective study analyzed 47 lateral cephalometric radiographs from patients aged 12–40 years. All the cephalometric radiographs were obtained from one radiology center using one cephalometric device (PaX-i3D Smart, D-052SB, Toshiba, Korea) with the following specifications: magnification factor 1.14 constant, pixel size: 127 µm, kVp: 50-99 kV, mA: 4-16, and exposure time: 12 s. 75 radiographs were reviewed, of which 47 met the eligibility criteria.

Inclusion criteria

Clear radiograph with no artifact or distortion. Class I, II, and III skeletal relations. Radiograph with standardized head position.

Exclusion criteria

Cephalograms with blurred landmarks. Cephalograms with significant double borders of the mandible. Individuals with craniofacial anomalies, asymmetries, or a history of craniofacial surgery. Cases with dental crowns, bridges, implants, and edentulousness.

Radiographic evaluation

Google Chrome was used as a typical web browser (Google LLC, California, USA). Digital pictures of cephalograms were uploaded to the appropriate patient profiles, which were generated in the Webceph platform using the newly formed account. The first step in the landmark detection was radiographic image calibration by fitting the calibration ruler on the digital image of the cephalogram to the 30 mm ruler that was shown on the platform screen.

Digital manual-operated analysis

At the beginning, the radiographs were calibrated. 15 hard tissue cephalometric landmarks were used in this study (Sella (S), Nasion (N), A-point (A), B-point (B), Porion (Po), Orbitale (Or), Pogonion (Pg), Gnathion (Gn), Menton (Me), Anterior Nasal Spine (ANS), Posterior Nasal Spine (PNS), Condylion (Co), Gonion (Go), Articular (Ar), and Condylar (Co), and 10 soft-tissue cephalometric landmarks (Glabella (G), Soft Tissue Nasion (Ns), Pronasale (Pn), Subnasale (Sn), Labrale Superius (Ls), Stomion (St), Labrale Inferius (Li), Soft Tissue Pogonion (Pog'), Soft Tissue Gnathion (Gn'), and Soft Tissue Menton (Me')) were visually located and labelled using manual digital tracing (WebCeph™, Cephio, and Ceppro DDH Inc.). Landmarks were located by the first orthodontist manually, then the accuracy of the localization of the landmarks was verified by a second orthodontist, and then the analysis was automatically derived.

Digital AI-operated analysis

The second set of tracings was performed immediately after the manually operated landmark analysis was computed by using the WebCeph® AI Digitization function to automatically identify and trace landmarks. In the end, the cephalometric measurement values for all parameters were downloaded in Excel format and then transferred into a single Microsoft Office Excel spreadsheet with data that was collected from the manually operated analysis. The manually operated cephalometric analysis was performed first, then the same radiograph was reanalyzed using the AI-operated analysis. Five patients' radiographs were analyzed in each section.

Statistical analysis

SPSS for Windows (version 23.0; SPSS Inc., Chicago, IL, USA) was used for statistical analyses. The level of significance was set at $p < 0.5$. Interclass correlation coefficients (ICC) were calculated to determine the reliability of the two approaches.

RESULTS

Table 1 shows the comparison of descriptive data in both the manual and AI approaches. The paired t-test was utilized to discover the differences between the groups. Starting with SNA, the mean difference was -0.770, indicating a significant difference between the two techniques for this assessment (*p*-value = 0.018). The ANB's mean difference was -0.964, with a *p*-value of < 0.001, showing a significant difference between the two approaches. Moving on to the Bjork sum, the mean difference was very small (0.508), and the p-value of 0.190 indicates no significant difference between the techniques.

