Head & Face Medicine

RESEARCH

Comparison of deep learning models to detect crossbites on 2D intraoral photographs

Beatrice Noeldeke[1], Stratos Vassis[2*], Mohammedreza Sefidroodi[2], Ruben Pauwels[3,4] and Peter Stoustrup[2]

Abstract

Background To support dentists with limited experience, this study trained and compared six convolutional neural networks to detect crossbites and classify non-crossbite, frontal, and lateral crossbites using 2D intraoral photographs.

Methods Based on 676 photographs from 311 orthodontic patients, six convolutional neural network models were trained and compared to classify (1) non-crossbite vs. crossbite and (2) non-crossbite vs. lateral crossbite vs. frontal crossbite. The trained models comprised DenseNet, EfficientNet, MobileNet, ResNet18, ResNet50, and Xception.

Findings Among the models, Xception showed the highest accuracy (98.57%) in the test dataset for classifying non-crossbite vs. crossbite images. When additionally distinguishing between lateral and frontal crossbites, average accuracy decreased with the DenseNet architecture achieving the highest accuracy among the models with 91.43% in the test dataset.

Conclusions Convolutional neural networks show high potential in processing clinical photographs and detecting crossbites. This study provides initial insights into how deep learning models can be used for orthodontic diagnosis of malocclusions based on intraoral 2D photographs.

Keywords Artificial intelligence, Orthodontic diagnosis, Deep learning, Neural networks, Crossbite

Introduction

Artificial Intelligence (AI) approaches are gaining increased attention to support orthodontic diagnosis and treatment planning [1]. A domain yet to be explored is the integration of AI to assist dentists and pediatricians to accurately diagnose orthodontic treatment need. This particular application is promising because dentists and pediatricians often play a crucial role in the initial diagnosis of malocclusion which prompts for referral of the patients to the orthodontic specialist [2]. However, research indicates that in up to 45% of the cases, the initial referral is incorrect due to inadequate application of national indication systems that determine treatment needs such as the British Index of Orthodontic Treatment Need (IOTN) [3]. Inadequate application is also

*Correspondence:
Stratos Vassis
stratos.vassis@dent.au.dk
[1]Leibniz University of Hanover, Königsworther Platz 1, 30167 Hanover, Germany
[2]Section of Orthodontics, Department of Dentistry and Oral Health, Aarhus University, Vennelyst Boulevard 9, Aarhus C, Aarhus 8000, Denmark
[3]Department of Dentistry and Oral Health, Aarhus University, Vennelyst Boulevard 9, Aarhus C, Aarhus 8000, Denmark
[4]Department of Radiology, Faculty of Dentistry, Chulalongkorn University, 34 Henri-Dunant Road, Wangmai, Patumwan, Bangkok 10330, Thailand

found when applying the national Kieferorthopädische Indikationsgruppe (KIG) indication system in Germany by various studies [4–6]. Incorrect use of these indication systems and a lack of basis for referral can increase pressure on the service providers and prolong waiting lists [7]. Guidelines do not seem to be able improve the quality of referrals significantly [3]. Hence, alternative strategies for improving detection of malocclusion among primary dental practitioners are required.

Malocclusion diagnosis requires the evaluation of images such as intra- and extraoral photographs, x-rays, and scans. Neural networks in particular have long been recognized for their potential in image analysis [8], but machine learning in orthodontics is still a developing field, leaving numerous unexplored opportunities [9].

The extent to which machin learning can assist clinicians in recognizing the indications for orthodontic treatment remains a less-explored domain within the broader landscape of AI applications in orthodontics. So far, AI has been successfully applied to assess crowding on occlusal intraoral pictures [10], to detect landmarks in cephalometric images [11], and to predict the necessity of orthognathic surgery [12] based on cephalograms [13]. Existing AI studies in orthodontics are predominantly centered on x-ray imaging. Machine learning applications on other sources of imaging data such as intra- and extraoral photographs and scans are still limited [10]. In general, demonstrating the true value of deep learning in clinical applications requires comprehensive studies that assess the robustness and generalizability of deep learning models on diverse datasets [11].

The aim of this study was to examine how effectively convolutional neural networks can detect crossbites as a malocclusion category using clinical intraoral photographs. Multiple convolutional neural networks were trained and compared regarding their accuracy in identifying crossbites as well as classifying the specific type of crossbite (frontal vs. lateral). The present study contributes to assessing the potential of machine learning approaches in orthodontics. The comparative analysis helps to identify suitable models for accurately detecting lateral and frontal crossbites as malocclusion categories termed KIG M4 and K4 in the German classification system, respectively. The present study provides a first step into developing AI-based systems that can assist examiners who initiate orthodontic treatment in determining if and when to refer a patient for orthodontic treatment.

