

# Development and Evaluation of a Convolutional Neural Network for Automated Detection of Lip Separation on Profile and Frontal Photographs

Yasamin Vazirizadeh<sup>1, 2</sup>, Hoori Mirmohammadsadeghi<sup>3</sup>, Mohammad Behnaz<sup>3</sup>, Shahab Kavousinejad<sup>2, 3</sup>  

<sup>1</sup> Department of Orthodontics, Faculty of Dentistry, Shahed University, Tehran, Iran.

<sup>2</sup> Dentofacial Deformities Research Center, Research Institute for Dental Sciences, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

<sup>3</sup> Department of Orthodontics, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

 **Corresponding author:**

Shahab Kavousinejad, Department of Orthodontics, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran, Dentofacial Deformities Research Center, Research Institute for Dental Sciences, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

dr.shahab.k93@gmail.com

**Article History**

Received: 30 July 2024

Accepted: 30 Sep 2024

**Abstract**

**Background and Aim:** Lip incompetence is defined as a habitual gap of more than 3-4 mm between the lips at rest, which can contribute to oral health issues and malocclusions. This study aimed to propose a deep learning-based model for automatic detection of lip separation on orthodontic photographs.

**Materials and Methods:** This retrospective observational study employed a balanced dataset of 800 clinical images, comprising 400 cases of lip incompetence and 400 cases of lip competence. An auto-cropping technique based on averaged manual cropping coordinates was used to isolate the lip region. The cropped images were resized to 70×70 pixels and normalized before feeding into a novel attention-based residual connection convolutional neural network (ARN-CNN). The model incorporated both residual connections and attention modules to enhance feature learning and training stability. Data augmentation (e.g., rotation and scaling) was applied to improve generalizability. Training was conducted using 5-fold cross-validation, with an external test set to evaluate performance and reduce overfitting. Metrics such as accuracy, precision, recall, F1 score, receiver-operating characteristic curve-area under the curve (ROC-AUC), and a confusion matrix were used for performance evaluation.

**Results:** The ARN-CNN achieved 95% accuracy on the test set. For the competent class, precision was 0.97, recall was 0.94, and F1 score was 0.96. These values were 0.94, 0.96, and 0.95, respectively, for the incompetent class with an AUC of 0.98.

**Conclusion:** The ARN-CNN model effectively identified lip incompetence, highlighting the potential of deep learning to support orthodontic diagnosis through image-based analysis.

**Keywords:** Artificial Intelligence; Deep Learning; Malocclusion; Convolutional Neural Networks; Orthodontics

**Cite this article as:** Vazirizadeh Y, Mirmohammadsadeghi H, Behnaz M, Kavousinejad S. Development and Evaluation of a Convolutional Neural Network for Automated Detection of Lip Separation on Profile and Frontal Photographs. *J Res Dent Maxillofac Sci.* 2025; **10(2):111-124.**

## Introduction

Facial soft tissue analysis is essential for evaluation of maxillofacial growth. The correlation of soft tissue indices and skeletal and occlusal changes in malocclusions has been previously documented [1], indicating that soft tissue analysis can provide information about skeletal and dental abnormalities, serving as a diagnostic tool. However, orthodontists must consider soft tissue adaptation of patients in their treatment plans, considering soft tissue limitations in terms of esthetics, stability, and function [2].

One critical aspect of facial soft tissue analysis is assessment of lip competence and the role of lip incompetence in growth and development of the craniofacial complex [3, 4]. Incompetent lips are characterized by lip separation by more than 3-4 mm [5], which can lead to significant oral complications due to inadequate lip sealing [6]. A relationship has been identified between incompetent lips and malocclusions, including vertical and sagittal skeletal and dental discrepancies [4]. In patients with normal occlusion who exhibit incompetent lips, dentofacial morphology is likely to be the underlying cause [1]. Thus, assessment of incompetent lips can provide important insights into the overall appearance and structure of the face.

Various methods, such as visual examination, cephalometric radiographs, and photography, are used to evaluate lip sealing [1, 7-11]. Recently, application of artificial intelligence (AI) has gained popularity for enhancement of the accuracy and efficiency in diagnosis and treatment planning [12]. AI, as a machine learning technology, learns from data and autonomously solves problems, offering rapid diagnosis and treatment planning capabilities. Deep learning, a subset of machine learning, utilizes multi-layered neural networks to automatically learn and represent complex

patterns and features from large datasets, enabling advanced capabilities in tasks such as image recognition, natural language processing, and more [13]. Convolutional neural networks (CNNs) are a class of deep learning models specifically designed for image processing [14]. By leveraging convolutional layers, CNNs automatically extract and learn hierarchical features from raw image data, enabling the detection of complex patterns and structures. Their ability to capture spatial hierarchies makes them highly effective for tasks like image classification and object detection. One challenge encountered in measuring the magnitude of lip separation (which is considered abnormal if it exceeds 4 mm) on photographs is lack of a suitable millimeter-scale reference. Therefore, it would be ideal to develop a deep learning-based method for detection of lip incompetence on photographs.

AI has shown promising applications in orthodontics, particularly in analyzing clinical photographs and assisting with diagnosis and treatment planning. AI models, especially CNNs, have demonstrated high accuracy in classifying orthodontic photographs according to their orientations [15]. These systems can also aid in detecting landmarks, categorizing dental crowding, and determining the necessity of tooth extraction with impressive precision [16]. AI has achieved state-of-the-art results in various orthodontic applications, including automated landmark detection on lateral cephalograms, skeletal classification, and decision-making regarding tooth extractions [17]. While AI shows potential to enhance orthodontic care by saving time and providing accuracy comparable to trained dentists, challenges remain in generalizability and standardization across studies [18]. As the field progresses, researchers are working towards implementing AI into clinical orthodontic workflows and addressing real-world evaluation concerns.

