European Journal of Orthodontics, 2022, **44**, 445–451
https://doi.org/10.1093/ejo/cjac016
Advance Access publication 9 May 2022
Original Article

Using artificial intelligence to determine the influence of dental aesthetics on facial attractiveness in comparison to other facial modifications

Dorothea Obwegeser[1], Radu Timofte[2,3], Christoph Mayer[2], Theodore Eliades[1], Michael M. Bornstein[4], Marc A. Schätzle[1] and Raphael Patcas[1]

[1]Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Switzerland
[2]Computer Vision Laboratory, Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland
[3]CAIDAS and Institute of Computer Science, Faculty of Mathematics and Computer Science, University of Wurzburg, Germany
[4]Department of Oral Health & Medicine, University Center for Dental Medicine Basel UZB, University of Basel, Switzerland

Correspondence to: Raphael Patcas, Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Plattenstrasse 11, 8032 Zurich, Switzerland. E-mail: raphael.patcas@zzm.uzh.ch

Summary

Background: Facial aesthetics is a major motivating factor for undergoing orthodontic treatment.

Objectives: To ascertain—by means of artificial intelligence (AI)—the influence of dental alignment on facial attractiveness and perceived age, compared to other modifications such as wearing glasses, earrings, or lipstick.

Material and methods: Forty volunteering females (mean age: 24.5) with near perfectly aligned upper front teeth [Aesthetic Component scale of the Index of Orthodontic Treatment Need (AC-IOTN) = 1 and Peer Assessment Rating Index (PAR Index) = 0 or 1] were photographed with a standardized pose while smiling, in the following settings (number of photographs = 960): without modifications, wearing eyeglasses, earrings, or lipstick. These pictures were taken with natural aligned dentition and with an individually manufactured crooked teeth mock-up (AC-IOTN = 8) to create the illusion of misaligned teeth. Images were assessed for attractiveness and perceived age, using AI, consisting of a face detector and deep convolutional neural networks trained on dedicated datasets for attractiveness and age prediction. Each image received an attractiveness score from 0 to 100 and one value for an age prediction. The scores were descriptively reviewed for each setting, and the facial modifications were tested statistically whether they affected the attractiveness score. The relationship between predicted age and attractiveness scores was examined with linear regression models.

Results: All modifications showed a significant effect (for all: $P < 0.001$) on facial attractiveness. In faces with misaligned teeth, wearing eyeglasses (−17.8%) and earrings (−3.2%) had an adverse effect on facial aesthetics. Tooth alignment (+6.9%) and wearing lipstick (+7.9%) increased attractiveness. There was no relevant effect of any assessed modifications or tooth alignment on perceived age (all: <1.5 years). Mean attractiveness score declined with predicted age, except when wearing glasses, in which case attractiveness was rated higher with increasing predicted age.

Conclusions: Alignment of teeth improves facial attractiveness to a similar extent than wearing lipstick, but has no discernable effect on perceived age. Wearing glasses reduces attractiveness considerably, but this effect vanishes with age.

Introduction

Each face is unique, and its features and characteristics distinguish a person. Its attractiveness, i.e. the degree to which a person's physical features are considered aesthetically pleasing or beautiful, strongly influences human interaction (1). An appealing face is commonly associated with positive characteristics, such as better social skills (2) or enhanced biological qualities like fertility and overall health (3). Thus, the pursuit for an attractive appearance is a strong desire for many people.

There are several possibilities for women and men to enhance their facial appearance. Options used in the strive for attractiveness range from simple interventions like cosmetic makeup, sophisticated haircuts, or jewellery, and accessories (4) to more elaborate and invasive procedures such as orthodontic treatment (5) or corrective surgical interventions (6).

The association between dental aesthetics and facial attractiveness has long been established (7–9), and so has the effect of dental attractiveness on perceived social and intellectual qualities (7, 10, 11). It is thus no surprise that improving aesthetics is manifestly the primary motivating factor for most patients to undergo orthodontic treatment (5, 11, 12).