Materials and methods

Dataset

The dataset used in this study was obtained from the Section of Orthodontics, Aarhus University, Denmark. It includes patients who underwent an initial orthodontic consultation at the Section of Orthodontics, Aarhus University, Denmark between 01.07.2018 and 31.07.2023. The dataset contains randomly selected clinical photographs taken for orthodontic diagnoses and treatment planning, hence representing the whole patient cohort seen during this time interval. Photos displaying the occlusion from anterior, left, and right sides were included for each patient. Exclusion criteria were crossbites in deciduous dentitions and orthodontic treatment in progress. For patient anonymity, the intraoral image dataset was used without personal information such as name, age, or gender. The images were first labelled as non-crossbite and crossbite. In a second step, the crossbite photographs were further labelled as lateral or frontal crossbite. The analysis applied the German classification system "Kieferorthopädische Indikationsgruppen", which determines health insurance cover of treatment and distinguishes between frontal crossbites (M4) (Fig. 1) and unilateral crossbites (K4) (Fig. 2) among others. If a patient exhibited both a lateral and frontal crossbite, the images were labelled as frontal crossbite in line with the KIG classification as M4 instead of K4 (Fig. 3). The labelling was initially done by S.V. and independently repeated by M.S. without any conflicts.

Fig. 1 Frontal crossbite

Fig. 2 Lateral crossbite

Fig. 3 Combination of a frontal crossbite and lateral crossbite. *Note:* According to the malocclusion category system (Kieferorthopädische Indikationsgruppe, KIG) this image would be classified as a M4 (frontal crossbite) and not K4 (lateral crossbite)

Preprocessing of the dataset

All preprocessing was performed using the PyTorch 2.0.1 framework (The Linux Foundation, San Francisco, CA, USA) for Python 3.10.12. For the training and testing of the models, 10% of the data was randomly split and only used for testing purposes, whereas the 90% of the images were used for training and validation. All images were resized to 224×224 or to 299×299 pixels to satisfy the respective model's input requirements.

To enhance the performance of the deep learning models with a limited number of original samples and to avoid overfitting, data augmentation was applied dynamically during the training process. Hence, each time an image was loaded during training, it was randomly modified using specified transformations. These transformations included random horizontal flips, rotations of up to 20°, and brightness adjustments by up to 20%. Such dynamic augmentation ensured that the model encountered varied versions of the images throughout the training. Hence, it ensures that the model learns to generalize from the underlying patterns in data, thereby improving its generalization capabilities without increasing the actual number of images in the dataset. This process was repeated across all folds during the k-fold cross-validation.

Classification models

Neural networks are a set of algorithms designed to recognize patterns. They function by processing input data such as images through layers of artificial neurons, which are inspired by human brain neurons. Convolutional neural networks (CNNs), a type of neural network, are increasingly used in medical image diagnostics for tasks such as detection, segmentation, and classification of anatomical structures. To classify the occlusions, we used several different CNN models which have previously been successfully applied in other image classification studies [13].

ResNet18 and ResNet50

The ResNet architecture, introduced by He et al. [14], uses residual blocks, which are blocks stacked on top of each other. The ResNet architecture incorporates skip connections, which enable the network to bypass certain layers. It also integrats batch normalization between layers, which makes the training process more stable and faster. ResNet has several variants differing in the number of neural network layers, such as the ResNet18 and ResNet50 with 18 and 50 layers, respectively, which are applied in this study.

MobileNet

Howard et al. developed MobileNet as an architecture for applications where computational resources and processing time are limited. Its key innovation is the use of depth-wise separable convolutions instead of standard convolutions used in many other neural networks. This approach splits the standard convolution into a summation of two distinct steps: a depth-wise convolution, which filters the input, and a pointwise convolution, which combines the filtered results using a 1×1-dimensional filter [15].

Xception

Xception model, a deep learning architecture proposed by Chollet, is inspired by the Inception architecture. Xception replaces the layers with depth-wise separable convolution to filter and combine information in a more efficient way. It combines these depth-wise separable convolutions with residual connections, which help the network to learn better by allowing information to skip certain layers [16].

DenseNet

Developed by Huang et al., DenseNet is a deep convolutional neural architecture with dense connections among its units, where each layers connects directly with each subsequent layer in a feed-forward manner [17]. This means that instead of just passing information from one layer to the next, each layer receives inputs from all previous layers and passes its own output to all subsequent layers.

EfficientNet

Introduced by Tan et al. [18], EfficientNet implements a scaling method that uniformly adjusts all three dimensions of the neural network: depth, width, and resolution. In this context, 'depth' refers to the number of layers, 'width' represents the number of channels in each layer, and 'resolution' indicates the input image size. This scaling methodology replaces arbitrary adjustments with a systematic approach, ensuring consistent scaling across all dimensions. This approach is based on a smaller based model, which is expanded using scaling coefficients predetermined through a grid search. In our implementation, we adopted the EfficientNet-B0 variant [18].

Model training

Model training was performed using Pytorch 2.0.1. In the initial step the models were trained to classify non-crossbite vs. crossbite. In the second step the models where trained to classify non-crossbite vs. lateral crossbite vs. frontal crossbite. Given the limited sample size of the dataset, we used k-fold cross-validation with k=10. To adjust the convolutional neural network models for our classification task, we used transfer learning. Consequently, we used the network layers of the respective model pre-trained on the ImageNet dataset [19]. We replaced the last layer (classifier) with a new output layer to conform to the number of classes in our dataset (e.g. two and three). For the output layer, SoftMax activation was used for all models. The initial learning rate was set to 0.001. The cross-entropy loss function (binary for two-classes and categorical for three-class classification) was used. AdamW optimizer was applied to reduce the risk of overfitting [20]. Batch size was 16. The number of epochs was set to 20 with an early stopping criterion if validation loss did not improve for three consecutive epochs.

