Peng et al. Head & Face Medicine (2025) 21:22
https://doi.org/10.1186/s13005-025-00499-5

Head & Face Medicine

RESEARCH

Open Access

Predicting changes of incisor and facial profile following orthodontic treatment: a machine learning approach

Jing Peng[1,2], Yan Zhang[1], Mengyu Zheng[1], Yanyan Wu[1], Guizhen Deng[2], Jun Lyu[3,4*†] and Jianming Chen[1*†]

Abstract

Background Facial aesthetics is one of major motivations for seeking orthodontic treatment. However, even for experienced professionals, the impact and extent of incisor and soft tissue changes remain largely empirical. With the application of interdisciplinary approach, we aim to predict the changes of incisor and profile, while identifying significant predictors.

Methods A three-layer back-propagation artificial neural network model (BP-ANN) was constructed to predict incisor and profile changes of 346 patients, they were randomly divided into training, validation and testing cohort in the ratio of 7:1.5:1.5. The input data comprised of 28 predictors (model measurements, cephalometric analysis and other relevant information). Changes of U1-SN, LI-MP, Z angle and facial convex angle were set as continuous outcomes, mean square error (MSE), mean absolute error (MAE) and coefficient of determination (R^2) were used as evaluation index. Change trends of Z angle and facial convex angle were set as categorical outcomes, accuracy, precision, recall, and F1 score were used as evaluation index. Furthermore, we utilized SHapley Additive exPlanations (SHAP) method to identify significant predictors in each model.

Results MSE/MAE/R^2 values for U1-SN were 0.0042/0.055/0.84, U1-SN, MP-SN and ANB were identified as the top three influential predictors. MSE/MAE/R^2 values for L1-MP were 0.0062/0.063/0.84, L1-MP, ANB and extraction pattern were identified as the top three influential predictors. MSE/MAE/R^2 values for Z angle were 0.0027/0.043/0.80, Z angle, MP-SN and LL to E-plane were considered as the top three influential indicators. MSE/MAE/R^2 values for facial convex angle were 0.0042/0.050/0.73, LL to E-plane, UL to E-plane and Z angle were considered as the top three influential indicators. Accuracy/precision/recall/F1 Score of the change trend of Z angle were 0.89/1.0/0.80/0.89, Z angle, Lip incompetence and LL to E-plane made the largest contributions. Accuracy/precision/recall/F1 Score of the change trend of facial convex angel were 0.93/0.87/0.93/0.86, key contributors were LL to E-plane, UL to E-plane and Z angle.

†Jun Lyu and Jianming Chen contributed equally to this work.

*Correspondence:
Jun Lyu
lyujun2020@jnu.edu.cn
Jianming Chen
orangeforest393@163.com
Full list of author information is available at the end of the article

Conclusion BP-ANN could be a promising method for objectively predicting incisor and profile changes prior to orthodontic treatment. Such model combined with key influential predictors could provide valuable reference for decision-making process and personalized aesthetic predictions.

Keywords Artificial neural network, Orthodontic treatment, Profile prediction, Incisor prediction, SHapley additive exPlanations

Background

Orthodontics is a discipline dedicated to achieve high standards of occlusion, aesthetics, and long-term stability. Nonetheless, with social development and the explosion of information, more patients seek orthodontic treatment primarily to improve their appearance, and the degree of attractiveness has been found to be somewhat related to personality, social interaction and life qualities [1, 2]. Although individuals' perceptions of beauty vary, the standard esthetic concern addressed by orthodontic therapy focuses on correcting sagittal skeletal discrepancies. In other words, the common goal when treating Class II or III patients is to minimize their deviation from Class I or attenuate the abnormal maxillomandibular relationship, thereby reducing facial disharmony [3]. Commonly, when explaining the customized treatment plan to patients, they often propose: "What changes will happen to my teeth and profile?" [4, 5] Since orthodontic therapy primarily focus on hard tissue, even for experienced professionals, the impact and the extent of teeth and soft tissue change remain largely empirical.

Given the close attachment and proximity of teeth, bone and muscle, the appearance of soft tissue is undeniably influenced by the underlying hard tissue structure. In the past, the change of facial profile was believed to adapt to the underlying dentoalveolar structures at an empirical ratio, which gave rise to the widespread use of Visual Treatment Objective (VTO) software [6]. Nonetheless, the emergence of clinical errors exceeding 2 mm has raised skepticism regarding the reliability and validity, as well as the robustness of multiple regression analysis [7]. Recent studies have shown that the least accurate predictions of this method tend to occur in the soft tissue regions, particularly in the chin and lower lip. This may be explained by the fact that it only integrates imaging data without considering the individual's overall information [8, 9]. With the booming progression of artificial intelligence (AI), its applications in the field of orthodontics have become increasingly widespread. From expert system-based automated cephalometric landmark detection and measurements to two-dimensional image classification, and further to the integration of complex data and architectures for generating comprehensive predictions, including factors such as gender, treatment duration, and growth and development [10–12]. Regrettably, few studies have employed such methodologies in forecasting the change of profile [13–15]. Due to the limited number of samples and predictors in previous research, along with the absence of ranking the predictors by their impact, the applicability of these models is constrained. Therefore, we collected as much information as possible including model measurements, cephalometric analysis and other relevant information to serve as predictors. The back-propagation artificial neural network (BP-ANN) model was selected due to its outstanding performance in handling issues of uncertainty, nonlinearity, lack of configuration and multiple-factor interactions [16].

With the application of interdisciplinary approach, we aim to propose a new holistic method to predict changes in incisor and profile for orthodontic patients prior to treatment, while identifying clinically significant influencing factors.

Methods

Study population

The study included 346 patients (adults and adolescents after pubertal growth peak) who sought fixed orthodontic consultation at the Affiliated Stomatology Hospital of Guangzhou Medical University in Guangzhou, China. Prior to orthodontic treatment, all participants received comprehensive information about the study and signed written informed consent. The study protocol was approved by the Ethical Committee (20240809171826). Extraction patterns included in the study were limited to no extraction, extraction of two premolars, and extraction of four premolars to minimize the heterogeneity of the study. Patients were excluded based on the following criteria: (1) under 14 years of age, (2) presence of missing teeth (excluding third molars) or malformed teeth, (3) history of previous orthodontic treatment or cleft lip and palate, (4) treatment options involving expanders, functional therapy, invisible therapy and orthognathic surgery.

Data collection

Pre-treatment data were collected as predictors, including model measurements (the crowding of upper and lower arches, molar relationship, anterior overbite, anterior overjet and curve of Spee), cephalometric analysis (analysis of bone, dental and soft tissue), and other relevant information (age, sex, lip incompetence, extraction pattern and anchorage mode), resulting in a total of 28 predictors. Only four specific measurements from post-treatment data (U1-SN, L1-MP, Z angle, and facial convex angle) were gathered, and changes in these measurements were calculated as outcomes, by subtracting the pre-treatment values from the post-treatment values. We set 4 continuous variables and 2 categorical variables as outcomes, specifically Y1: the change of U1-SN, Y2: the change of L1-MP, Y3: the change of Z angle, Y4: the change of facial convex angle, Y5: the change trend of Y3 (using −0.5 and 0.5 as the cut-off points, results were divided into three categories: 0 for unchanged, 1 for decreased and 2 for increased) and Y6: the change trend of Y4 (classification criteria consistent with Y5).

Since there is no uniform formula for the sample size calculation of ANN model, our analysis followed the most recent suggestions [17, 18]: a minimum of 50 samples are required to start any meaningful machine learning based data analysis, and at least 10 samples per degree of freedom (predictor) is reasonable [19], which would require a total of 280 samples in the research. Accordingly, our sample size was 346, which met the minimum requirement and was theoretically feasible.