Table 1: Descriptive statics illustration for all readings of the groups with statistical evaluation (n=47 for all landmarks)

	Landmark	Manual	AI	Mean Difference	*p*-value
SKELETAL	SNA	81.701±4.524	82.471±4.022	-0.770	0.018
	SNB	78.467±4.297	78.272±3.890	0.195	0.445
	ANB	3.234±3.316	4.198±2.833	-0.964	0.000
	Bjork sum	392.514±5.989	392.006±5.533	0.508	0.190
	FMA	23.062±7.051	23.195±5.685	-0.133	0.808
	Gonial angle	123.262±8.623	122.013±6.838	1.249	0.121
	A to N-Perp (FH)	2.485±10.434	2.622±7.751	-0.137	0.890
	B to N-Perp (FH)	-6.704±14.729	-9.599±12.942	2.894	0.036
	Pog to N-Perp (FH)	-4.356±16.696	-6.993±14.41	2.637	0.074
	FH to AB	82.258±7.644	80.956±6.468	1.302	0.018
	A-B to mandibular plane	74.679±8.809	75.851±7.063	-1.171	0.031
	Wits appraisal	4.264±9.296	5.200±8.005	-0.936	0.171
DENTAL	Overjet	8.174±4.983	7.443±4.059	0.731	0.024
	Overbite	3.377±5.19	3.651±4.531	-0.274	0.367
	U1 to FH	117.702±9.547	113.750±7.71	3.952	0.000
	U1 to SN	108.249±9.197	104.938±7.969	3.311	0.000
	U1 to UOP	56.499±6.843	59.207±5.767	-2.708	0.000
	IMPA	96.210±9.486	95.782±8.561	0.429	0.419
	L1 to LOP	62.587±8.742	63.686±8.015	-1.099	0.091
	Interincisal angle	123.027±12.661	127.274±12.173	-4.247	0.000
	Canting of occlusal plane	4.076±6.121	5.083±4.849	-1.007	0.088
	U1 to NA (mm)	10.245±5.635	6.840±3.722	3.405	0.000
	U1 to NA (deg)	26.549±8.636	22.468±7.002	4.080	0.000
	L1 to NB (mm)	10.237± 6.023	10.596±5.924	-0.359	0.170
	L1 to NB (deg)	27.19±8.625	25.725±7.925	0.627	0.581

Values were expressed as mean±SD. Data was analyzed using paired t-test; p-values <0.05 are significantly different.

Similarly, FMA yields a non-significant result (*p*-value = 0.808), showing no major difference between the approaches. The analysis of the gonial angle revealed a mean difference of 1.249, and while the p-value of 0.121 suggests some difference, it did not reach statistical significance. The *p*-value for A to N-Perp (FH) was 0.890, indicating no significant difference. However, there was a substantial difference between B and N-Perp (FH), with a mean difference of 2.894 and a p-value of 0.036. Pog to N-Perp (FH) yielded a borderline result, with a *p*-value of 0.074, indicating that the difference was not statistically significant. In the instance of FH to AB, the mean difference of 1.302 turned out to be statistically significant (*p*-value = 0.018). A-B to mandibular planes also showed a significant difference (*p*-value = 0.031). The Wits appraisal had a mean difference of -0.936; however, the *p*-value of 0.171 indicated that there was no significant difference. However, Overjet demonstrated a substantial difference, with a mean difference of 0.731 and a *p*-value of 0.024. Overbite, on the other hand, did not reveal a significant difference (*p*-value = 0.367). For U1 to FH, the mean difference was quite large, with 3.952, and resulted in the *p*-value of < 0.001 to be a highly significant difference between the methods. Similarly, U1 to SN and U1 to UOP both showed highly significant differences, with *p*-values of < 0.001. Meanwhile, IMPA did not show a significant difference (*p*-value = 0.419), nor did L1 to LOP. The interincisal angle has a highly significant difference, with a *p*-value of < 0.001 and a mean difference of -4.247. Canting of the occlusal plane also approaches significance with a p-value of 0.088, though it does not quite reach the threshold. Furthermore, substantial differences between U1 to NA (mm) and U1 to NA (deg) occurred, with *p*-values < 0.001. However, there were no significant differences from L1 to NB (mm) and L1 to NB (deg) (*p*-values of 0.170 and 0.581, respectively). In summary, the results highlight significant differences for several measurements, particularly SNA, ANB, ODI, combination factor, B to N-Perp(FH), FH to AB, overjet, U1 to FH, U1 to SN, U1 to UOP, interincisal angle, U1 to NA (mm), U1 to NA (deg), and nasolabial angle, while others like SNB, FMA, Bjork sum, and overbite showed no significant differences. These findings suggested that while the two methods may be comparable in some cases, there were notable discrepancies in specific measurements. The results of the Intraclass Correlation Coefficient (ICC) test provided important insights into the level of agreement or reliability between the two measurement methods (manual and AI), as presented in Table 2.