Model evaluation

The models were tested on the remaining (10%) test data that were not included in the training. Accuracy, precision, recall (sensitivity), specificity, F1 score, and Cohen's Kappa were calculated, and confusion matrices were determined for each model and tasks. Additionally, we mapped the Receiver Operating Characteristic (ROC) and calculated the corresponding Area Under the Curve (AUC) value.

Results

Dataset

The inclusion and exclusion criteria resulted in a dataset of 676 photographs from 311 patients. The labelling resulted in 260 non-crossbite images, 258 frontal crossbite images and 158 lateral crossbite images.

Crossbite vs. non-crossbite

All models trained to classify crossbite vs. non-crossbite showed high accuracy. The best performance over all k-folds were achieved by Xception with 98.57% accuracy in the validation dataset (Table 1). This was closely followed by MobileNet (98.55%), ResNet18 (97.14%), DenseNet (97.10%), and EfficientNet (97.10%). Slightly performing worse than the other architectures, ResNet50 model exhibited the lowest accuracy with a maximum of 91.43% over all k-folds. Specificity ranged from 100.00% for Xception and ResNet18 to 90.91% for ResNet50. For precision, again Xception again outperformed the other architectures with a value of 98.94%. All models demonstrated robust recall values over 90%. In terms of the F1-score, MobileNet achieved the highest score of 98.49%. The high Cohen's Kappa values across models indicate a strong agreement between the predicted and actual classifications; only ResNet50 showed a lower performance with a value of 81.93%.

Results on the test set and mean metrics over all k-folds are presented in Tables A1 and A2 in the appendix. The confusion matrices (Fig. 4) displayed a high number of true positives and true negatives, with a low count of false positives and false negatives. The Receiver Operating Characteristic (ROC) curves as well as the observed Area Under the Curve (AUC) indicated strong performance of all models (Fig. 5).

Lateral crossbite vs. frontal crossbite vs. non-crossbite

The models which were trained to not only detect a crossbite, but also differentiate between lateral and frontal crossbites performed slightly worse on average than the models that only classified crossbite vs. non-crossbite. The highest accuracy over all k-folds was achieved by the DenseNet model with 91.43%, closely followed by MobileNet (91.30%), EfficientNet (90.00%), and Xception (88.57%) in the validation dataset (Table 2). ResNet18 as well as ResNet50 lagged behind with 76.81% and 74.29%, respectively. Notably, the accuracy metrics for the frontal crossbite group were lower than in the other groups

Table 1 Highest model accuracy metrics on the test set for two-class classification over all k-folds. *Note:* Highest values for each metric highlighted in bold

Model	Accuracy (in %)	Specificity (in %)	Precision (in %)	Recall (sensitivity) (in %)	F1-score (in %)	Cohen's Kappa (in %)
ResNet18	97.14	**100**	98.00	95.45	96.60	93.20
ResNet50	91.43	90.91	90.48	91.61	90.96	81.93
MobileNet	98.55	97.62	98.21	**98.81**	**98.49**	**96.98**
Xception	**98.57**	**100**	**98.94**	97.92	98.40	96.80
DenseNet	97.10	95.24	96.55	97.52	96.99	93.99
EfficientNet	97.10	95.65	96.00	97.83	96.81	93.62