Considering the limited number of studies on the automatic detection of soft tissue problems on photographs using deep learning, this study was conducted to investigate lip competence through deep learning models, and also contribute to development of a fully automated system for generating a problem list in orthodontics. As a preliminary step, the present study aimed to develop a CNN model for automatic detection of lip incompetence on frontal and profile photographs.

## Materials and Methods

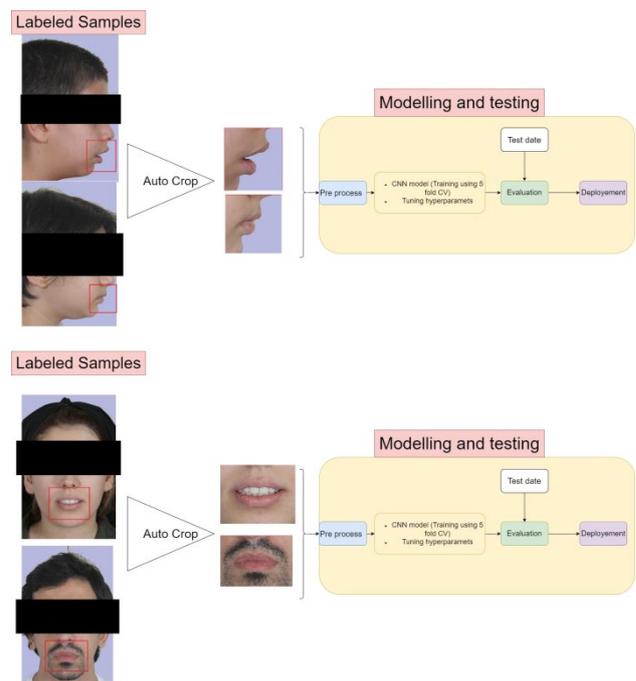
### Dataset preparation:

This study was a retrospective observational study with a deep learning component. It was approved by the ethics committee of Shahid Beheshti University of Medical Sciences (IR.SBMU.DRC.REC.1402.102). Images that met the standard criteria for orthodontic photography were selected from the archives of the Orthodontics Department of Shahid Beheshti Dental School. These criteria included specific angles, natural head position, appropriate lighting, and high resolution, ensuring that all critical facial features, particularly the mouth and its surrounding areas, were clearly visible [19, 20]. Each image was thoroughly reviewed and validated to ensure that it met the necessary standards for precise analysis.

A dataset of 800 images (at rest, natural head position) was created, comprising of 400 images of patients with lip incompetence and 400 patients with lip competence. The dataset included patients between 8 to 50 years of age, including 57% females and 43% males. Lip incompetence was defined as lip separation by more than 4 mm, based on the patients' orthodontic records and direct clinical measurements. This process was validated by three experienced orthodontists. Next, the images were labeled according to the lip separation status (incompetent or competent) by

three orthodontists, and a consensus approach was used to confirm labeling accuracy.

All data were utilized for model training and validation through 5-fold cross-validation. To further evaluate the model's performance and mitigate overfitting, an additional test set comprising of 61 new profile images (25 images with lip incompetence and 36 images with lip competence) and 56 new frontal images (31 images with lip incompetence and 25 images with lip competence) were collected after finalizing the model and hyperparameters. These test samples included images sourced from Shahid Beheshti Dental School. By separating the test data from the training and validation sets used in the 5-fold cross-validation, we ensured data leakage prevention and robust evaluation of the model's generalizability (Figure 1).



**Figure 1.** Pipeline of the study

Importantly, the test data were evaluated only once to avoid any possibility of tuning the model based on the test results, thereby maintaining the integrity of the model's performance assessment. This approach ensured that the final evaluation reflected the model's true ability to generalize to

new, unseen data without any bias introduced by repeated testing or adjustments.

#### *Preprocessing:*

The image data underwent preprocessing. Both frontal and profile images contained significant extraneous regions outside the lip area. Therefore, we utilized a technique called "auto-crop" to separate the lip region (region of interest) on both frontal and profile images. To implement automatic cropping, we initially identified a rectangular region around the lip area on each image manually. Specifically, we determined a rectangular region within the lip area for each image (defined by 4 coordinates, including X and Y). Subsequently, we computed the average coordinates of these rectangle corners across the training samples to serve as the final coordinates for automatic image cropping. For each dimension (X and Y), this involved:

- Calculating the mean start and end X coordinates relative to image width across all images.
- Calculating the mean start and end Y coordinates relative to image height across all images.

Considering potential variations in aspect ratios of the images, we used the following formula for automatic cropping of the lip area on profile and frontal images:

- Frontal cropping area (rectangle): The average coordinates for start and end points on the image were defined as follows:
  - Start x = image width  $\times$  0.30
  - End x = image width  $\times$  0.72
  - Start y = image height  $\times$  0.60
  - End y = image height  $\times$  0.85
- Profile cropping area (rectangle): The average coordinates for start and end points on the image were defined as follows:
  - Start x = image width  $\times$  0.65
  - End x = image width  $\times$  0.92

- Start y = image height  $\times$  0.64
- End y = image height  $\times$  0.91

These values were determined by averaging the manually defined cropping coordinates across the training dataset. The formulae use relative proportions within the image rather than absolute dimensions. This ensures that, even with changes in aspect ratio, the cropped region remains close to the target area, effectively handling variations in image aspect ratio. These dimensions were derived from the average cropping regions observed in the training set of frontal and profile images. For the test images, cropping was performed automatically based on these average coordinates, without manual adjustment, ensuring consistency with the training phase. This approach enabled simultaneous evaluation of the automatic cropping process and the model's performance. Therefore, the auto-crop technique was employed while adhering to the standard principles of orthodontic photography on profile and frontal images (including head position, etc.). This method has limitations, which are discussed in the Discussion section.