Artificial intelligence (AI) has in recent years been introduced to diagnose dentofacial aesthetics, and convolutional neural networks (CNNs) can now accurately evaluate facial attractiveness or predict an apparent age (13–15). Attractiveness is the quality of being pleasing or appealing. Unlike beautifulness, attractiveness does not describe the face *per se*, but rather how the face is being perceived. As such, facial attractiveness can be evaluated, rated, and ranked by observers or by CNNs. Such deep learning models are not limited to identify and interpret facial features relevant to

attractiveness, but also those that are related to a person's ageing. Most CNNs have a diagnostic performance comparable to medical experts (16) and allow to circumvent the common drawbacks and limitations of rater-based evaluations.

It seems, the impact of dental corrections in comparison to other facial modifications has so far never been investigated before by means of AI. Hence, the objective of this study was to explore—by means of an existing face detector and newly trained attractiveness/age prediction regression networks—the influence of anterior tooth alignment (i.e. a misaligned vs. an aligned upper front) on facial attractiveness and apparent age, in comparison to other common facial modifications. To determine the influence of dental alignment on facial attractiveness compared to other modifications, such as wearing glasses, earrings, or lipsticks, the null hypothesis was formulated that neither tooth alignment nor any facial modifications would have any significant noticeable influence on facial attractiveness.

Materials and methods

This investigation is based on 960 frontal facial images of 40 volunteering female participants (age range: 20–30 years; mean age: 24.5 years) comprising dentists, dental students, and dental nurses from the Center of Dental Medicine of the University of Zurich, Switzerland.

All participants were informed about the research purposes in form of a flyer and a short briefing. All participants signed an informed consent form, including their approval to have their faces possibly printed and published.

Inclusion criteria were perfectly aligned upper incisors and a signed informed consent form. Plaster models were made to ascertain the perfect alignment of the upper front, corresponding to a grade 1 of the Aesthetic Component (AC) scale of the Index of Orthodontic Treatment Need (IOTN) (17) and a displacement score of 0 or 1 of the Peer Assessment Rating Index (PAR Index) (18).

The gypsum models were further used to produce 40 individual, perfectly adapted mock-ups, which displayed all the same dental misalignment: artificial teeth (MAJOR DENT, mould 16, shade A3.5) were positioned to match grade 8 of the IOTN's AC scale (17), and a silicon index was used to transfer the exact same misalignment from one setup to the other. Pink denture plastics were used to mimic the gingiva, and the completed mock-ups were polished to high gloss prior insertion.

Photographs were taken under standardized conditions in front of a monochrome dark blue background, at a 2-m distance from the photographer. Throughout all of the photo shooting sessions, the same single-lens reflex camera with a dedicated flash and fixed settings (aperture F11, shutter 1/200, ISO 200) was used.

Four different settings were evaluated, each with aligned and misaligned teeth (i.e. with inserted mock-up), equaling in total to eight independent observations of a smiling face: 1. without any accessories and modifications, 2. wearing golden hoop earrings, 3. wearing horn-rimmed glasses, and 4. wearing a classical red lipstick (Figure 1). Three separate pictures were taken from each of the eight settings (in order to minimize the impact of possible distractors like a momentarily facial expressions) of all forty participants (total number of photographs = 960). Images were not further processed, only rotated and cropped to the same size of 256 × 256 pixels to obtain an aligned face and a 40% background margin. Photo shooting sessions took approximately 20 minutes per participant. Participants were compensated with a theatrical movie voucher, their individual mock-up, and the red lipstick.

Facial attractiveness of each picture was evaluated using AI (19). The applied AI consisted of an existing face detector (20) and newly developed, dedicated deep CNNs employing a DenseNet-201 architecture (21) for image and feature recognition. To align the facial images in a unified setting, the midpoint between both eyes was determined for all images individually by the face detector (20), and the faces were rotated around this point to align them horizontally and scale the faces according to the distance between both eyes such that the eyes of all faces were aligned.