Among these, model measurements were collected from examination records, while other relevant information were gathered from both examination records and medical record system by YZ and YW. To ensure accuracy, the model measurements were re-examined by GD using intraoral scan data. Cephalograms tracings were made by 1 investigator (MZ) and repeated twice at intervals of 2 weeks to minimize measurement errors. Before data analysis, JP reviewed the tracings, and any disagreement would be discussed with JC to reach a consensus. The reference points were digitized with the Dolphin Imaging (v11.95, Dolphin Imaging and Management Solutions Inc., Chatsworth, CA, USA), twenty-six landmarks and 17 measurements were chosen (Fig. 1).

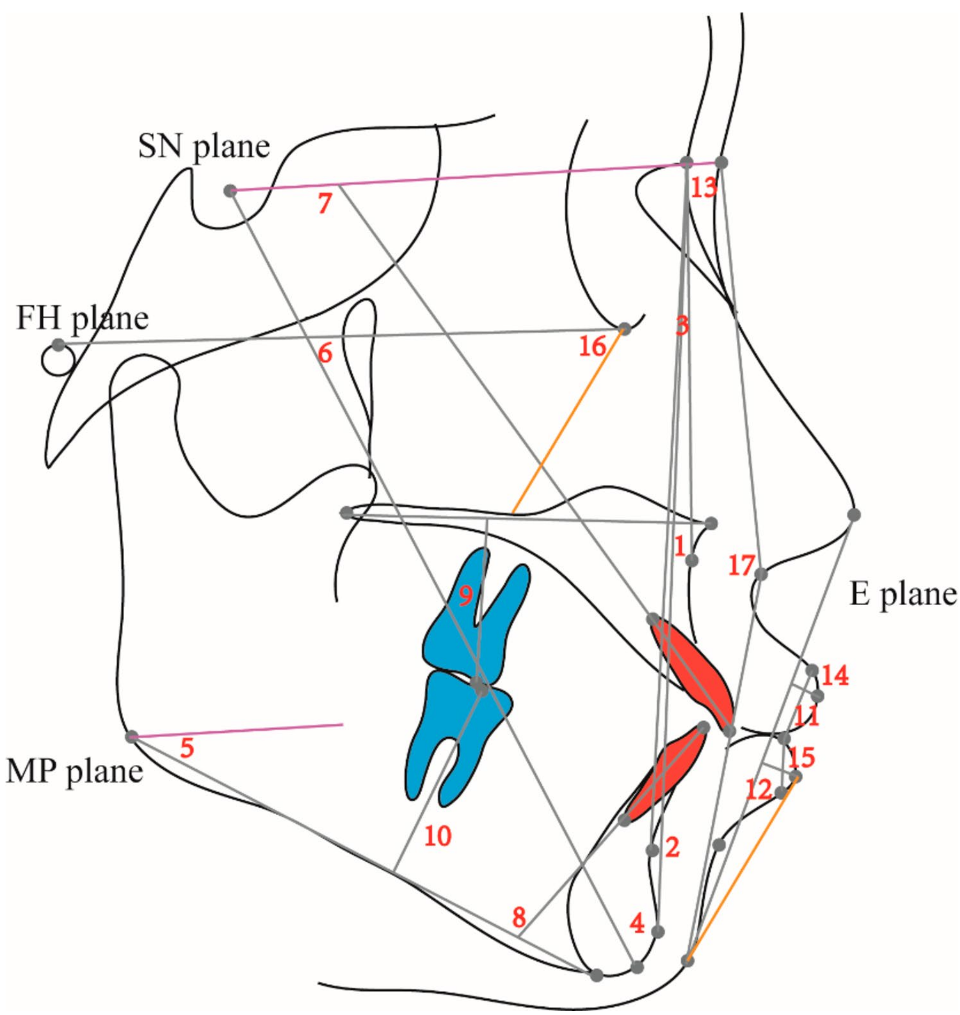

Fig. 1 Index of cephalometric measurements. 1 SNA (°), 2 SNB (°), 3 ANB (°), 4 SNP (°), 5 MP-SN (°), 6 Y-axis (°), 7 U1-SN (°), 8 L1-MP (°), 9 U6-PP (mm), 10 L6-MP (mm), 11 UL-E plane (mm), 12 LL-E plane (mm), 13 SN-Sn (°), 14 UL height (mm), 15 LL height (mm), 16 Z angle (°), 17 facial convex angle (°). Purple and orange lines indicate shifted lines

Network construction

We utilized the Python programming language to construct, train, and test the BP-ANN model, with the process documented in open-access Jupyter Notebooks (https://cimcb.github.io/MetabProjectionViz/). A 3-layer neural network, consisting of 28 input neurons in the input layer and 7 neurons in the hidden layer, was employed for the machine learning task. The hidden neurons functioned as interneurons, learning by adjusting their weighted values, and the number of these neurons was determined through a trial-and-error approach [20].

Continuous input data were normalized to the range [-1, 1] using maximum-minimum normalization before being processed by the neural network. The dataset was randomly divided into three cohorts: 70% for training, 15% for validation and 15% for testing [21]. This BP-ANN model employed the error backward propagation learning algorithm, each layer "shared" the error with its neurons, allowing the reference errors for each layer to be obtained. These reference errors were then used to adjust the connection weights, aiming to minimize the error as much as possible [16]. Iterative learning was halted at the minimum error point of the validation set, specifically at 0.01, and training was preemptively terminated when the mean square error (MSE) of the validation set reached its minimum [22]. Moreover, the momentum parameter of 0.9 was employed to smooth the optimization path in parameter space, mitigating common issues like oscillations and local minima entrapment. The sigmoid function was selected as the activation function for the hidden layers, while linear activation was employed for regression tasks and softmax activation for classification tasks [23].

After selecting the best-fit model, the performance was evaluated on the testing set using appropriate evaluation metrics. For continuous variables (Y1-Y4), the metrics included MSE, mean absolute error (MAE) and the coefficient of determination (R^2). For categorical variables (Y5 and Y6), accuracy, precision, recall, and F1 Score were used to verify the model's accuracy and precision.

Statistical analysis

R-software (version 4.4.0, www.r-project.org) was used to perform the baseline characteristics analyses. The Chi-square test was applied for categorical variables, the student's t-test for continuous variables with a normal distribution, and the Wilcoxon rank-sum test for continuous variables without a normal distribution. Continuous variables were reported as the mean with standard deviation or the median with interquartile range.

We employed the SHapley Additive exPlanations (SHAP) method to interpret the outputs of our machine learning models. This approach could quantify and rank the contribution of each factor to individual predictions, facilitating a comprehensive understanding of model behavior. For further analysis, we identified and focused on the top three indicators that contributed most significantly to each model [24].

Results

Patient characteristics

The population characteristics are shown in Table 1. A total of 346 eligible patients were randomly divided into training cohort ($n = 242$), validation cohort ($n = 52$), and testing cohort ($n = 52$) in the ratio of 7:1.5:1.5.

The median age of patients at diagnosis was 23.5 (21, 26) years, with the majority being female (86.1%). Most had incompetent lips (61.0%) and underwent the extraction model involving removal of four premolars (84.4%). Additionally, most patients presented with mild crowding (68.2% for upper arch and 57.8% for lower arch) and class I molar relationship (56.6%). Among the features associated with cephalometric analysis, the median ANB was 3.75 (2.3, 5.4), the average MP-SN was 34.7 ± 6.9, the median of U1-SN was 107.1 (101.3, 112.1), the average L1-MP was 97.3 ± 9.4, the median of Z angle was 69.1 (64.0, 73.1) and facial convex angle was 164.3 (160.1, 167.9). Regarding anchorage mode, 46.2% of patients received mild or moderate anchorage, while 53.8% received maximum anchorage using implants. However, there was no significant difference among different groups ($P > 0.05$).