As the next step in the preprocessing pipeline, the images were converted to grayscale and then enhanced using contrast-limited adaptive histogram equalization (CLAHE) to mitigate contrast variations. CLAHE is a technique that enhances the contrast of images by adapting histogram equalization to local regions and limiting amplification of noise, which improves the visibility of features in both bright and dark areas [21]. CLAHE can enhance CNN learning as part of the data preprocessing pipeline, especially in cases where the data suffer from low contrast or uneven illumination [22-24]. Subsequently, the images were resized to 70x70 pixels and normalized (scaling pixel values between 0 and 1) to prepare them for input into the CNN. We also applied various data augmentation

techniques, including rotation, horizontal flipping (for frontal images only), translation, scaling, and shearing (Figure 2).

*Model architecture:*

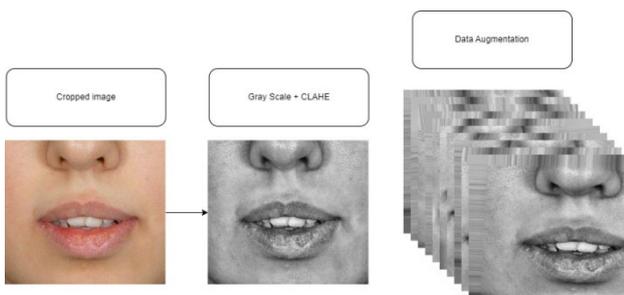
This study employed a hybrid architecture that integrates attention modules with residual connections [25] (ARN-CNN) (Figure 3). This model consists of two primary components:

*1. Attention module:*

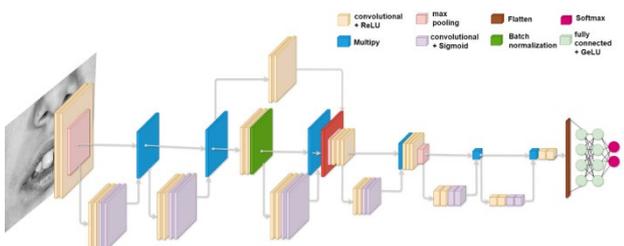
The attention module is designed to enhance feature representation by applying an attention mechanism. It includes Conv2D layers with specified filters, followed by ReLU and sigmoid activation functions. The attention mechanism is implemented by performing element-wise multiplication between the processed features and the original input.

*2. Residual block:*

Residual blocks are constructed using Conv2D layers with varying filter sizes and batch normalization. Each residual block is followed by the application of the attention module. Residual connections are incorporated to sum the output of Conv2D layers with the input tensor, facilitating improved gradient flow and feature learning.



**Figure 2.** Data augmentation techniques applied in this study



**Figure 1.** Proposed architecture. The final Dense layer reaches two probabilities: incompetence and competence

The model processes inputs through a series of Conv2D layers, MaxPooling2D operations, and attention modules. Residual blocks, which form the central part of the network, progressively increase the filter size and count. The network concludes with Flatten, Dense, and Dropout layers to generate the final output, employing a softmax activation function for classification purposes. For comparative evaluation, we also utilized the ResNet50 architecture with fine-tuned transfer learning. Regularization techniques were applied to control model complexity and prevent overfitting. Specifically, L2 regularization was applied to the Dense layer kernels to constrain the model's complexity, while L1 regularization was used for the activities and biases of the Dense layers to maintain simplicity and enhance generalization. These regularization strategies were implemented during training to improve model performance and reduce overfitting.

*Model training and evaluation:*

The coding for this project and training-validation were conducted in Python 3.8 using TensorFlow 2.6 and Keras 2.4, utilizing a NVIDIA GeForce RTX 3090 graphics card. For validation, the 5-fold cross-validation method was employed. In this approach, the data were divided into 5 equal parts. In each training iteration, one of these parts served as the test data, while the remaining parts were used for training of the CNN model. The model was trained on the training data and evaluated on the test data. This process was repeated 5 times, with each part used once as the test data, and the average performance across these 5 iterations was reported as the validation metric.

This method was beneficial for hyperparameter tuning, which was performed using the Grid search technique. This involved optimizing parameters such as the number and size of filters

in Conv2D layers and attention modules, selecting the optimal optimizer for training the network (e.g., Adamax or Adam), learning rate determination, setting dropout rates to mitigate overfitting, and choosing kernel sizes for feature extraction from images. A dropout rate of 0.45 and a learning rate of 0.001 were selected. These parameters were tuned by executing the model across all possible combinations and comparing the resulting accuracies on different validation sets.

To prevent data leakage and ensure that the test data did not influence the model training process, the following measures were implemented: Test data were completely separated from the training and validation data. After finalizing the model and hyperparameters, independent test data were collected from other centers to ensure no influence or tuning based on the test data. The test data were evaluated only once to prevent model optimization based on test results, preserving the integrity of the final assessment. For the final evaluation, new and independent data were used, which were not involved in the training or validation phases at any point.

Finally, the model was tested only once on separate test datasets collected independently from a dedicated center. The automatic cropping successfully cropped the correct region in all test data, clearly isolating the lip area. These datasets were classified by three orthodontists based on lip competence or incompetence. The model's performance on these test datasets reflected its final accuracy in this study. For the final analysis and evaluation of the model's performance, the following metrics were used:

- **Accuracy:** The ratio of correctly classified samples to the total number of samples. Accuracy serves as the primary metric for evaluation of model performance.

- **Confusion matrix:** Provides a detailed analysis of the results, including the number of true positives (TPs), false positives (FPs), true negatives (TNs), and false negatives (FNs) for each class.
- **Precision and recall:** Calculated for each class, these metrics help evaluate the model's ability to correctly identify positive and negative samples.
- **F1-score:** The harmonic mean of precision and recall, providing a balanced measure between the two.
- **Receiver-operating characteristic (ROC) curve and area under the curve (AUC):** The ROC curve and the AUC assess the model's performance in distinguishing between classes [26].
- **Log loss:** Log loss is a metric used to evaluate the performance of probabilistic classification models, measuring the discrepancy between predicted probabilities and actual outcomes, with lower values indicating better model accuracy [27].
- **Gradient-weighted class activation mapping (Grad CAM):** Grad CAM is a technique used in CNNs to visualize and interpret which regions of an input image are most influential in predicting the target class, providing insights into the model's decision-making process [28, 29].