Table 2: Intraclass correlation coefficient test results

Readings	ICC Value	*p*-value	95% CI for ICC	
			Lower Bound	Upper Bound
SNA	0.861	0.000	0.752	0.922
SNB	0.911	0.000	0.846	0.949
ANB	0.828	0.000	0.604	0.917
Bjork sum	0.895	0.000	0.820	0.940
FMA	0.833	0.000	0.718	0.903
Gonial angle	0.752	0.000	0.595	0.854
APDI	0.864	0.000	0.767	0.922
A to N-Perp (FH)	0.733	0.000	0.566	0.842
B to N-Perp (FH)	0.766	0.000	0.610	0.864
Pog to N-Perp (FH)	0.791	0.000	0.653	0.879
FH to AB	0.857	0.000	0.745	0.920
A-B to mandibular plane	0.890	0.000	0.804	0.938
Wits appraisal	0.856	0.000	0.756	0.917
Overjet	0.879	0.000	0.785	0.933
Overbite	0.911	0.000	0.846	0.949
U1 to FH	0.771	0.001	0.338	0.903
U1 to SN	0.803	0.000	0.474	0.912
U1 to UOP	0.650	0.000	0.366	0.808
IMPA	0.921	0.000	0.863	0.955
L1 to LOP	0.859	0.000	0.760	0.919
Interincisal angle	0.865	0.000	0.521	0.947
Canting of occlusal plane	0.735	0.000	0.569	0.843
U1 to NA (mm)	0.583	0.016	0.048	0.811
U1 to NA (deg)	0.709	0.002	0.255	0.871
L1 to NB (mm)	0.956	0.000	0.922	0.975
L1 to NB (deg)	0.667	0.000	0.433	0.817

The ICC values represent the consistency of the readings, with values closer to 1 indicating high reliability and values near 0 indicating low reliability. The intraclass correlation coefficient (ICC) data showed that the two sets of measurements were very consistent for most readings. Variables such as SNB (ICC = 0.911), SNA (ICC = 0.861), and overbite (ICC = 0.911) demonstrated very high agreement, with ICC values greater than 0.9, indicating strong dependability in both techniques. Several others, including ANB (ICC = 0.828) and gonial angle (ICC = 0.752), had significant agreement, with ICC values ranging from 0.75 to 0.89, showing strong consistency. However, some readings, like U1 to UOP (ICC = 0.650) and U1 to NA (mm) (ICC = 0.583), illustrated moderate agreement, suggesting more variation between methods. Overall, the ICC results revealed that most variables were assessed consistently by both methods, but a few individual readings stated significant differences, particularly in dental and face soft tissue measurements.