To regress the attractiveness score and an estimated age for each image, the regression problem was transformed into a classification problem by binning both values into a fixed number of bin-classes (22). As mentioned, DenseNet-201 was used as a feature extractor followed by a linear layer with softmax activation to predict the bin class for each input image. To regress a continuous value instead of a discrete value, Rothe *et al.*'s (22) approach was adopted and the expected value was computed over the output distribution p for each value to predict the final output value:

$$E\ [p] = \sum_{i=0}^{p_i} i \cdot p_i$$

Training for attractiveness score

Two distinct networks (for age and attractiveness regression) were trained on a single Nvidia TitanXp GPU. The CNN feature extractor of both networks was initialized with pretrained weights on the ImageNet dataset (23) and trained for attractiveness on the dataset of the BLINQ dating application (19). This dataset contains a large diversity of casual facial images of mostly Caucasians (aged range: 18–37 years, mean age: 25 years), including the annotation of over 17 million corresponding attractiveness ratings performed by the application users. The dataset was filtered and only facial images showing females (>13 000 facial pictures) were used, as only these were relevant.

The obtained normalized attractiveness scores were split into 10 bins. Subsequently, the network was trained for 12 epochs with a batch size of 32 and a learning rate of 0.003 and the learning rate was decayed by a factor of 10 after 8 epochs. The same data augmentation strategy as Rothe *et al.* (22) was used, namely random scaling by 10%, random translation by 10%, and random rotation by 10 degrees and resizing each image to a resolution of 256 × 256 pixels before feeding it to the network. Training was performed in a fully supervised fashion, using cross-entropy as the loss function. The trained network achieved a Pearson correlation of 0.641 on the constructed test set.

To obtain a robust attractiveness score, the detected facial images were augmented 10 times using the same data augmentation employed during training, resulting in 100 predictions per image, which were then averaged. Every picture received one attractiveness score from 0 to 100 (0 = extremely unattractive, 100 = extremely attractive).

Figure 1. Sample pictures of a test person showing the eight different settings: four with aligned (left side) and four with misaligned mock-up teeth (right side. Clockwise: smile, smile and earrings, smile and lipstick, smile and glasses).

Training for age prediction

For age estimation (age prediction), the DenseNet-201 was pre-trained on the ImageNet dataset (23), and then trained on the IMDb-Wiki dataset (age range 0–100 years; containing >221 000 facial images annotated with their respective age) (22) for 12 epochs using cross-entropy with 101 bins, where each year corresponds to one bin. The CNN model was further fine-tuned on the APPA-REAL dataset (age range: 0–95) (24), which contains 7591 images including associated real and apparent age labels with nearly 300 000 votes, for an ensemble of 10 different models with different train-validation splits but with the same test set. For fine-tuning, the same training setting as above (for pre-training on the IMDb-Wiki dataset) was executed on the APPA-REAL dataset, but now for 24 epochs with a 10 times smaller learning rate of 0.0003 that was decayed after 18 epochs by a factor of 10. The ensemble averages the age prediction of all models at test time. The fine-tuned method trained for this investigation achieved a mean absolute error of 3.926 on the test set of APPA-Real, improving over previously reported accuracy of 4.082 (24).

Facial images were augmented 10 times and age was predicted with all trained networks, leading to 100 different estimates in total for inference, resulting in 100 predictions per image which were then respectively averaged.

Ethics

Each participant signed an informed consent form, explicitly authorizing the use and publication of facial images for research purposes. The protocol of the study was reviewed by the governmental ethics committee prior to the execution of the study and a formal waiver was issued (BASEC Req. 2019-0030). Guidelines in medical ethics, as specified in the WMA declaration of Helsinki (64th WMA General Assembly, Fortaleza, Brazil, October 2013) were rigorously followed, together with all relevant juridical regulations.