## Results

Table 1 presents the performance of ARN-CNN and ResNet50 models (profile and frontal images) on the test datasets. For profile images, ARN-CNN achieved a log loss of 0.17, F1-score of 0.96, recall of 0.94, and precision of 0.97 for competent lips; while, ResNet50 had a log loss of 0.18, F1-score of 0.96, recall of 0.97, and precision of 0.95. For incompetent lips, ARN-CNN scored a F1-score of 0.94, while ResNet50 scored 0.94. For

frontal images, ARN-CNN had a log loss of 0.05, F1-score of 0.98, recall of 1.00, and precision of 0.96 for competent lips; while ResNet50 had a log loss of 0.08, F1-score of 0.98, recall of 1.00, and precision of 0.96. Both models showed similar performance, with ARN-CNN slightly outperforming ResNet50 in frontal view.

Figures 4 and 5 correspond to the profile and frontal model performances, respectively. In Figures 4A and 4B, the learning curve graphs illustrate training and validation metrics including accuracy and loss for the profile classification model across epochs. Figure 4C displays a confusion matrix highlighting three misclassified samples among all test instances. An example of misclassification is provided in Figure 4C, showing the probability percentage associated with competence and incompetence lips to be 57% and 43%, respectively. Figure 4D showcases the ROC curve and AUC value, providing a comprehensive evaluation of model performance. Figure 4E presents the model's confidence scores for each test sample, with incorrect classifications marked by red circles. The overall confidence level for the test data classification was 95%. In Figure 4's Grad CAM section, the images reveal the specific areas where the model concentrates its attention on selected test data samples. Additionally, in the Feature Maps section, an example of model input and output feature maps of each convolutional

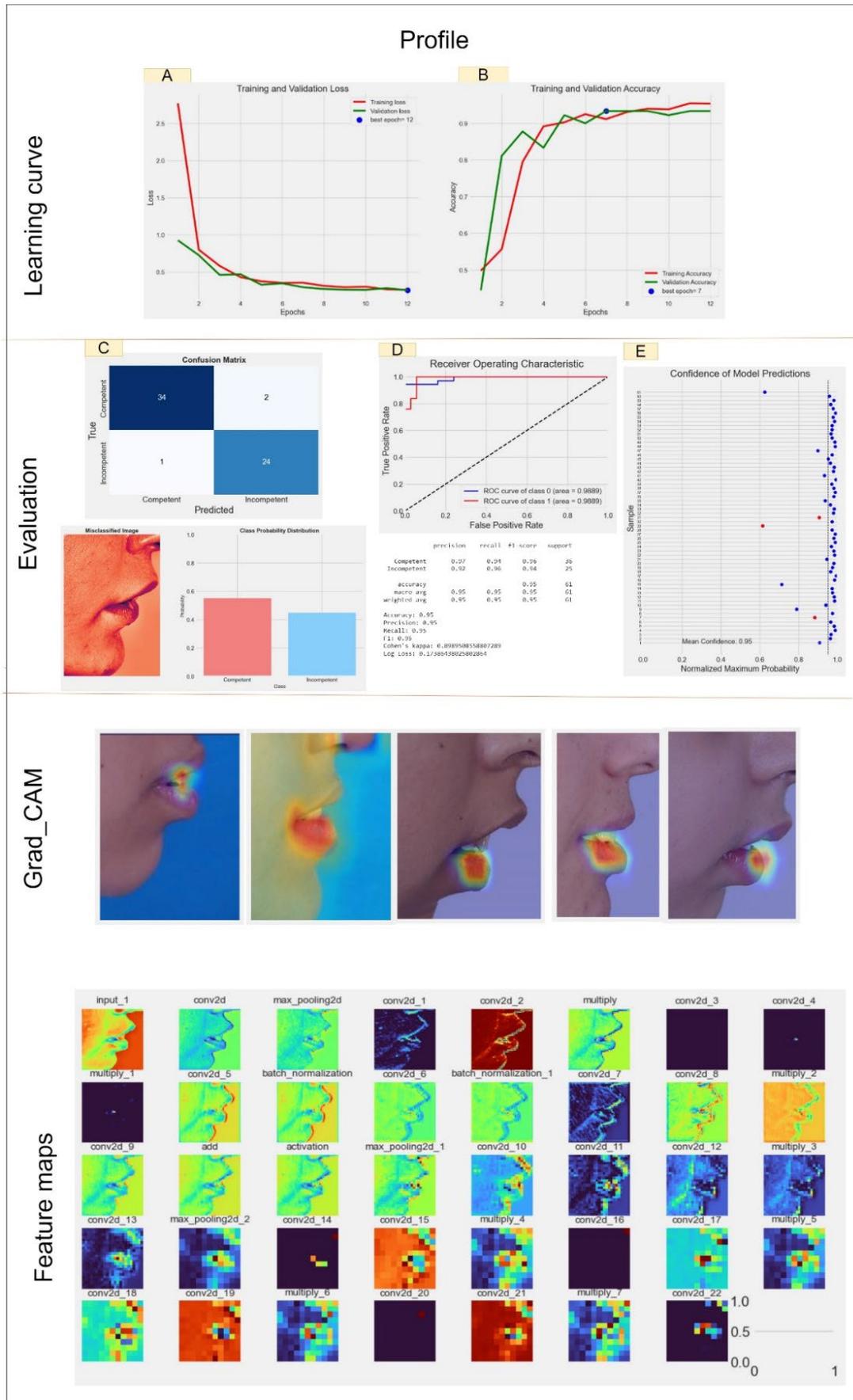
layer is presented for interpretation of model behavior.