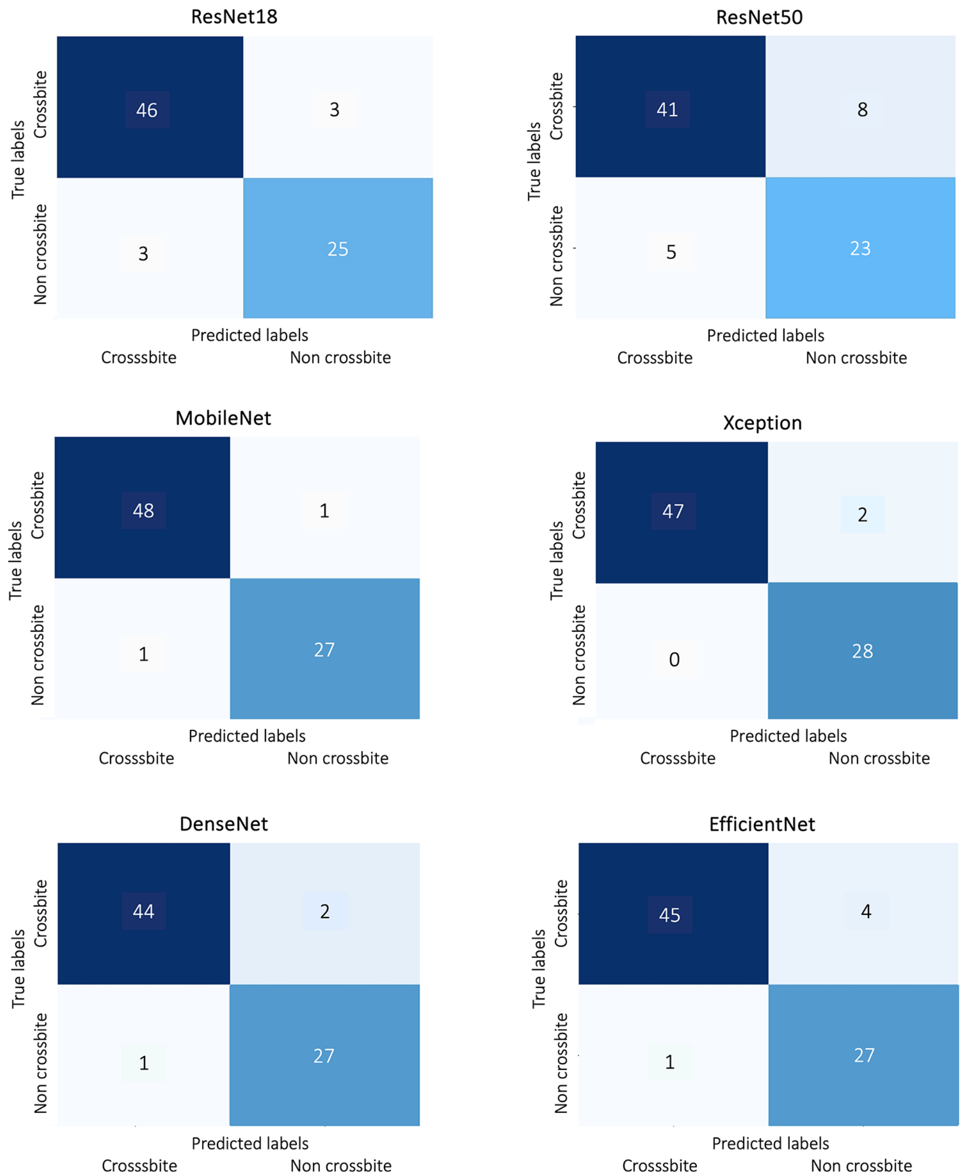

Fig. 4 Crossbite vs. non-crossbite: Confusion matrices. *Note:* Illustration of respective model with highest accuracy on test data among all k-folds

across the models. Confusion matrices for each model and ROC Curve including the computed AUC value are visualized in Figs. 6 and 7. Tables A3 and A4 in the appendix present further results.

Discussion

This study implemented various deep learning architectures to compare their performance to classify crossbites. For distinguishing between cross-bite vs. non-crossbite, all models achieved high accuracy suggesting that the different models were effective in learning the distinguishing features between the two classes. The excellent accuracy achieved by Xception and MobileNet particularly underscores the potential of convolutional neural networks in capturing occlusal and orthodontic features in 2D intraoral images. Their strong performance indicates that depth-wise separable convolutions and skip connections as architectural choices can effectively extract the relevant features from the images. EfficientNet, DenseNet, and ResNet18 demonstrate similarly high accuracy, which suggests that multiple approaches can effectively capture the essential features for this binary classification task. The high accuracies for these models are in line with applications of neural networks to diagnose the indications of orthognathic surgery [21–23]. Only ResNet50 slightly lagged behind in terms of performance in this binary classification task, possibly due to the ResNet50's overly complex structure leading to overfitting.