DISCUSSION

This study was primarily designed to evaluate the accuracy and reliability of using the fully automated WebCeph™ cephalometric analysis in comparison to manually operated cephalometric analysis, as the latter one had been proved in previous studies to have good reliability and reproducibility [7]. AI models' effectiveness is based on their ability to engage, help, and please users in addition to their capacity to provide correct or comprehensive information [8]. The AI performance in the identification system of the landmarks within 2mm error variance is considered to be clinically accepted in the literature [7,9]. While AI may reach high accuracy in landmark identification (within 2 mm), the consequences for angular measures are considerable, emphasizing the need to consider both detection accuracy and clinical effect [10]. In this study, the fully automated landmark was compared to the digital manual localization of the landmarks, as the researchers had reported that the AI was capable of identifying landmarks similar to moderate-level experience in landmark detection [11,12]. To evaluate the accuracy of WebCeph™, cephalometric measures were utilized over landmark identification. These measurements are the result of the cephalometric tracing process providing information to assist with treatment planning, and errors in landmark location during measuring may cancel out or aggravate the discrepancy [13,14]. The objective of this research was to give a valid and useful comparison that is based on real-life orthodontic needs by including many important cephalometric analyses. The null hypothesis showed that there was no significant difference between manually operated and AI-operated digital measurements. The current findings included 22 of 25 parameters that had an ICC value above 0.70, which means high reliability. Only U1 to NA, U1 to UOP, and L1 to NB (deg) had moderate reliability. The reason for the lower ICC value may be attributed to faulty identification of landmarks by the software in a few cases. These findings go with a recent study that compared fully automated linear and angular measurement using WebCeph™ and manual tracing [15]. Other researchers who studied web-based platforms or computer software had reported having reproducibility like CephX, FACAD® computer software [16]. On the other hand, a study compared 12 cephalometric measurements obtained by Dolphin Imaging, CephNinja, CephX, and manual tracing and reported that fully automatic analysis with CephX needs to be more reliable, but with manual adjustment of the landmark, the degree of reliability was increased [17]. This study reported that there was a significant difference between the SNA, ANB, B to N perp (FH), U1 to FH, U1 to SN, U1 to UOP, interincisal angle, U1 to NB (mm), and U1 to NA (deg). But only B to N perp (FH), U1 to FH, U1 to SN, U1 to UOP, interincisal angle, U1 to NA (mm), and U1 to NA (deg) are considered clinically significant. The errors in landmark positions have a great impact on both linear and angular measurements. This meant that the manual adjustment of the landmark is a must when using the WebCeph™ AI-based cephalometric analysis. This finding goes with a very recent study that found that the angular and linear measurements based on AI *vs*. human landmark localization were significantly different and the use of automated methods should be done with care since any mistake in landmark identification may lead to serious errors in cephalometric measurements, which can impact orthodontic diagnosis and planning [18].

Study limitations

This study was performed using Webceph platform, comparing these findings to other platforms and software is recommended.

Conclusion

Most of the measurements were assessed consistently by both methods, but a few individual readings stated significant differences, particularly in dental measurements. For that reason, AI-assisted WebCeph can be used as an assistive tool, but the landmarks should be adjusted before any cephalometric analysis starts.

Conflict of interests

The authors declared no conflict of interest.

Funding source

The authors did not receive any source of funds.

Data sharing statement

Supplementary data can be shared with the corresponding author upon reasonable request.