Figures 5A and 5B display the learning curve graphs for accuracy and loss during training and validation epochs for the frontal image classification model. Figure 5C exhibits a confusion matrix indicating one misclassified sample among all test instances. An example of misclassification is provided in Figure 5C, where the probability percentage related to competent and incompetent lips was 41% and 59%, respectively (correct label: incompetent). Figure 5D presents the ROC curve and AUC value, providing an assessment of model performance on the frontal image classification task. Figure 5E shows the model's confidence scores for each test sample, with incorrect classifications marked by red circles. The overall confidence level for the test data classification was 97%. In Figure 5's Grad CAM section, the images highlight the regions where the model directs its attention to selected test data samples. Similarly, in the Feature Maps section, an example of model input and output feature maps of each convolutional layer is provided for interpretation of model behavior.

Figure 6 shows a comparison of Grad CAM across different models. It illustrates the regions of input images that each model emphasizes when making predictions, providing a visual comparison of how attention is allocated among different models.

**Table 1.** Performance of ARN-CNN and ResNet50 models (profile and frontal) on the test datasets

	Model	Class	Accuracy	Precision	Recall	F1-Score	Log loss
Profile	ARN-CNN	Competent	0.95	0.97	0.94	0.96	0.17
		Incompetent		0.92	0.96	0.94	
	Resnet50	Competent	0.95	0.95	0.97	0.96	0.18
		Incompetent		0.96	0.92	0.94	
Frontal	ARN-CNN	Competent	0.98	0.96	1.00	0.98	0.05
		Incompetent		1.00	0.97	0.98	
	Resnet50	Competent	0.98	0.96	1.00	0.98	0.08
		Incompetent		1.00	0.97	0.98	



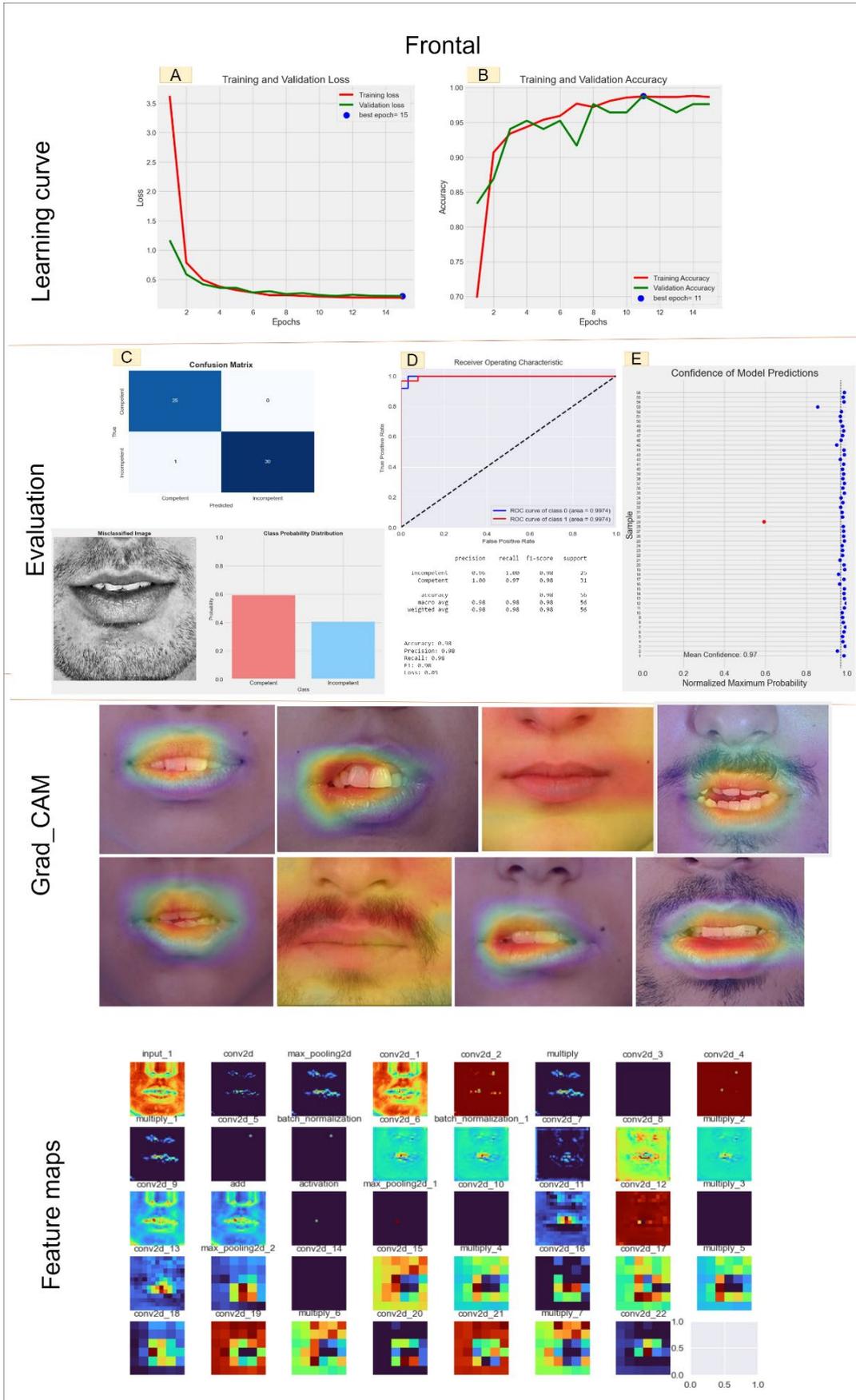
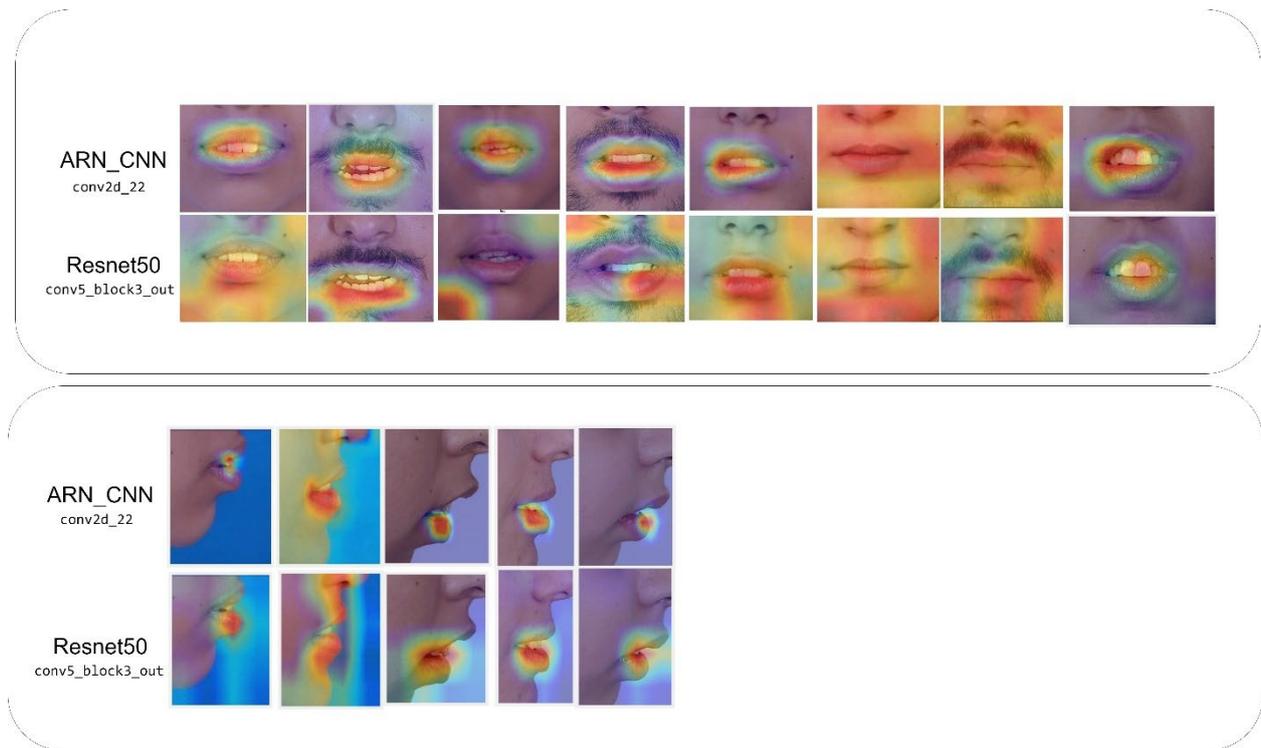