Network establishment and evaluation

Figures 2, 3, 4 and 5 illustrate the neural network predictions for the changes in Y1 to Y4. Several indicators were conducted to evaluate the performance of each prediction. Specifically, The MSE for the training/validation/testing cohort of Y1 were 0.0036/0.0050/0.0042, the MAE were 0.048/0.054/0.055, and the R^2 were 0.85/0.82/0.84. For Y2, the MSE for the training/validation/testing cohort were 0.0059/0.0075/0.0062, the MAE were 0.060/0.070/0.063 and the R^2 were 0.83/0.80/0.84. As for Y3, the MSE for the training/validation/testing cohort were 0.0033/0.0039/0.0027, the MAE were 0.046/0.049/0.043 and the R^2 were 0.77/0.75/0.80. Lastly, for Y4, the MSE for the training/validation/testing cohort were 0.0024/0.0039/0.0042, the MAE were 0.040/0.050/0.050 and the R^2 were 0.72/0.71/0.73.

We also verified the performance of the classification outcomes. For Y5, the accuracy for the training/validation/testing cohorts were 0.80/0.81/0.89. The precision for the validation/testing cohorts were 1.0/1.0, the recall were 0.75/0.80, and the F1 score were 0.86/0.89. For Y6, the accuracy for the training/validation/testing cohorts were 0.88/0.92/0.93. The precision for the validation/testing cohorts were 0.93/0.87, the recall were 0.78/0.93, and the F1 score were 0.85/0.86.

Table 1 Baseline clinical and imaging characteristics of 346 eligible patients

Predictors	Overall (346 patients)	Training cohort (242 patients)	Validation cohort (52 patients)	Testing cohort (52 patients)	*P* value
Sex					0.66
male	48(13.9)	31(12.8)	9(17.3)	8(15.4)	
female	298(86.1)	211(87.2)	43(82.7)	44(84.6)	
Extraction pattern					0.77
no extraction	32(9.2)	24(9.9)	4(7.7)	4(7.7)	
two premolars	22(6.4)	13(5.4)	4(7.7)	5(9.6)	
four premolars	292(84.4)	205(84.7)	44(84.6)	43(82.7)	
Lip incompetence					0.97
no	135(39.0)	94(38.8)	21(40.4)	20(38.5)	
yes	211(61.0)	148(61.2)	31(59.6)	32(61.5)	
Upper crowding					0.13
mild	236(68.2)	164(67.8)	32(61.6)	40(76.9)	
moderate	70(20.2)	54(22.3)	9(17.3)	7(13.5)	
severe	21(6.1)	14(5.8)	6(11.5)	1(1.9)	
space	19(5.5)	10(4.1)	5(9.6)	4(7.7)	
Lower crowding					0.14
mild	200(57.8)	132(54.6)	28(53.9)	30(57.7)	
moderate	118(34.1)	88(36.3)	20(38.4)	19(36.5)	
severe	19(5.5)	15(6.2)	3(5.8)	2(3.8)	
space	9(2.6)	7(2.9)	1(1.9)	1(1.9)	
Molar relationship					0.41
I	196(56.6)	141(58.3)	30(57.7)	25(48.1)	
II	93(26.9)	65(26.9)	15(28.8)	13(25.0)	
III	54(15.6)	33(13.6)	7(13.5)	14(26.9)	
other	3(0.9)	3(1.2)	0	0	
Overbite					0.85
normal	210(60.7)	150(62.0)	32(61.4)	28(53.9)	
deep I	59(17.0)	40(16.5)	9(17.3)	10(19.2)	
deep II	20(5.8)	13(5.4)	3(5.8)	4(7.7)	
deep III	22(6.4)	17(7.0)	2(3.9)	3(5.8)	
reverse I	23(6.6)	13(5.4)	4(7.7)	6(11.5)	
reverse II	4(1.2)	2(0.8)	2(3.9)	1(1.9)	
open I	6(1.7)	5(2.1)	0	0	
open II	1(0.3)	1(0.4)	0	0	
open III	1(0.3)	1(0.4)	0	0	
Overjet					0.96
normal	130(37.6)	89(36.8)	19(36.5)	22(42.3)	
deep I	125(36.1)	90(37.2)	20(38.5)	15(28.8)	
deep II	70(20.2)	48(19.8)	11(21.2)	11(21.2)	
deep III	13(3.8)	8(3.3)	2(3.8)	3(5.8)	
reverse I	7(2.0)	6(2.5)	0	1(1.9)	
reverse II	1(0.3)	1(0.4)	0	0	
Anchorage mode					0.54
mild	135(39.0)	95(39.3)	21(40.4)	19(36.6)	
moderate	25(7.2)	22(9.1)	2(3.8)	1(1.9)	
maximum	186(53.8)	125(51.6)	29(55.8)	32(61.5)	
Age	23.5(21, 26)	23(20.3, 25)	24(19.8, 26)	24(22, 25.3)	0.64
Spee curve	2.5(2, 3.5)	2.5(2, 3.5)	3(2, 3.5)	2.8(2, 3.6)	0.84
ANB	3.75(2.3, 5.4)	3.8(2.0, 5.4)	3.6(2.5, 5.3)	4.2(2.9, 5.6)	0.28
SNA	82.7 ± 3.6	82.6 ± 3.6	83.2 ± 3.5	82.6 ± 3.7	0.52
SNB	79.0 ± 3.6	79.0 ± 3.7	79.8 ± 3.4	78.4 ± 3.4	0.13
SNP	80.2 ± 3.4	80.2 ± 3.5	80.3 ± 3.6	80.2 ± 3.1	0.95

Table 1 (continued)

Predictors	Overall (346 patients)	Training cohort (242 patients)	Validation cohort (52 patients)	Testing cohort (52 patients)	*P* value
MP-SN	34.7 ± 6.9	34.8 ± 6.9	35.1 ± 6.7	33.8 ± 7.5	0.58
Y Axis	64.4 ± 3.6	64.3 ± 3.7	64.5 ± 3.4	64.2 ± 3.5	0.91
L1-MP	97.3 ± 9.4	97.3 ± 9.2	95.4 ± 9.4	98.8 ± 10.3	0.17
U1-SN	107.1(101.3, 112.1)	107.1(101.3, 111.7)	107.3(102.4, 112.1)	107.1(101.0, 112.6)	0.97
U6-PP	20.5 ± 2.3	20.6 ± 2.2	20.2 ± 2.3	20.4 ± 2.3	0.44
L6-MP	30.3(28.6, 32.1)	30.1(28.6, 31.9)	30.4(28.3, 32.0)	30.7(28.6, 32.2)	0.56
Ul height	10.3(9.4, 11.3)	10.4(9.5, 11.4)	10.1(9.2, 11.0)	10.6(9.4, 11.5)	0.32
Ll height	9.0(7.6, 10.2)	9.1(7.7, 10.2)	8.8(7.6, 10.3)	8.9(7.5, 10.2)	0.68
Ll to E-plane	3.0 ± 3.1	2.9 ± 3.0	2.6 ± 3.4	3.4 ± 3.4	0.25
Ul to E-plane	1.4(0.5 ± 3.2)	1.3(-0.7,3.2)	1.9(0, 3.6)	1.65(0.15, 3.15)	0.35
SN-Sn	74.3 ± 4.9	74.1 ± 4.7	74.8 ± 5.2	74.6 ± 5.3	0.54
Z angle	69.1(64.0, 73.1)	69.5(64.4, 73.2)	69.6(65.0, 73)	66(63.1, 71.5)	0.07
facial convex angle	164.3(160.1, 167.9)	164.1(160.2, 167.9)	165.5(160.3, 168.1)	165.8(159.9, 167.8)	0.77