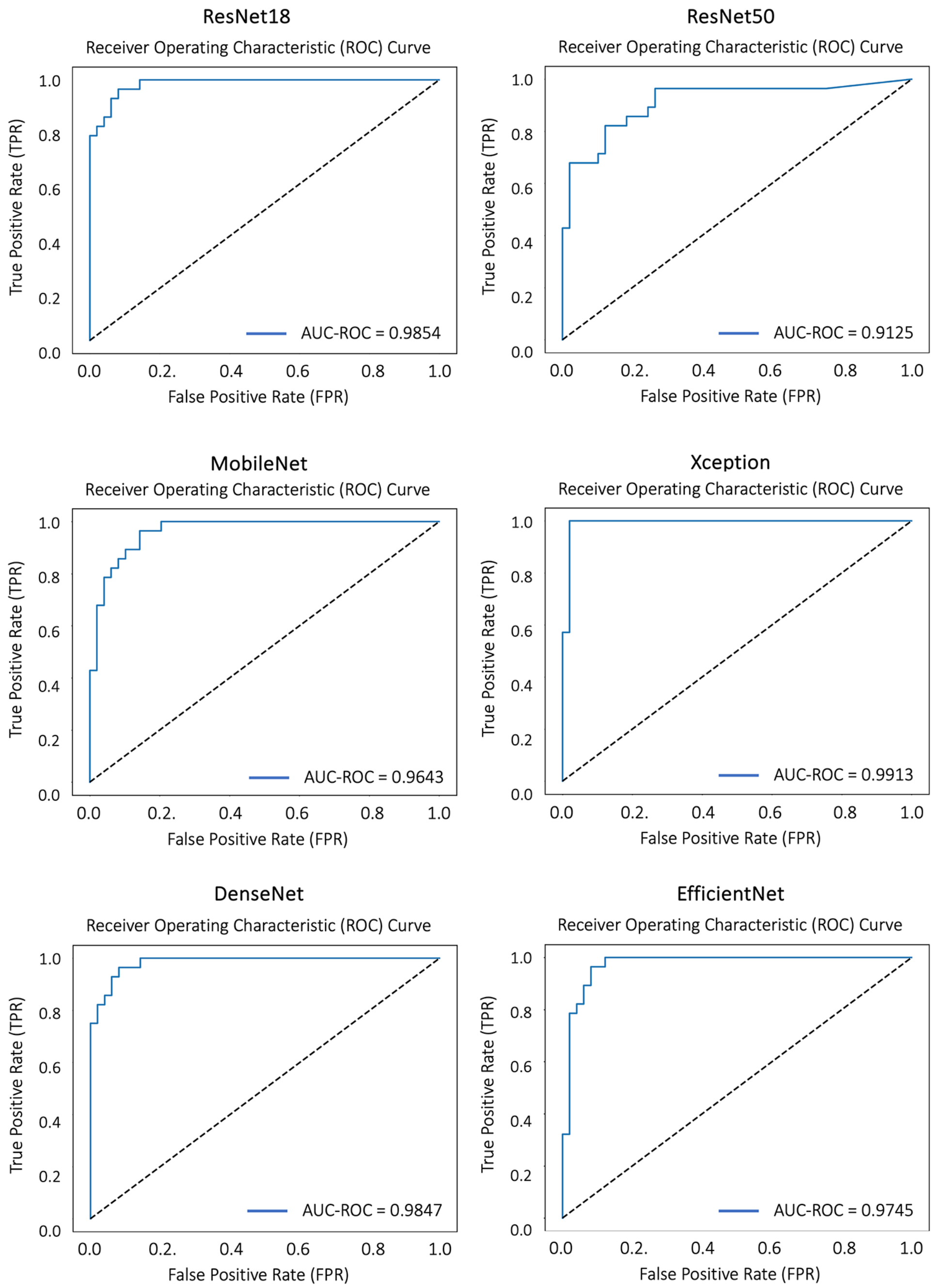

Fig. 5 Crossbite vs. non-crossbite: Receiver Operating Characteristic (ROC) curves. *Note:* Illustration of respective model with highest accuracy on test data among all k-folds

The reduced accuracy in the three-class classification task highlights the increased difficulty for models when distinguishing between frontal and lateral crossbites, compared to the simpler binary classification of crossbite versus non-crossbite. Besides, the reduced accuracy is likely related to the reduced number of images available for training in the crossbite groups as the crossbite group is split into lateral and frontal crossbites. This is supported by the comparatively lower accuracy metrics for these groups. To successfully train neural networks, the training dataset needs to be sufficiently large [24]. The models faced an additional challenge related to the lateral crossbite groups because this class can exhibit features overlapping with other classes: If the patient exhibited both a frontal and lateral crossbite, the malocclusion was classified as a frontal crossbite (KIG M4) in line with the KIG classification system [25]. Other authors have also reported a drop in performance when adding additional

Table 2 Highest model accuracy metrics for three-class classification over all k-folds. *Note:* Highest values for each metric highlighted in bold

Model	Accuracy (in %)	Specificity (class 2, in %)	Precision (average, in %)	Recall (Sensitivity) (average, in %)	F1-Score (in %)	Cohen's Kappa (in %)
ResNet18	76.81	**100**	86.07	59.12	55.38	55.50
ResNet50	74.29	**100**	83.24	61.62	54.51	56.89
MobileNet	91.30	91.67	**91.72**	83.60	84.00	**91.67**
Xception	88.57	93.94	85.80	82.60	83.20	81.52
DenseNet	**91.43**	93.94	88.46	88.90	88.60	86.38
EfficientNet	90.00	96.15	89.84	**89.00**	**89.18**	84.84