Figure 5. Performance of the frontal image classification model on the test data



**Figure 6.** Comparison of Grad CAM across different models

## Discussion

This study aimed to design and develop an AI model for automatic detection of lip incompetence, with performance evaluated on a separate test dataset. The results demonstrated that the model was more efficient for detecting lip incompetence on the frontal view compared to the profile view. Although both architectures showed similar performance levels, Grad CAM visualizations revealed that the ARN-CNN architecture specifically focused on lip areas when detecting lip separation. The trained model of this study was integrated into an orthodontic diagnostic software developed in C# within the Visual Studio environment.

Lip incompetence, which exposes the teeth and increases the risk of traumatic dental injuries [30], can also lead to gingivitis and periodontal disease due to constant mouth breathing [6]. Proper lip closure is essential for maintaining oral health by preventing excessive gingival and tooth show. Clinicians addressing lip incompetence should consider factors such as

airway obstructions, protruded upper incisors, and skeletal abnormalities, as correcting these issues can enhance lip competence and overall oral function. AI algorithms, when trained on extensive datasets of lip incompetence cases, have the potential to improve the detection of subtle signs associated with this condition [31]. By analyzing facial images, including lip position and movement, AI systems could offer valuable diagnostic insights and quantify lip parameters such as height, symmetry, and dynamics, complementing traditional clinical evaluations.

To remove extraneous regions from the images, we utilized an automatic cropping (auto-crop) technique. Despite its effectiveness, the automatic cropping method may struggle with extreme variations in facial positioning or image quality, potentially leading to suboptimal cropping in certain cases. Additionally, reliance on fixed proportional coordinates may not account for all individual variations in lip positioning, which could affect accuracy. If the images do not adhere to the standards of

orthodontic photography, this issue may become more pronounced. Comparison of our automatic image cropping method with Al-Mahadeen et al.'s [32] approach reveals notable differences. Our method relies on fixed relative proportions based on image dimensions, ensuring consistent cropping across varying aspect ratios and simplifying implementation. Conversely, Al-Mahadeen et al.'s method [32] employs pixel-based object detection to define cropping coordinates, which can achieve higher accuracy in targeting specific features but involves greater complexity. However, this method successfully performed automatic cropping on all test data, ensuring consistency and accuracy throughout the testing phase. For future studies, it is recommended to use the MediaPipe model [33] for more accurate facial landmark detection and cropping, which can adapt better to individual variations and maintain higher precision in diverse image conditions. In addition to our method and Al-Mahadeen et al.'s approach [32], other prominent image cropping methods include R-CNN, Fast R-CNN, Faster R-CNN, YOLO, and deformable part models, each offering distinct advantages [34]. However, the method discussed in the present study performed well on separate test data and appears well-suited for orthodontic photography, adhering to a specific standard.