SHAP analysis

SHAP summary plot provides a visual representation of the impact of various predictors on the outcome. As for Y1, larger values of U1-SN, MP-SN and ANB were more likely to negatively affect Y1, which corresponded with more retraction of upper incisors. Regarding Y2, larger L1-MP values, smaller ANB angles and closer proximity to Class III malocclusion were more likely to negatively affect Y2, which related to more retraction of lower incisors. For Y3, larger values of Z angle and MP-SN, combined with smaller values of LL to E-plane, were more likely to negatively affect Y3, which indicated the decrease of Z angle. Similarly, for Y4, smaller values of LL to E-plane, larger values of UL to E-plane and Z angle were more likely to negatively affect Y4, which implied the decrease of facial convex angle. (Figures 6, 7, 8 and 9)

The SHAP bar plot visually displays the mean absolute SHAP values for various features on the categorical outcomes. For Y5, soft tissue Z angle, Lip incompetence and LL to E-plane made the largest contributions. In the case of Y6, the key contributors were LL to E-plane, UL to E-plane and Z angle. (Figures 10 and 11)

Discussions

Since its inception in the 1950s, AI has advanced rapidly and is now widely used in orthodontics [10, 25]. For intricate clinical questions, it not only enhances efficiency and productivity, but also assists researchers in identifying key points that may have been overlooked in large datasets, thereby reducing the subjective bias commonly found in clinical practice [3, 26, 27].

Our results revealed that constructed BP-ANN models demonstrated strong capability in analyzing patients' comprehensive information and forecasting changes of incisor and profile before orthodontic treatment. To be specific, the fitting degree for incisors were better than that for soft tissue. This may due to teeth are generally designed to achieve or approach the standard value. As for borderline cases with severe skeletal deformities, teeth would retain compensatory labial or lingual inclination to improve the profile without requiring orthognathic surgery [28, 29]. However, factors such as the available alveolar space, the design and control of anchorage, the length of the teeth roots, the aesthetic standards and oral hygiene conditions may contribute to the deviation of results [30, 31].

We further found that compared with the quantitative prediction of profile change, the qualitative prediction was more accurate. The reason may be that qualitative prediction relies more on guidelines and observation, whereas the actual clinical process involves a variety of uncertain and dynamic factors that could complicate the precision of quantitative predictions. For instance, when a patient presents with small Z Angle and convex profile, orthodontists often try to retract incisors and rotate the mandibular counterclockwise to increase the Z Angle while reduce facial convexity [32]. Conversely, if a patient has large Z Angle and concave profile, the focus shifts to increase incisor inclination with careful consideration of tooth extraction, to decrease the Z Angle while improve facial convexity [33]. Moreover, the deviation in quantitative prediction may result from complex biomechanical responses and multiple interactions among bone, teeth and soft tissue during orthodontic tooth movement. For example, implants designed to achieve maximum anchorage in the sagittal direction may inevitably impact the vertical control, potentially leading to change of occlusal plane and three-dimensional soft tissue [34]. Even in orthodontic-orthognathic surgical treatment, which aims to remove dental compensations and correct overall skeletal discrepancies, the success rate of achieving the predicted facial morphology within a 1 mm error margin was only 54%, due to the complex and nonlinear response of soft tissues to underlying hard-tissue changes

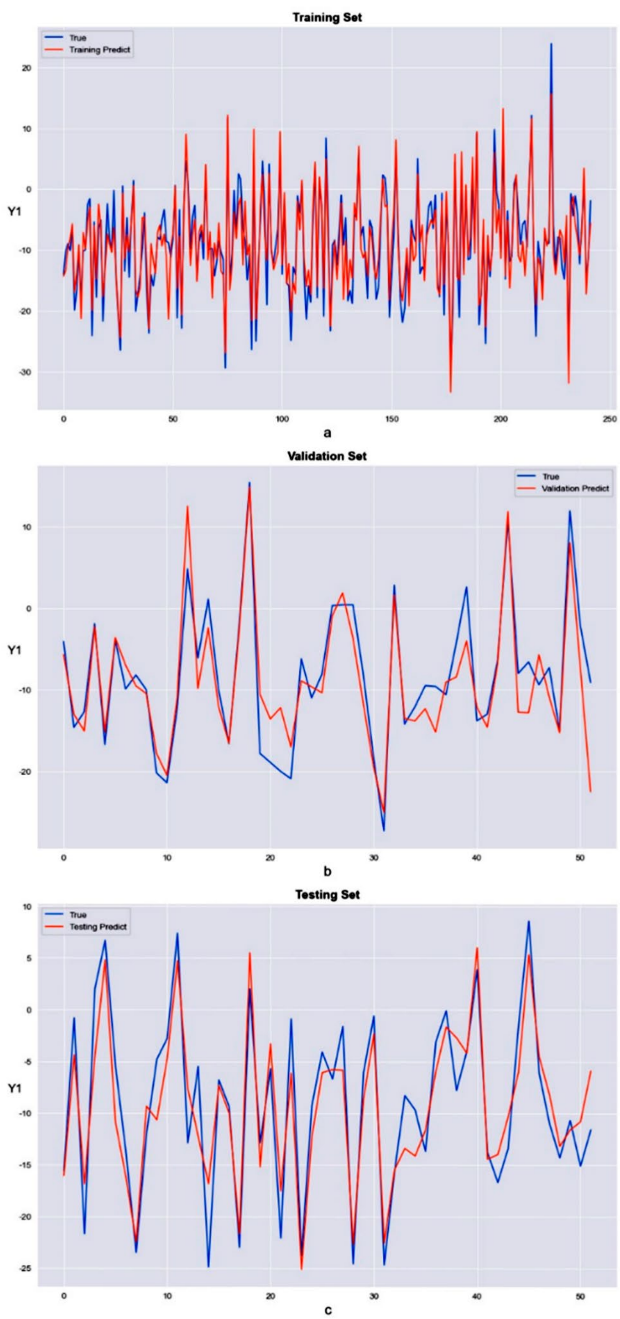

Fig. 2 The neural network predictions for the changes in Y1, (**a**) training cohort, (**b**) validation cohort, (**c**) testing cohort

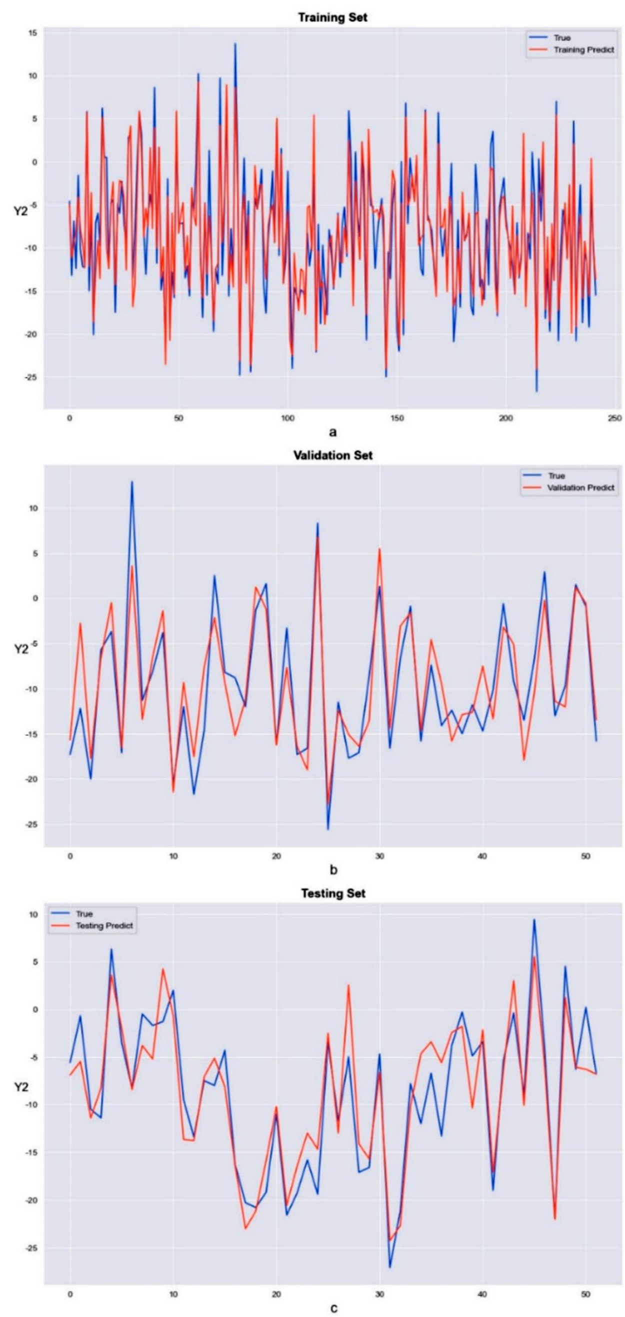

Fig. 3 The neural network predictions for the changes in Y2, (**a**) training cohort, (**b**) validation cohort, (**c**) testing cohort