Statistical analysis

Data were analysed in SPSS software [IBM SPSS Statistics for Windows, version 24 (IBM Corp., Armonk, New York, USA)]. The attractiveness scores and age predictions were descriptively reviewed for all eight different settings separately, and a smiling face without modifications was considered as reference. All investigated variables (wearing glasses, earrings etc.) were presumed to be independent variables. Exploratory analyses of the scores were performed, including box-and-whisker plots for each setting. Kolmogorov–Smirnov tests were computed to attest normal distribution, and—wherever appropriate—a Student's *t*-test or Wilcoxon signed-rank test was applied to disclose any effect of facial modifications. Attractiveness scores were plotted against predicted age, and

linear regression models were computed for each setting individually. *P*-values of <0.05 were considered significant.

Results

The attractiveness scores for each modification are presented in Table 1. The setting with the highest score is the smiling face with aligned teeth and lipstick, with an average attractiveness score of 51.5 (range: 19.9–91.5). The setting with the lowest score is a smiling face with misaligned teeth and eyeglasses, resulting in an average attractiveness score of 38.2 (rage: 20.2–69.3). A comparison of the attractiveness scores reveals statistical differences between the modifications (Table 1). While it is evident that aligned teeth produce a more attractive face (mean attractiveness score: 49.7) than misaligned teeth (46.5), this tendency can be observed throughout all facial modifications (Figure 2a and 2b): whatever facial modification used, the attractiveness score is always slightly lower with misaligned teeth.

Each modification has a very distinct effect on facial attractiveness (Table 1): lipstick increases facial attractiveness (considerably so in cases of misaligned teeth), earrings moderate facial attractiveness only slightly, and glasses provoke a substantial reduction of facial attractiveness.

The positive effect of wearing lipstick in cases with misaligned teeth [50.2% (+7.9%)] is greater that the alignment of the teeth [49.7 (+6.9%)], but the difference is not statistically significant ($P = 0.382$, evaluated by using Student's *t*-test for paired samples).

The impact of the different modifications on the estimated age is minimal (Table 2). All differences are below ±1.5 years of the mean perceived age, and as such not clinically relevant. Certain tendencies are, however, discernible: while earrings do seemingly not have any influence on estimated age, wearing lipstick makes one look slightly older and wearing glasses younger.

Figure 3a and 3b disclose the relationship between estimated age and attractiveness. An observable overall trend of

Table 1. Mean, median, and minimum and maximum attractiveness scores for all settings

	Attractiveness (%)			Normal distribution	Difference to baseline without any modifications	
	Mean	Median	Min; max	(*P*-value of K–S test)	Absolute (%)	*P*-value
Misaligned teeth	46.5	47.8	16.6; 84.5	No (0.031)		
Wearing earrings	45.0	45.8	17.3; 83.3	No (0.045)	−1.5 (−3.2%)	<0.001*
Wearing glasses	38.2	36.6	20.2; 69.3	No (0.033)	−8.3 (−17.8%)	<0.001*
Wearing lipstick	50.2	51.0	22.7; 86.5	Yes (0.060)	+3.7 (+7.9%)	<0.001**
Aligned teeth	49.7	48.5	16.5; 93.2	Yes (0.153)		
Wearing earrings	48.0	48.2	17.3; 91.5	Yes (0.200)	−1.7 (−3.4%)	<0.001**
Wearing glasses	41.7	39.8	21.6; 77.2	Yes (0.082)	−8.0 (−16.1%)	<0.001**
Wearing lipstick	51.5	51.4	19.9; 91.5	Yes (0.200)	+1.8 (+3.6%)	<0.001**

Kolmogorov–Smirnov test (K–S test) to unveil normal distribution. Wilcoxon signed-rank test and Student's *t*-test for paired samples to disclose statistically significant differences.
*Wilcoxon signed-rank test.
**Student's *t*-test for paired samples.