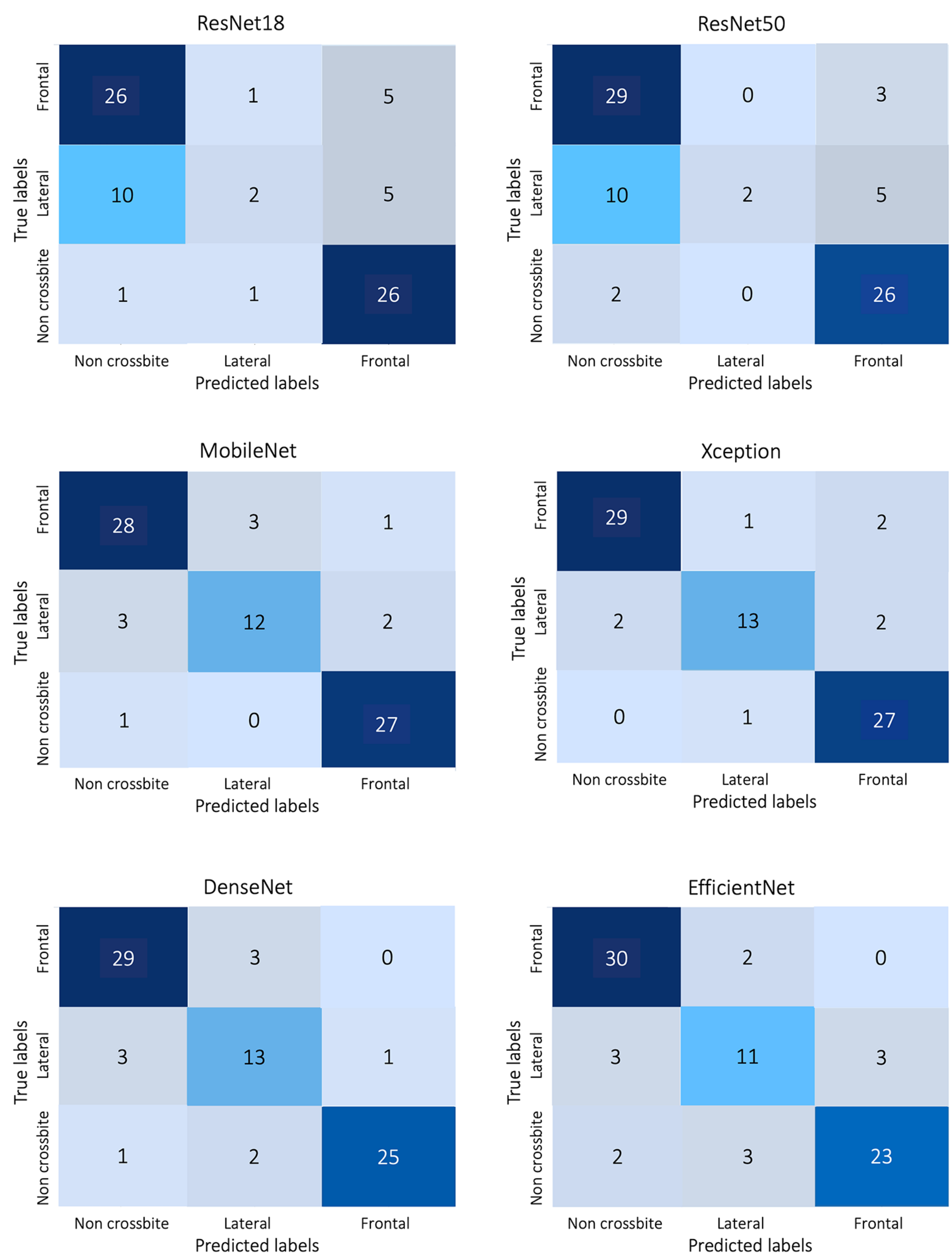

Fig. 6 Crossbite vs. frontal crossbite vs. lateral crossbite: Confusion matrices. *Note:* Illustration of respective model with highest accuracy on test data among all k-folds