[14]. Additionally, patient-specific factors such as age, skeletal type, lip thickness, and habits like lip biting and mouth breathing can affect muscle function and the repositioning of soft tissue, further complicating accurate predictions [35–38].

Since our model can meticulously capture the subtle correlation between soft tissue and orthodontic tooth movement, and optimizing soft tissue camouflage is a common goal for both orthodontists and patients, we further analyzed and ranked the factors using SHAP analysis. Interestingly, we found that patients with high angle (MP-SN) positively impacts the retraction of upper incisors, as well as has a negative effect on the change in Z Angle, which presents a contradiction for patients with convex profile. Generally, high angle cases have always been a challenge for orthodontists due to issues like anterior alveolar hypoplasia, lip incompetence, molar extrusion, and clockwise rotation of the mandible, which may lead to downward and posterior rotation of the chin and compromising facial esthetics [39]. Research has found

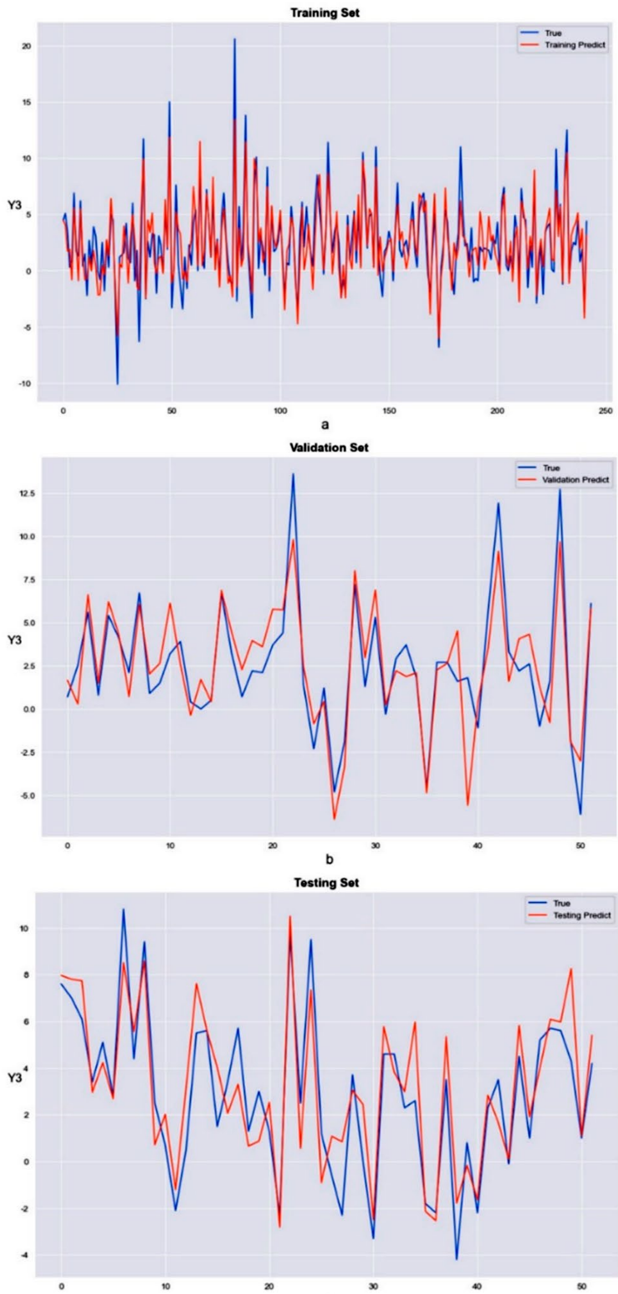

Fig. 4 The neural network predictions for the changes in Y3, (**a**) training cohort, (**b**) validation cohort, (**c**) testing cohort

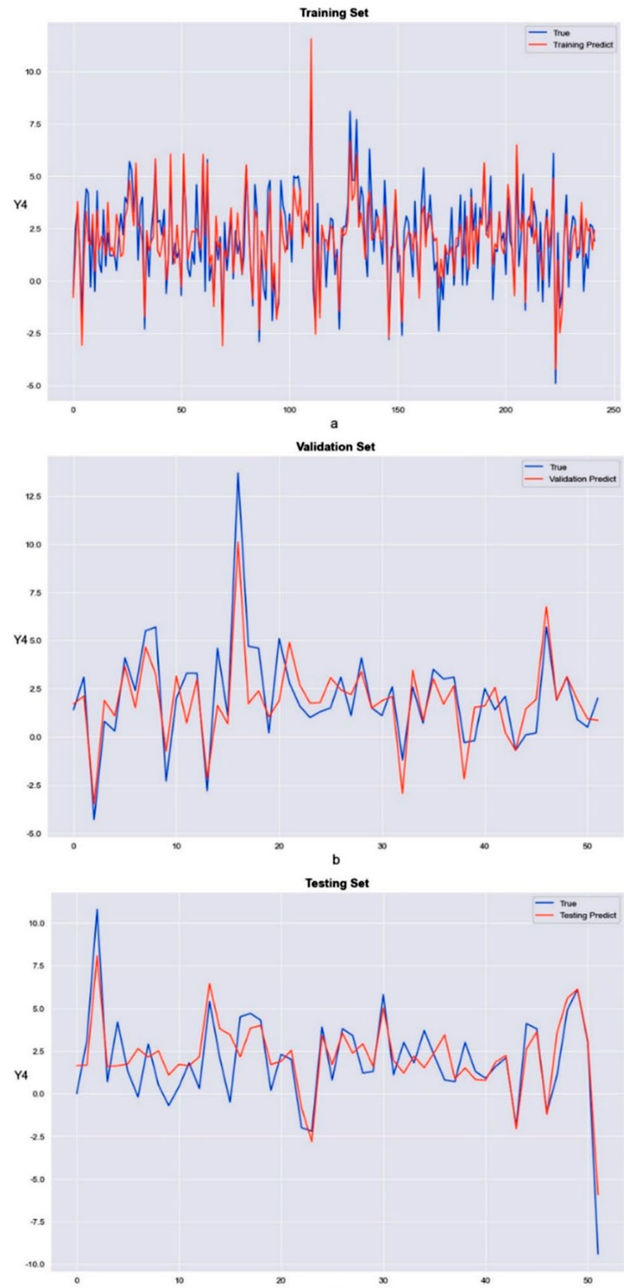

Fig. 5 The neural network predictions for the changes in Y4, (**a**) training cohort, (**b**) validation cohort, (**c**) testing cohort

that effective incisor retraction and good vertical control are beneficial for such patients [34]. Nevertheless, comparative analysis revealed that in patients with high-angle growth pattern, the maxillary palatal alveolar bone was significantly thinner, and the distance between incisor root and incisive canal was relatively small, which restricted incisor retraction [40]. These evidences highlight the unique characteristics of high angle patients in prediction models and emphasize the need for careful risk management when planning incisor retraction.