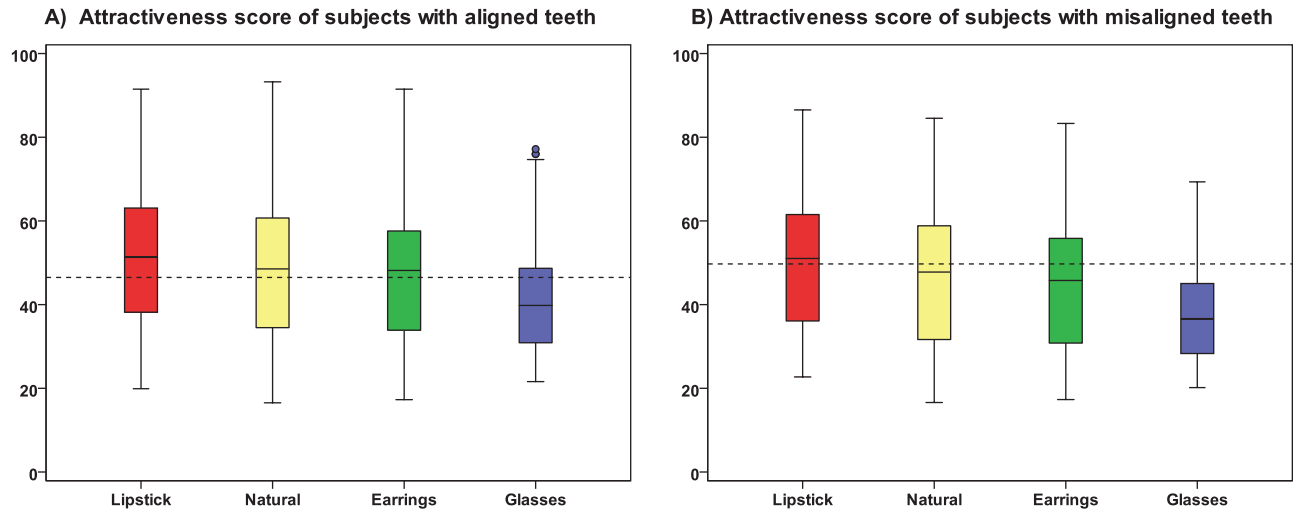

Figure 2. Box-and-whisker plots for facial attractiveness of participants wearing different modifications with aligned teeth (A) and misaligned teeth (B), respectively. Dotted line: mean attractiveness score of all subjects, with misaligned teeth (A) and aligned teeth (B), respectively.

reduced attractiveness in older-looking participants can be detected both for aligned (correlation coefficient: –0.36% per year) and misaligned cases (correlation coefficient: –0.37% per year). The particularly attractive faces (score >90) can only be found in cases with aligned teeth.

The association between apparent age and attractiveness is dependent on the four different settings. Interpreting the linear regressions lines individually, the rule 'the higher the estimated age, the lower the attractiveness score', holds true for faces without modification, or with lipstick or earrings. Only wearing glasses demonstrates a positive correlation between apparent age and attractiveness.

Discussion

This study is the first to use AI to assess the impact of anterior tooth alignment on attractiveness and age, in comparison to the effects of different facial adjuvants known to alter facial appearance. The results are revealing.

Examining each modification separately, this investigation was able to establish that every modification has a very distinct effect on facial attractiveness. All investigated modifications provoked modest yet always statistically significant alterations in facial attractiveness. Thus, the null hypothesis must admittedly be rejected. While mostly moderate in amplitude, the diversity of the different effects demonstrates how serious the combination of the various modifications controls and affects attractiveness. With a mean score of 38.2 for faces with misaligned teeth and glasses, the score increased considerably to 51.4 when the teeth were aligned and no glasses but lipstick was worn.