Several studies have explored AI accuracy in orthodontics [35, 36]. For example, Yu et al. [37] developed a robust skeletal diagnostic system using CNNs with lateral cephalograms. Nakornnoi and Chanmanee [38] assessed the accuracy of Digital Imaging software in predicting soft tissue changes during orthodontic treatments, including non-extraction and extraction treatments, and orthognathic surgery. Jeong et al. [39] demonstrated that CNNs could accurately assess soft tissue profiles for orthognathic surgery using only facial photographs, which aligns with our findings.

However, their study would benefit from a larger sample size to enhance accuracy. Tanikawa and Yamashiro [40] confirmed the clinical acceptability of AI systems in predicting facial morphology post-treatment, consistent with our results. The novelty of our research lies in its focus on AI for detecting lip incompetence, an area previously unexplored. Our study utilized deep learning techniques that are well-recognized and widely applied in both research and industry. Reinforcement learning algorithms, which have been utilized in the medical field to optimize clinical diagnostic algorithms, hold significant potential for advancing clinical orthodontics in the future [41]. A significant challenge is the availability of extensive datasets, which is crucial for training effective AI models due to ethical and confidentiality concerns [42].

The present study highlights the potential of deep learning for automated lip competence analysis using orthodontic photographs. Given that these models are classification-based, future research should consider exploring regression methods and key point detection techniques for enhanced accuracy. Further research should evaluate the method's efficacy and accuracy in patients with malocclusions, variations in vertical dimension, and craniofacial issues. Expanding this analysis to include frontal and profile images, as well as cephalometric radiographs, with larger sample sizes, could improve predictive models and explore changes in orthodontic metrics. AI's role in medicine is evolving from associative inference to predictive treatment planning, and enhancing the interpretability of deep learning models will provide deeper insights into these predictions. AI models require robust training data and validation to ensure reliable performance. Ethical considerations, patient privacy, and informed consent are paramount in clinical AI applications. AI should complement, not replace, clinical judgment, and collaboration

with experts in AI and oral medicine is crucial for successful implementation [21, 43].

## Conclusion

This study demonstrated the significant potential of deep learning algorithms for automated detection of lip incompetence on orthodontic photographs. While AI presents promising advancements in the analysis and detection of lip incompetence, ongoing research and development are essential to refine these models and ensure their practical utility in clinical settings.

## Acknowledgment

This study was part of a thesis for a postgraduate degree in Orthodontics at Shahid Beheshti University of Medical Sciences, Tehran, Iran. The study protocol was approved by the institutional Ethics Committee under the code IR.SBMU.DRC.REC.1402.102. Funding for this study was provided by the Dentofacial Deformities Research Center at the Research Institute for Dental Sciences, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

## Author Contributions

Author Y.V. conducted the literature review and contributed to the methodology section. Author H.M. assisted in data collection and analysis and contributed to the discussion. Author M.B. supervised the orthodontic aspects of the study and contributed to the interpretation of results. Author S.K. led the AI-related work, including the development and implementation of the convolutional neural network for image analysis, and contributed to the overall writing of the manuscript.

## References

1. Kook MS, Jung S, Park HJ, Oh HK, Ryu SY, Cho JH, Lee JS, Yoon SJ, Kim MS, Shin HK. A comparison study of different facial soft tissue analysis methods. *J Craniomaxillofac Surg*. 2014 Jul;42(5):648-56.
2. Siécola GS, Capelozza L Filho, Lorenzoni DC, Janson G, Henriques JFC. Subjective facial analysis and its correlation with dental relationships. *Dental Press J Orthod*. 2017 Mar-Apr;22(2):87-94.
3. Drevensek M, Stefanac-Papić J, Farcnik F. The influence of incompetent lip seal on the growth and development of craniofacial complex. *Coll Antropol*. 2005 Dec;29(2):429-34.
4. Leonardo SE, Sato Y, Kaneko T, Yamamoto T, Handa K, Iida J. Differences in dento-facial morphology in lip competence and lip incompetence. *Orthod Waves*. 2009 Mar;68(1):12-9.
5. Proffit WR, Fields H, Larson B, Sarver DM. *Contemporary orthodontics*. 5th ed. St. Louis: Elsevier; 2018.
6. Hassan AH, Turkistani AA, Hassan MH. Skeletal and dental characteristics of subjects with incompetent lips. *Saudi Med J*. 2014 Aug;35(8):849-54.
7. Siécola GS, Capelozza L Filho, Lorenzoni DC, Janson G, Henriques JFC. Subjective facial analysis and its correlation with dental relationships. *Dental Press J Orthod*. 2017 Mar-Apr;22(2):87-94.
8. Zacharopoulos GV, Manios A, Kau CH, Velagrakis G, Tzanakakis GN, de Bree E. Anthropometric Analysis of the Face. *J Craniofac Surg*. 2016 Jan;27(1):e71-5.
9. Juerchott A, Saleem MA, Hilgenfeld T, Freudlsperger C, Zingler S, Lux CJ, Bendszus M, Heiland S. 3D cephalometric analysis using Magnetic Resonance Imaging: validation of accuracy and reproducibility. *Sci Rep*. 2018 Aug 29;8(1):13029.
10. Ahmed M, Shaikh A, Fida M. Diagnostic validity of different cephalometric analyses for assessment of the sagittal skeletal pattern. *Dental Press J Orthod*. 2018 Sep-Oct;23(5):75-81.
11. Sukhia RH, Nuruddin R, Azam SI, Fida M. Predicting the sagittal skeletal pattern using dental cast and facial profile photographs in children aged 9 to 14 years. *J Pak Med Assoc*. 2022 Nov;72(11):2198-203.
12. Monill-González A, Rovira-Calatayud L, d'Oliveira NG, Ustrell-Torrent JM. Artificial intelligence in orthodontics: Where are we now? A scoping review. *Orthod Craniofac Res*. 2021 Dec;24 Suppl 2:6-15.
13. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015 May 28;521(7553):436-44.
14. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25.