Another point worth noting is that LL to E-plane has the most significant effect in prediction of soft tissue profile, which ranked in top three in either qualitative or quantitative model. Specifically, patients with convex lower lip are likely to have a positive effect on the Z angle and facial convex angle. Conversely, reduced Z angle and face convex angle are preferable for patients with concave lower lip. This could be attributed to its privileged location as the adjacent esthetic subunit to the chin, which exhibits greater adjustment from tooth relocation [41].

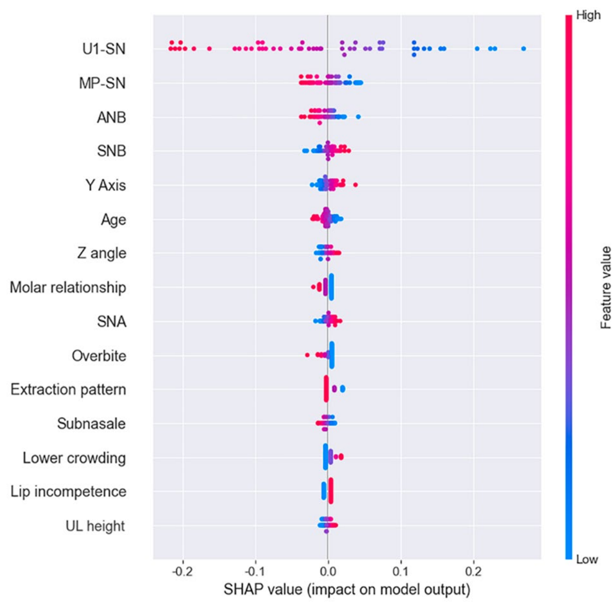

Fig. 6 SHAP summary plot of Y1

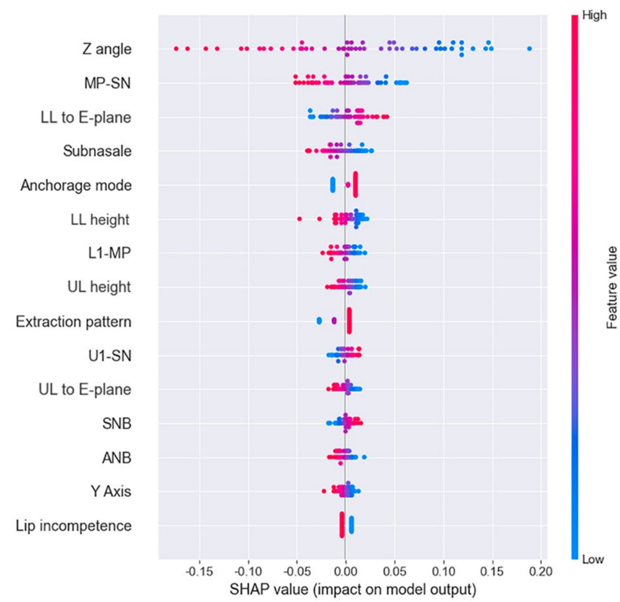

Fig. 8 SHAP summary plot of Y3

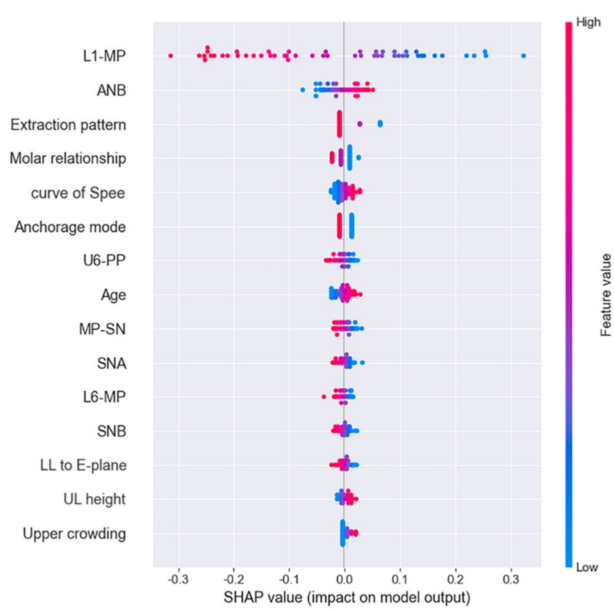

Fig. 7 SHAP summary plot of Y2

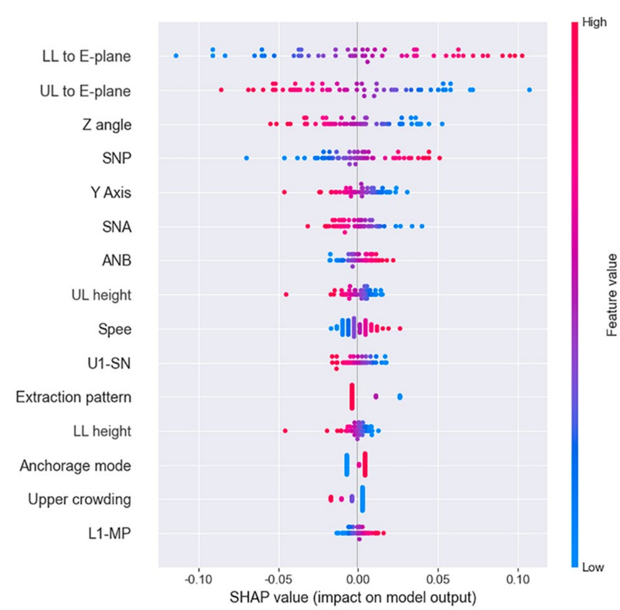

Fig. 9 SHAP summary plot of Y4

In addition to orthodontic tooth movement, factors such as initial incisor inclination, lip tension, thickness and height could also count for the difference [42, 43]. Interestingly, our findings revealed that the effect of UL to E-plane on soft tissue is opposite to that of LL to E-plane. This contradicts the common understanding that upper incisor retraction can alleviate upper lip protrusion, accompanied by the backward movement of the Subnasale (Sn) point and improvement of profile. Nonetheless, remarkable upper lip protrusion is often associated with severe protrusion of the upper incisors and alveolar bone, simply retracting the incisor may not achieve optimal results and may require more complex anchorage, along with higher risk of relapse [44]. Additionally, soft tissues may not fully adapt to the new support structure triggered by tooth relocation, resulting in discrepancies in lip shape and facial contour [45].

Except for the top three influential predictors, other model measurements such as lower arch crowding, the curve of Spee and molar relationships also had important impact on the outcome and ranked top five. It is consistent with previous research [46, 47]. Severe crowding and