The overall range of the attractiveness scores of the forty assessed faces was rather large, giving testimony that the sample was particularly heterogeneous, containing faces with all degrees of attractiveness. Despite this diversity, the tests for paired differences all uncovered highly significant results concerning the impact of the modifications, and underline the generalizability of the observed effects. Thus, the study adds empirical evidence to the observation that the alignment of front teeth (9, 10, 25) and use of lipstick (26) both increase facial attractiveness, while wearing glasses cause a decrease (27).

One of the important novelties presented here is the possibility to directly juxtapose the different effects. The observation that the benefits of a simple and inexpensive modification such as lipstick are similar to those of a costly orthodontic treatment, is indeed disturbing for orthodontists and raises uncomfortable questions. One must bear in mind, however, that orthodontic alignment improves more than just oral attractiveness, and may provide other substantial advantages, such as enhanced masticatory function (28), self-esteem (29), and quality of life (30). Furthermore, bright red lipstick is usually not worn at all times and may even be inappropriate or undesirable under certain circumstances. While these aspects remain unaccounted for in this investigation, the clinician must be cognizant that when it comes to assess purely facial attractiveness, the outlined results suggest that the benefits of tooth alignment over lipstick are seemingly nonexistent.

In fact, some previous authors already claimed that while the oral region significantly contributes to overall facial attractiveness, other parts of the face—including the eyes or skin complexion—appear to be more influential cues than dental aesthetics (31), and that orthodontic correction may at times improve dental aesthetics without having any impact on facial aesthetics (25). The present contribution seems to echo these assumptions.

The results are also in line with former publications on the use of lipstick or eyeglasses and their effect on facial attractiveness. Women who wear lipstick are considered more attractive, albeit the intensity and shade of the colour may influence attractiveness, too (26, 32). Furthermore, people who wear glasses are reportedly labelled as less attractive (27), less friendly, but also as more intelligent and duller (33) and were

Table 2. Mean, median, and minimum and maximum estimated age scores of all settings

	Estimated age (years)		
	Mean	Median	Min; max
Misaligned teeth	28.2	28.7	19.2; 37.4
Wearing earrings	28.4	28.8	19.8; 35.9
Wearing glasses	26.9	27.1	17.6; 35.7
Wearing lipstick	28.6	28.7	20.5; 36.1
Aligned teeth	28.6	28.8	20.4; 37.0
Wearing earrings	28.4	29.2	20.0; 34.8
Wearing glasses	27.2	26.9	19.0; 33.0
Wearing lipstick	28.9	29.3	19.9; 37.0

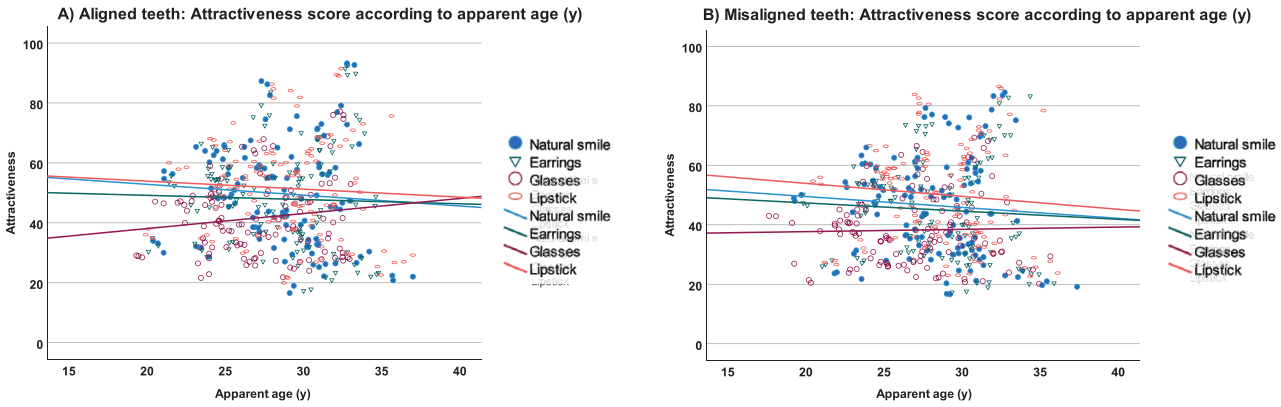

Figure 3. Scatterplots of all individuals—and all settings—with aligned teeth (A) and misaligned teeth (B). Attractiveness plotted against estimated age. Color-coded regression lines according to the respective setting.