REFERENCES

1. Mercier JP, Rossi C, Sanchez IN, Renovales ID, Sahagún PM, Templier L. Reliability and accuracy of Artificial intelligence-based software for cephalometric diagnosis. A diagnostic study. *BMC Oral Health*. 2024;24(1):1309. doi: 10.1186/s12903-024-05097-6.
2. Narkhede S, Rao P, Sawant V, Sachdev SS, Arora S, Pawar AM, et al. Digital versus manual tracing in cephalometric analysis: A systematic review and meta-analysis. *J Pers Med*. 2024;14(6):566. doi: 10.3390/jpm14060566.
3. Jihed Mh, Dallel I, Tobji S, Amor AB. The impact of artificial intelligence on contemporary orthodontic treatment planning-a systematic review and meta-analysis. *Sch J Dent Sci*. 2022;5:70-87. doi: 10.36347/sjds.2022.v09i05.001.
4. Hügle M, Omoumi P, van Laar JM, Boedecker J, Hügle T. Applied machine learning and artificial intelligence in rheumatology. *Rheumatol Adv Pract*. 2020;4(1):rkaa005. doi: 10.1093/rap/rkaa005.
5. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data*. 2021;8(1):53. doi: 10.1186/s40537-021-00444-8.
6. Khosravani S, Esmaeili S, Malek Mohammadi N, Eslamian L, Dalaie K, Motamedian SR. Inter and intra-rater reliability of lateral cephalometric analysis using 2D dolphin imaging software. *J Dent School*. 2020;38(4):148-152. doi: 10.22037/jds.v38i4.35384.
7. Hwang HW, Park JH, Moon JH, Yu Y, Kim H, Her SB, et al. Automated identification of cephalometric landmarks: Part 2-Might it be better than human*? Angle Orthod*. 2020;90(1):69-76. doi: 10.2319/022019-129.1.
8. Daraqel B, Wafaie K, Mohammed H, Cao L, Mheissen S, Liu Y, et al. The performance of artificial intelligence models in generating responses to general orthodontic questions: ChatGPT vs Google Bard. *Am J Orthod Dentofacial Orthop*. 2024;165(6):652-662. doi: 10.1016/j.ajodo.2024.01.012.
9. Tsolakis IA, Tsolakis AI, Elshebiny T, Matthaios S, Palomo JM. Comparing a fully automated cephalometric tracing method to a manual tracing method for orthodontic diagnosis. *J Clin Med*. 2022;11(22):6854. doi: 10.3390/jcm11226854.
10. Schwendicke F, Chaurasia A, Arsiwala L, Lee JH, Elhennawy K, Jost-Brinkmann PG, et al. Deep learning for cephalometric landmark detection: systematic review and meta-analysis. *Clin Oral Investig*. 2021;25(7):4299-4309. doi: 10.1007/s00784-021-03990-w.
11. Kim J, Kim I, Kim YJ, Kim M, Cho JH, Hong M, et al. Accuracy of automated identification of lateral cephalometric landmarks using cascade convolutional neural networks on lateral cephalograms from nationwide multi-centres. *Orthod Craniofac Res*. 2021;24 Suppl 2:59-67. doi: 10.1111/ocr.12493.
12. Lee Y, Pyeon JH, Han SH, Kim NJ, Park WJ, Park JB. A comparative study of deep learning and manual methods for identifying anatomical landmarks through cephalometry and cone-beam computed tomography: A systematic review and meta-analysis. *Appl Sci*. 2024;14(16):7342. doi: 10.3390/app14167342.
13. Ongkosuwito EM, Katsaros C, van 't Hof MA, Bodegom JC, Kuijpers-Jagtman AM. The reproducibility of cephalometric measurements: a comparison of analogue and digital methods. *Eur J Orthod*. 2002;24(6):655-665. doi: 10.1093/ejo/24.6.655. PMID: 12512783..
14. Sayinsu K, Isik F, Trakyali G, Arun T. An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. *Eur J Orthod*. 2007;29(1):105-198. doi: 10.1093/ejo/cjl065.
15. Mahto RK, Kafle D, Giri A, Luintel S, Karki A. Evaluation of fully automated cephalometric measurements obtained from web-based artificial intelligence driven platform. *BMC Oral Health*. 2022;22(1):132. doi: 10.1186/s12903-022-02170-w.
16. Alqahtani H. Evaluation of an online website-based platform for cephalometric analysis. *J Stomatol Oral Maxillofac Surg*. 2020;121(1):53-57. doi: 10.1016/j.jormas.2019.04.017.
17. Meriç P, Naoumova J. Web-based fully automated cephalometric analysis: Comparisons between App-aided, computerized, and manual tracings. *Turk J Orthod*. 2020;33(3):142-149. doi: 10.5152/TurkJOrthod.2020.20062.
18. Kang S, Kim I, Kim YJ, Kim N, Baek SH, Sung SJ. Accuracy and clinical validity of automated cephalometric analysis using convolutional neural networks. *Orthod Craniofac Res*. 2024;27(1):64-77. doi: 10.1111/ocr.12683.