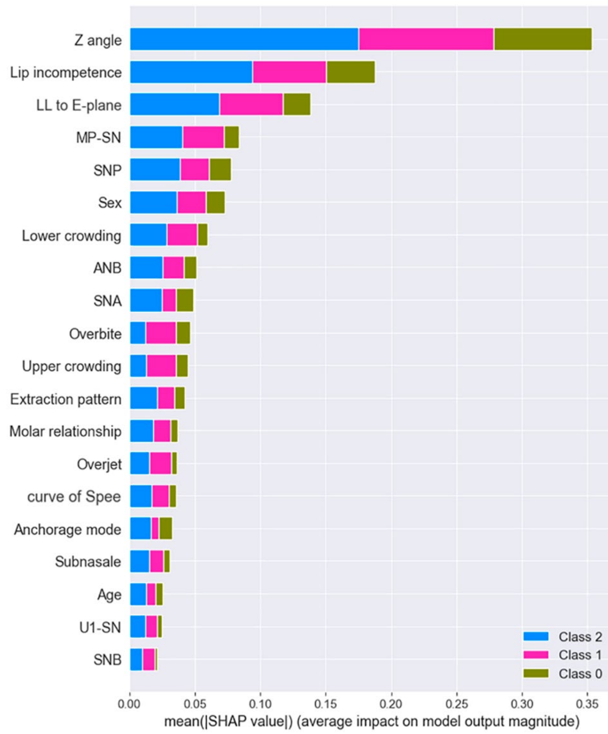

Fig. 10 SHAP bar plot of Y5

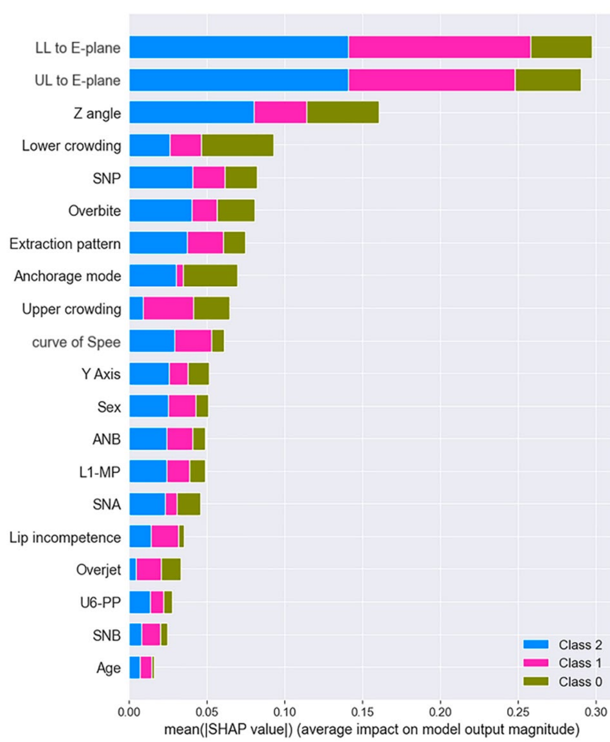

Fig. 11 SHAP bar plot of Y6

deep curve of Spee usually indicate extra need of space, which may interfere with the adjustment of incisor inclination [47]. Also, pre-treatment molar relationship often implies irregular intermaxillary relationship. To achieve the Class I molar relationship may require some compromise in the adjustment of incisor inclination [48]. These findings highlight the importance of gathering comprehensive information and conducting integrated patient evaluation.

Our study underscores the importance of personalized prediction before orthodontic treatment for patients with varying characteristics and highlights the most significant factors. Given the acceptable accuracy of our research results, another clinical utility of the system lies in serving valuable reference for patients and young physicians who are uncertain about extraction strategies, since such decision is a common and important aspect in clinical practice. Orthodontists can assist patients in selecting the extraction pattern that best aligns with their chief complaint and expectation by comparing the predictions of different modes.

Limitation

1) Though cephalometric analysis has long been considered as the key method for profile evaluation and is easily obtained, three-dimensional measurements could provide more comprehensive information, and we are already working on it.
2) Types of extraction were restricted to three modes to reflect the clinical characteristics of the orthodontic patients as much as possible while minimize the heterogeneity of the study. However, it inevitably leads to some loss of patient information and reduction in sample size.
3) A larger sample size, more detailed variables and external testing cohort are expected to validate the practicality of the model.
4) Transversal issues are an important aspect of orthodontics and may be associated with sagittal issues. Although we have excluded cases with expanders, functional and invisible therapy, future studies are expected to explore the effects of transversal issues in greater detail.

Conclusion

Based on the theoretical and clinical significance, we constructed the BP-ANN model to anticipate changes of incisor and profile under comprehensive parameters, as well as identify potential significant factors prior to treatment. This approach will serve as a valuable reference for personalized aesthetic predictions, particularly in cases where exists uncertainty about the necessity of extraction or which kind of patient may benefit from profile changes. Furthermore, it could offer theoretical support for in-depth exploration of the potential correlations among the structures of craniofacial bones, teeth, and soft tissue.

Abbreviations

AI	Artificial intelligence
BP-ANN	Back-propagation artificial neural network model
MSE	Mean square error
MAE	Mean absolute error
R^2	Coefficient of determination
SHAP	SHapley Additive exPlanations
VTO	Visual Treatment Objective

Acknowledgements
Not applicable.

Author contributions
JP put forward the point, reviewed cephalometric tracings, analyzed and interpreted the data and was the major contributor in writing the first draft of the manuscript. YZ, MZ, YW and GD have made substantial contributions to the acquisition of the data; JL has made key contributions to the methodology, validation of results, and conduction of the investigation. JC provided the data resources and supervision of the draft. All authors have read and approved the final manuscript.

Funding
1. The General Guiding Project of Guangzhou Municipal Health Commission (20241A011096).
2. Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization (2021B1212040007).

Data availability
The datasets used and/or analysed during the study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Ethics approval was obtained from the Ethics Committee of Stomatology Hospital Affiliated to Guangzhou Medical University (20240809171826), and informed consent was acquired from each patient before starting orthodontic treatment.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Orthodontics, School and Hospital of Stomatology, Guangdong Engineering Research Center of Oral Restoration and Reconstruction & Guangzhou Key Laboratory of Basic and Applied Research of Oral Regenerative Medicine, Guangzhou Medical University, Guangzhou, China
[2]Department of Stomatology, LianZhou People's Hospital, Qingyuan, China
[3]Department of Clinical Research, The First Affiliated Hospital of Jinan University, Guangzhou, Guangdong 510630, China
[4]Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization, Guangzhou, Guangdong 510630, China

Received: 29 August 2024 / Accepted: 12 March 2025
Published online: 28 March 2025