even found less guilty when committing a crime (34). Here, too, the beneficial outcome of lipstick wear and detrimental effect of eyeglasses on facial attractiveness is seen, but these influences seem modulated by age. The older a person is seen as, the smaller these influences become.

Perhaps surprisingly, earrings did not enhance facial attractiveness. This finding is difficult to interpret, and no scientific investigation can be found in the English literature on earrings' impact on facial attractiveness. A possible interpretation would be that earrings tend to distract from attractive facial traits by emphasizing parts of the face that are generally less attractive, or that the selected pair of earrings of golden hoops were simply not appealing, as they were relatively large, and more delicate earrings would have been deemed more attractive.

The effects of tooth position or other modifications on perceived age (which is based on facial appearance and does not necessarily correspond to chronological age) were all under 1.5 years, and as such clinically not of any major relevance. However, certain tendencies are discernable, but restraint must be applied so as to not over-interpret these findings. Wearing lipstick made the participant look slightly older, i.e. more adult, while wearing glasses made her appear younger. This observation was unrelated to whether the teeth were aligned or not.

The results portray a clear picture regarding the relationship between attractiveness and apparent age, as a distinct overall trend of reduced attractiveness in older-looking participants can be detected both for aligned and misaligned cases alike. This comes as no surprise. A youthful facial appearance makes women look more attractive, and facial attractiveness is known to decrease in the age range that was assessed in the present study (31).

Limitations

Only females were assessed, so while this study may corroborate the validity of CNN-based scoring of facial attractiveness and age prediction, the actual results can obviously not be inferred on males.

This present research also differs in some aspects from other previous studies that have investigated facial and dental aesthetics. First, the tooth position was changed directly in the subject's mouth using an individual mock-up. While mock-ups are convenient to mimic different malocclusions or various degrees of crowding, other features affecting attractiveness, such as the smile arc, cannot be altered. Also, a single (rather darker) tooth colour was chosen, which did certainly not match the tooth colour of all participants. As such, the subtle change of tooth brightness cannot be excluded as possible influencing factor. Therefore, as a general caveat, any direct comparison of the results of this investigation to results of studies that are based on digitally changed tooth position or on patients before and after an orthodontic treatment, should be made with caution.

Lastly, it cannot be overemphasized that using AI to determine facial attractiveness will not generate any new truth, and it will remain only as indicative as the data it relies on. The advantage of using a multitude of subjective ratings as data, even in the millions, lies thus solely in being able to deliver a reliable portrait of what society currently considers beautiful, or—more accurately—what its members are attracted to. The CNNs are trained and expected to reflect the millions of ratings they were trained on. The situation is somewhat more straight-forward when it comes to apply AI for the evaluation of a perceived age, as age prediction is based on facial features that are not based on human ratings.

As a final general note, one last important caveat must be addressed. The present research paper does not propose the use of the presented AI solution as a medical application for diagnostics or treatment planning, and one must be cognisant that any AI-based software should be licensed by an appropriate regulatory body if its intended usage is for medical purposes.

Conclusions

CNNs can be trained to reliably assess the impact of tooth alignment and diverse facial modifications on facial attractiveness and perceived age in females. Alignment of upper front teeth improves facial attractiveness significantly, to a comparable degree than wearing lipstick. Wearing eyeglasses substantially reduces attractiveness. There was no discernable effect on perceived age for neither tooth alignment, nor all other assessed facial modifications.

Funding

This study was supported with a research grant of the Swiss Association of Dentomaxillofacial Radiology/SADMFR (grant no. 19/01).