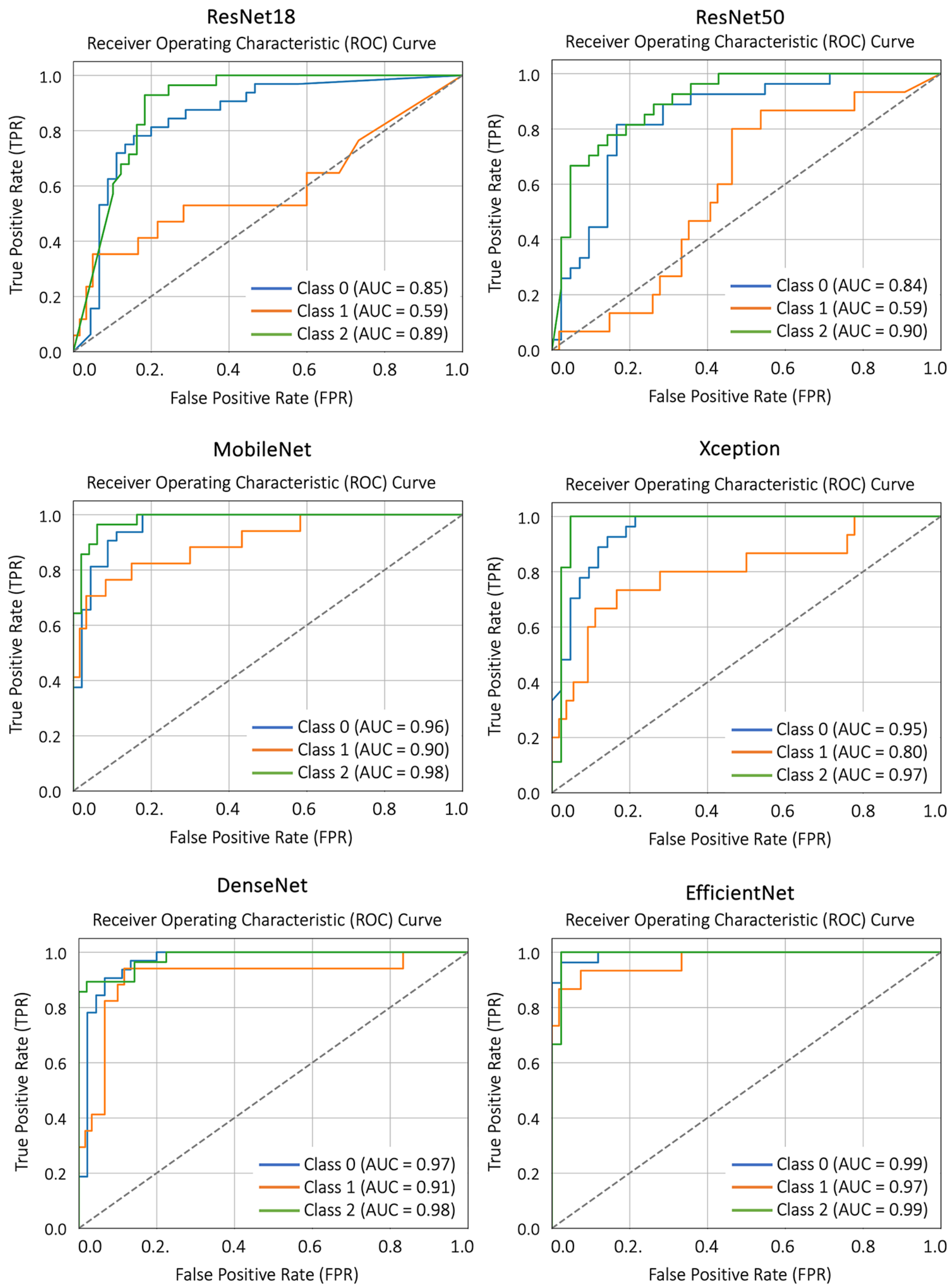

Fig. 7 Crossbite vs. frontal crossbite vs. lateral crossbite: Receiver Operating Characteristic (ROC) curves. *Note:* Illustration of respective model with highest accuracy on test data among all k-folds. Class 0 = frontal crossbite, class 1 = lateral crossbite, class 2 = non-crossbite

classes to a classification task, for example in the context of predicting the binary decision of performing orthognathic surgery vs. the more detailed diagnosis of surgery type and extraction decision [22]. Although accuracy decreases compared to the binary classification crossbite vs. non-crossbite, the accuracy is still over 90% for some of the neural networks, demonstrating their potential to further distinguish between frontal and lateral crossbites.

In the present dataset, DenseNet's architecture seems to be best suited to learn and predict the features of non-crossbite, lateral and frontal crossbite, despite the small sample size. The comparatively poor performance by both the ResNet50 and ResNet18 is in line with Ryu et al. [26] who found that the ResNet architecture does not perform as efficiently as other neural networks to detect crowding in orthodontic images.

The performance dropped when differentiating between frontal and lateral crossbites as opposed to the binary classification crossbite vs. non-crossbite prompt a critical reflection on the trade-offs in clinical applications. The lower accuracy when differentiating lateral from frontal crossbites might be acceptable if the clinical implications of differentiating crossbites outweigh the simplicity and higher accuracy of a binary classification.

Overall, the results highlight that CNN have high potential to reliably support the detection of crossbites. Possible applications include the remote diagnosis of dentofacial deformities and virtual treatment monitoring, for example Dental Monitoring. Preliminary assessments through CNNs on uploaded patient images can benefit patients in remote areas with limited access to orthodontic specialists. Virtual treatment monitoring allows for continuous patient assessment without the need for frequent in-person visits. Such applications of CNN can also support dentists or pediatricians as frequent initiators for orthodontic treatment to identify orthodontic issues early, allowing for timely intervention and potentially reducing the severity and duration of treatment.

The study has several limitations. Firstly, our evaluation was restricted to lateral and frontal crossbites and did not include other types of malocclusions; a model's performance on a single task does not necessarily predict its performance on another. The results might be limited due to utilization of 2D intraoral photographs. For example, detecting a crossbite on the second molars based on 2D images can be challenging because the second molars which are not fully erupted at treatment start might not be fully captured. Additionally, the clinical photographs were taken using intraoral mirrors, which can bias the images and lead to under- or overestimation of the presence and severity of posterior and anterior crossbites [27]. Hence, a direct clinical examination can provide a more insightful analysis, and the value of image analysis is more pronounced in the absence of such a clinical examination. Although the results show high accuracy, the relatively small dataset size might limit the generalizability of the models. Furthermore, a visual explanation of the models' output (e.g. Gradient-weighted Class Activation Mapping) was not included in this study.

To overcome these limitations, future research should also assess other malocclusion classifications commonly captured in national indication systems, based on 3D scans. Further deep learning models should be tested and compared in larger datasets to demonstrate the full potential, generalizability, and explainability of these approaches. In particular, future research should focus on improving accuracy for distinguishing between frontal and lateral crossbites.

Conclusions

This study introduces several deep learning models designed to detect specific malocclusion traits and differentiate between frontal and lateral crossbites. The models classifying non-crossbite vs. crossbite show very high accuracy, which highlights their potential in detecting this malocclusion. The models that additionally distinguish between lateral and frontal crossbites show slightly lower accuracy, indicating that they are limited by the smaller sample size and additional challenge of distinguishing within the crossbite group. Overall, the results suggest that convolutional neural network models are capable of learning and processing intraoral 2D photographs for orthodontic diagnosis. This can provide a first step to employ AI-based systems to support examiners with little experience in making referrals effectively and optimize utilization of services.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13005-024-00448-8.

Supplementary Material 1

Acknowledgements

We would like to thank all orthodontic residents for their efforts in capturing high-quality images, which served as the foundation of this study.
We would like to thank all postgraduate residents at our department for taking intraoral pictures that have been necessary for the analysis.

Author contributions

BN, SV and RP trained the deep learning models. MS and SV labelled the intraoral pphotographs. BN, SV, RP and PS were major contributors in writing and editing the manuscript. All authors read and approved the manuscript.