References

1. Foo YZ, Simmons LW, Rhodes G. Predictors of facial attractiveness and health in humans. Sci Rep. 2017;7:39731.
2. Zhang Y, Zheng M, Wang X. Effects of facial attractiveness on personality stimuli in an implicit priming task: an ERP study. Neurol Res. 2016;38:685–91.
3. Cai J, Deng Y, Min Z, Zhang Y, Zhao Z, Jing D. Revealing the representative facial traits of different sagittal skeletal types: Decipher what artificial intelligence can see by Grad-CAM. J Dent. 2023;138:104701.
4. Tang X, Cai J, Lin B, Yao L, Lin F. Motivation of adult female patients seeking orthodontic treatment: an application of Q-methodology. Patient Prefer Adherence. 2015;9:249–56.
5. Christensen L, Luther F. Adults seeking orthodontic treatment: expectations, periodontal and TMD issues. Br Dent J. 2015;218:111–7.
6. Sample LB, Sadowsky PL, Bradley E. An evaluation of two VTO methods. Angle Orthod. 1998;68:401–8.
7. Maetevorakul S, Viteporn S. Factors influencing soft tissue profile changes following orthodontic treatment in patients with class II division 1 malocclusion. Prog Orthod. 2016;17:13.
8. Xing K, Mei H, Feng Q, Quan S, Zhang G, Jia A, Ge H, Mei D, Li J. Accuracy in predicting soft tissue changes of orthodontic class III cases using Dolphin® software. Clin Oral Investig. 2023;27:4531–9.
9. Zhang X, Mei L, Yan X, Wei J, Li Y, Li H, Li Z, Zheng W, Li Y. Accuracy of computer-aided prediction in soft tissue changes after orthodontic treatment. Am J Orthod Dentofac Orthop. 2019;156:823–31.
10. Nordblom NF, Büttner M, Schwendicke F. Artificial intelligence in orthodontics: critical review. J Dent Res. 2024;103:577–84.
11. Anic-Milosevic S, Medancic N, Calusic-Sarac M, Dumancic J, Brkic H. Artificial neural network model for predicting sex using dental and orthodontic measurements. Korean J Orthod. 2023;53:194–204.
12. Parrish M, O'Connell E, Eckert G, Hughes J, Badirli S, Turkkahraman H. Short- and Long-Term prediction of the Post-Pubertal mandibular length and Y-Axis in females utilizing machine learning. Diagnostics (Basel). 2023; 13.
13. Park YS, Choi JH, Kim Y, Choi SH, Lee JH, Kim KH, Chung CJ. Deep Learning-Based prediction of the 3D postorthodontic facial changes. J Dent Res. 2022;101:1372–9.
14. Tanikawa C, Yamashiro T. Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients. Sci Rep. 2021;11:15853.
15. Zhu J, Yang Y, Wong HM. Development and accuracy of artificial intelligence-generated prediction of facial changes in orthodontic treatment: a scoping review. J Zhejiang Univ Sci B. 2023;24:974–84.
16. Xie X, Wang L, Wang A. Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment. Angle Orthod. 2010;80:262–6.
17. Spee BTM, Leder H, Mikuni J, Scharnowski F, Pelowski M, Steyrl D. Using machine learning to predict judgments on Western visual Art along content-representational and formal-perceptual attributes. PLoS ONE. 2024;19:e0304285.
18. Mikuni J, Spee BTM, Forlani G, Leder H, Scharnowski F, Nakamura K, Watanabe K, Kawabata H, Pelowski M, Steyrl D. Cross-cultural comparison of beauty judgments in visual Art using machine learning analysis of Art attribute predictors among Japanese and German speakers. Sci Rep. 2024;14:15948.
19. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165:710–8.
20. Sadati Tilebon SM, Emamian SA, Ramezanpour H, Yousefi H, Özcan M, Naghib SM, Zare Y, Rhee KY. Intelligent modeling and optimization of titanium surface etching for dental implant application. Sci Rep. 2022;12:7184.
21. Xu Y, Goodacre R. On splitting training and validation set: A comparative study of Cross-Validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. J Anal Test. 2018;2:249–62.
22. Cai J, Min Z, Deng Y, Jing D, Zhao Z. Assessing the impact of occlusal plane rotation on facial aesthetics in orthodontic treatment: a machine learning approach. BMC Oral Health. 2024;24:30.
23. Liu X, Parhi KK. Molecular and DNA artificial neural networks via fractional coding. IEEE Trans Biomed Circuits Syst. 2020;14:490–503.
24. Al Turkestani N, Li T, Bianchi J, Gurgel M, Prieto J, Shah H, Benavides E, Soki F, Mishina Y, Fontana M, et al. A comprehensive patient-specific prediction model for temporomandibular joint osteoarthritis progression. Proc Natl Acad Sci U S A. 2024;121:e2306132121.
25. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc. 2020;92:807–12.

26. Che W-Q, Li Y-J, Tsang C-K, Wang Y-J, Chen Z, Wang X-Y, Xu A-D, Lyu J. How to use the surveillance, epidemiology, and end results (SEER) data: research design and methodology. Mil Med Res. 2023;10:50.
27. Xu Y, Zheng X, Li Y, Ye X, Cheng H, Wang H, Lyu J. Exploring patient medication adherence and data mining methods in clinical big data: A contemporary review. J Evid Based Med. 2023;16:342–75.
28. Raposo R, Peleteiro B, Paço M, Pinho T. Orthodontic camouflage versus orthodontic-orthognathic surgical treatment in class II malocclusion: a systematic review and meta-analysis. Int J Oral Maxillofac Surg. 2018;47:445–55.
29. Alhammadi MS, Almashraqi AA, Khadhi AH, Arishi KA, Alamir AA, Beleges EM, Halboub E. Orthodontic camouflage versus orthodontic-orthognathic surgical treatment in borderline class III malocclusion: a systematic review. Clin Oral Investig. 2022;26:6443–55.
30. Lu J, Wang Z, Zhang H, Xu W, Zhang C, Yang Y, Zheng X, Xu J. Bone graft materials for alveolar bone defects in orthodontic tooth movement. Tissue Eng Part B Rev. 2022;28:35–51.
31. Yamaguchi M, Fukasawa S. Is inflammation a friend or foe for orthodontic treatment?? inflammation in orthodontically induced inflammatory root resorption and accelerating tooth movement. Int J Mol Sci. 2021; 22.
32. Guo R, Tian Y, Li X, Li W, He D, Sun Y. Facial profile evaluation and prediction of skeletal class II patients during camouflage extraction treatment: a pilot study. Head Face Med. 2023;19:51.
33. Elias KG, Sivamurthy G, Bearn DR. Extraction vs nonextraction orthodontic treatment: a systematic review and meta-analysis. Angle Orthod. 2024; 94.
34. Peng J, Lei Y, Liu Y, Zhang B, Chen J. Effectiveness of micro-implant in vertical control during orthodontic extraction treatment in class II adults and adolescents after pubertal growth peak: a systematic review and meta-analysis. Clin Oral Investig. 2023;27:2149–62.
35. Oliver BM. The influence of lip thickness and strain on upper lip response to incisor Retraction. Am J Orthod. 1982;82:141–9.
36. Yogosawa F. Predicting soft tissue profile changes concurrent with orthodontic treatment. Angle Orthod. 1990;60:199–206.
37. Kinzinger GSM, Lisson JA, Buschhoff C, Hourfar J. Age-dependent effects on palate volume and morphology during orthodontic RME treatment. Clin Oral Investig. 2023;27:2641–52.
38. Nagaiwa M, Gunjigake K, Yamaguchi K. The effect of mouth breathing on chewing efficiency. Angle Orthod. 2016;86:227–34.
39. Inami T, Nakano Y, Miyazawa K, Tabuchi M, Goto S. Adult skeletal class II high-angle case treated with a fully customized lingual bracket appliance. Am J Orthod Dentofac Orthop. 2016;150:679–91.
40. Al-Rokhami RK, Sakran KA, Alhammadi MS, Mashrah MA, Cao B, Alsomairi MAA, Al-Worafi NA. Proximity of upper central incisors to incisive Canal among subjects with maxillary Dentoalveolar protrusion in various facial growth patterns. Angle Orthod. 2022;92:529–36.
41. Modarai F, Donaldson JC, Naini FB. The influence of lower lip position on the perceived attractiveness of chin prominence. Angle Orthod. 2013;83:795–800.
42. Qadeer TA, Jawaid M, Fahim MF, Habib M, Khan EB. Effect of lip thickness and competency on soft-tissue changes. Am J Orthod Dentofac Orthop. 2022;162:483–90.
43. Trisnawaty N, Ioi H, Kitahara T, Suzuki A, Takahashi I. Effects of extraction of four premolars on vermilion height and lip area in patients with bimaxillary protrusion. Eur J Orthod. 2013;35:521–8.
44. Zhou Q, Gao J, Guo D, Zhang H, Zhang X, Qin W, Jin Z. Three dimensional quantitative study of soft tissue changes in nasolabial folds after orthodontic treatment in female adults. BMC Oral Health. 2023;23:31.
45. Lee Y, Lim S-W, Chan V, Hong P, Han S-B, Chae HS. The surgical outcomes of anterior segmental osteotomy in Asian skeletal class II patients. Oral Maxillofac Surg. 2024;28:289–98.
46. Leavitt L, Volovic J, Steinhauer L, Mason T, Eckert G, Dean JA, Dundar MM, Turkkahraman H. Can we predict orthodontic extraction patterns by using machine learning? Orthod Craniofac Res. 2023;26:552–9.
47. Li P, Kong D, Tang T, Su D, Yang P, Wang H, Zhao Z, Liu Y. Orthodontic Treatment Planning based on Artificial Neural Networks. Sci Rep. 2019; 9:2037.
48. Sangcharearn Y, Ho C. Maxillary incisor angulation and its effect on molar relationships. Angle Orthod. 2007;77:221–5.

Publisher's note