15. Ryu J, Lee YS, Mo SP, Lim K, Jung SK, Kim TW. Application of deep learning artificial intelligence technique to the classification of clinical orthodontic photos. *BMC Oral Health*. 2022 Oct 25;22(1):454.
16. Ryu J, Kim YH, Kim TW, Jung SK. Evaluation of artificial intelligence model for crowding categorization and extraction diagnosis using intraoral photographs. *Sci Rep*. 2023 Mar 30;13(1):5177.
17. Mohammad-Rahimi H, Nadimi M, Rohban MH, Shamsoddin E, Lee VY, Motamedian SR. Machine learning and orthodontics, current trends and the future opportunities: A scoping review. *Am J Orthod Dentofacial Orthop*. 2021 Aug;160(2):170-192.e4.
18. Nordblom NF, Büttner M, Schwendicke F. Artificial Intelligence in Orthodontics: Critical Review. *J Dent Res*. 2024 Jun;103(6):577-84.
19. Ahmad I. Digital dental photography. Part 7: extra-oral set-ups. *Br Dent J*. 2009 Aug 8;207(3):103-10.
20. Samawi SS. Clinical Digital Photography in Orthodontics. *Jordan Dental Journal*. 2011;18(1).
21. Zuiderveld K. Contrast limited adaptive histogram equalization. In: *Graphics Gems IV*. 1994. p. 474-85.
22. Alwakid G, Gouda W, Humayun M. Deep Learning-Based Prediction of Diabetic Retinopathy Using CLAHE and ESRGAN for Enhancement. *Healthcare (Basel)*. 2023 Mar 15;11(6):863.
23. Chakraverti S, Agarwal P, Pattanayak HS, Chauhan SP, Chakraverti AK, Kumar M. De-noising the image using DBST-LCM-CLAHE: A deep learning approach. *Multimed Tools Appl*. 2024;83(4):11017-42.
24. Hayati M, Muchtar K, Maulina N, Syamsuddin I, Elwirehardja GN, Pardamean B. Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning. *Procedia Comput Sci*. 2023;216:57-66.
25. Kavousinejad S. An Attention-Based Residual Connection Convolutional Neural Network for Classification Tasks in Computer Vision. *Journal of Dental School, Shahid Beheshti University of Medical Sciences*. 2024 Jan 1;42(1):14-25.
26. Narkhede S. Understanding AUC-ROC curve. *Towards Data Science*. 2018;26(1):220-7.
27. Vovk V. The fundamental nature of the log loss function. *Fields of logic and computation II: Essays dedicated to Yuri Gurevich on the Occasion of His 75th Birthday*. 2015:307-18.
28. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision 2017* (pp. 618-626).
29. Wang J, Khan MA, Wang S, Zhang Y. SNSVM: SqueezeNet-Guided SVM for Breast Cancer Diagnosis. *Comput Mater Contin*. 2023 Aug 30;76(2):2201-2216.
30. Wig M, Kumar A, Chaluvaiiah MB, Yadav V, Shyam R. Lip incompetence and traumatic dental injuries: a systematic review and meta-analysis. *Evid Based Dent*. 2022 Jul 11.
31. Patil S, Albogami S, Hosmani J, Mujoo S, Kamil MA, Mansour MA, Abdul HN, Bhandi S, Ahmed SSSJ. Artificial Intelligence in the Diagnosis of Oral Diseases: Applications and Pitfalls. *Diagnostics (Basel)*. 2022 Apr 19;12(5):1029.
32. Al-Mahadeen B, AlTarawneh M, AlTarawneh IH. Signature region of interest using auto cropping. *arXiv preprint arXiv:10043549*. 2010.
33. Lugaresi C, Tang J, Nash H, C Clanahan C, Uboweja E, Hays M, et al. Mediapipe: a framework for building perception pipelines. *arXiv preprint arXiv:190608172*. 2019.
34. Yan J, Lin S, Bing Kang S, Tang X. Learning the change for automatic image cropping. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2013* (pp. 971-978). Chang Q, Bai Y, Wang S, Wang F, Wang Y, Zuo F, Xie X. Automatic soft-tissue analysis on orthodontic frontal and lateral facial photographs based on deep learning. *Orthod Craniofac Res*. 2024 Jul 5.
35. Chang Q, Bai Y, Wang S, Wang F, Wang Y, Zuo F, Xie X. Automatic soft-tissue analysis on orthodontic frontal and lateral facial photographs based on deep learning. *Orthod Craniofac Res*. 2024 Dec;27(6):893-902.
36. Surendran A, Daigavane P, Shrivastav S, Kamble R, Sanchla AD, Bharti L, Shinde M. The Future of Orthodontics: Deep Learning Technologies. *Cureus*. 2024 Jun 10;16(6):e62045.
37. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J. Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. *J Dent Res*. 2020 Mar;99(3):249-56.
38. Nakornnoi T, Chanmanee P. Accuracy of Digital Imaging Software to Predict Soft Tissue Changes during Orthodontic Treatment. *J Imaging*. 2024 May 31;10(6):134.

39. Jeong SH, Yun JP, Yeom HG, Lim HJ, Lee J, Kim BC. Deep learning based discrimination of soft tissue profiles requiring orthognathic surgery by facial photographs. *Sci Rep.* 2020 Oct 1;10(1):16235.
40. 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 Aug 4;11(1):15853.
41. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature.* 2015 Feb 26;518(7540):529-33.
42. Rousseau M, Retrouvey JM. Machine learning in orthodontics: Automated facial analysis of vertical dimension for increased precision and efficiency. *Am J Orthod Dentofacial Orthop.* 2022 Mar;161(3):445-50.
43. Kavousinejad S, Ameli-Mazandarani Z, Behnaz M, Ebadifar A. A Deep Learning Framework for Automated Classification and Archiving of Orthodontic Diagnostic Documents. *Cureus.* 2024 Dec 28;16(12):e76530.