Funding

No funding was received.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethical approval

Ethical approval was granted from the Danish National Committee on Health Research Ethics (NVK) 2400379/2923796.

Consent for publication

Consent for publishing individuals' personal data has been obtained.

Data sharing

The data used in this study will not be shared as it contains sensitive patient information.

Competing interests

The authors declare no competing interests.

Received: 30 May 2024 / Accepted: 26 August 2024
Published online: 02 September 2024

References

1. Kunz F, Stellzig-Eisenhauer A, Boldt J. Applications of Artificial Intelligence in Orthodontics—An overview and perspective based on the current state of the art. Appl Sci. 2023;13(6).

2. Di Blasio M, Vaienti B, Pedrazzi G, Cassi D, Magnifico M, Meneghello S, Di Blasio A. Are the reasons why patients are referred for an Orthodontic visit correct? Int J Environ Res Public Health. 2021;18(10).
3. O'Brien K, Wright J, Conboy F, Bagley L, Lewis D, Read M, et al. The effect of orthodontic referral guidelines: a randomised controlled trial. Br Dent J. 2000;188(7):392–7.
4. Furhmann RAW. Genehmigungsfähigkeit bei unklarem KIG-Befund Obergutachten zur Überprüfung von GKV-Gutachten. 2019.
5. Gesch D, Kirbschus A, Schröder W, Bernhardt O, Proff P, Bayerlein T, et al. Influence of examiner differences on KIG-classification when assessing malocclusions. J Orofac Orthop = Fortschr Der Kieferorthopadie: Organ/official J Dtsch Gesellschaft fur Kieferorthop. 2006;67(2):81–91.
6. Stolze A, Goldbecher H. Der optimale Behandlungsbeginn | Quintessenz Verlags-GmbH. 2012. pp. 271 – 84.
7. Reddy S, Derringer KA, Rennie L. Orthodontic referrals: why do GDPs get it wrong? Br Dent J. 2016;221(9):583–7.
8. Weide E. Die Macht der künstlichen Intelligenz. Ein gelungener Einstieg ins nächste Jahrtausend. Ullstein Taschenbuchvlg; 1993.
9. Auconi P, Gili T, Capuani S, Saccucci M, Caldarelli G, Polimeni A, Di Carlo G. The validity of machine learning procedures in Orthodontics: what is still missing? J Pers Med. 2022;12(6).
10. 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;13(1):5177.
11. 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):4299–309.
12. Taraji S, Atici SF, Viana G, Kusnoto B, Allareddy VS, Miloro M, Elnagar MH. Novel machine learning algorithms for prediction of treatment decisions in adult patients with Class III Malocclusion. J Oral Maxillofac Surg. 2023.
13. Lee K-S, Ryu J-J, Jang HS, Lee D-Y, Jung S-K. Deep Convolutional Neural Networks Based Analysis of Cephalometric Radiographs for Differential diagnosis of orthognathic surgery indications. Appl Sci. 2020;10(6).
14. He K, Zhang X, Ren S, Sun J, editors. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Recognition P. (CVPR); 2016 27–30 June 2016.
15. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 2017.
16. Chollet F, Xception. Deep Learning with Depthwise Separable Convolutions2017. 1800-7 p.
17. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ, editors. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.
18. Tan M, Le Q, editors. Efficientnet: Rethinking model scaling for convolutional neural networks. International conference on machine learning; 2019: PMLR.
19. Deng J, Dong W, Socher R, Li LJ, Kai L, Li F-F, editors. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition; 2009 20–25 June 2009.
20. Lee K-S, Ryu J-J, Jang HS, Lee D-Y, Jung S-K. Deep Convolutional Neural Networks Based Analysis of Cephalometric Radiographs for Differential diagnosis of orthognathic surgery indications. Appl Sci [Internet]. 2020; 10(6).
21. Lee C, Jeon KJ, Han SS, Kim YH, Choi YJ, Lee A, Choi JH. Ct-like MRi using the zero-te technique for osseous changes of the tMJ. Dentomaxillofacial Radiol. 2020;49(3).
22. Choi H-I, Jung S-K, Baek S-H, Lim WH, Ahn S-J, Yang I-H, Kim T-W. Artificial Intelligent Model with neural network machine learning for the diagnosis of orthognathic surgery. J Craniofac Surg. 2019;30(7).
23. Shin W, Yeom H-G, Lee GH, Yun JP, Jeong SH, Lee JH, et al. Deep learning based prediction of necessity for orthognathic surgery of skeletal malocclusion using cephalogram in Korean individuals. BMC Oral Health. 2021;21(1):130.
24. Alwosheel A, van Cranenburgh S, Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J Choice Modelling. 2018;28:167–82.
25. Schopf P. Kieferorthopädische Abrechnung (BEMA und GOZ GOÄ) mit Erläuterung der ab 1.1.2002 gültigen Kieferorthopädischen Indikationsgruppen (KIG) sowie der ab 1.1.2004 geltenden Fassung des BEMA und der Richtlinien des Bundesausschusses der Zahnärzte und Krankenkassen.
26. 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;13(1).
27. Jackson T, Kirk C, Phillips C, Koroluk L. Diagnostic accuracy of intraoral photographic orthodontic records. J Esthetic Restor Dentistry. 2018;31.