Conflicts of interest

None to declare.

Data availability

A code repository providing trained models and codes for attractiveness and age predictions is publicly available at https://github.com/2006pmach/facial_attractiveness_prediction.

References

1. Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M. and Smoot, M. (2000) Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.
2. Goldman, W. and Lewis, P. (1977) Beautiful is good: evidence that the physically attractive are more socially skillful. *Journal of Experimental Social Psychology*, 13, 125–130.
3. Foo, Y.Z., Simmons, L.W. and Rhodes, G. (2017) Predictors of facial attractiveness and health in humans. *Scientific Reports*, 7, 39731.
4. Mulhern, R., Fieldman, G., Hussey, T., Lévêque, J.-L. and Pineau, P. (2003) Do cosmetics enhance female Caucasian facial attractiveness?. *International Journal of Cosmetic Science*, 25, 199–205.
5. Kiyak, H.A. (2000) Cultural and psychologic influences on treatment demand. *Seminars in Orthodontics*, 6, 242–248.
6. Patcas, R., *et al.* (2017) Motivation for orthognathic treatment and anticipated satisfaction levels—a two-centre cross-national audit. *Journal of Cranio-Maxillofacial Surgery*, 45, 1004–1009.
7. Olsen, J.A. and Inglehart, M.R. (2011) Malocclusions and perceptions of attractiveness, intelligence, and personality, and behavioral intentions. *American Journal of Orthodontics and Dentofacial Orthopedics*, 140, 669–679.

8. Richards, M.R., *et al.* (2015) Contribution of malocclusion and female facial attractiveness to smile esthetics evaluated by eye tracking. *American Journal of Orthodontics and Dentofacial Orthopedics*, 147, 472–482.
9. Havens, D.C., McNamara, J.A., Jr., Sigler, L.M. and Baccetti, T. (2010) The role of the posed smile in overall facial esthetics. *The Angle Orthodontist*, 80, 322–328.
10. Papio, M.A., Fields, H.W., Jr., Beck, F.M., Firestone, A.R. and Rosenstiel, S.F. (2019) The effect of dental and background facial attractiveness on facial attractiveness and perceived integrity and social and intellectual qualities. *American Journal of Orthodontics and Dentofacial Orthopedics*, 156, 464–474 e1.
11. Shaw, W.C., Rees, G., Dawe, M. and Charles, C.R. (1985) The influence of dentofacial appearance on the social attractiveness of young adults. *American Journal of Orthodontics*, 87, 21–26.
12. Wędrychowska-Szulc, B. and Syryńska, M. (2010) Patient and parent motivation for orthodontic treatment—a questionnaire study. *European Journal of Orthodontics*, 32, 447–452.
13. Patcas, R., Bernini, D.A.J., Volokitin, A., Agustsson, E., Rothe, R. and Timofte, R. (2019) Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age. *International Journal of Oral and Maxillofacial Surgery*, 48, 77–83.
14. Patcas, R., *et al.* (2019) Facial attractiveness of cleft patients: a direct comparison between artificial-intelligence-based scoring and conventional rater groups. *European Journal of Orthodontics*, 41, 428–433.
15. Peck, C.J., *et al.* (2021) Commercial artificial intelligence software as a tool for assessing facial attractiveness: a proof-of-concept study in an orthognathic surgery cohort. *Aesthetic Plastic Surgery*. doi: 10.1007/s00266-021-02537-4
16. Shen, J., *et al.* (2019) Artificial intelligence versus clinicians in disease diagnosis: systematic review. *JMIR Medical Informatics*, 7, e10010.
17. Shaw, W.C., Richmond, S., O'Brien, K.D., Brook, P. and Stephens, C.D. (1991) Quality control in orthodontics: indices of treatment need and treatment standards. *British Dental Journal*, 170, 107–112.
18. Richmond, S., *et al.* (1992) The development of the PAR Index (Peer Assessment Rating): reliability and validity. *European Journal of Orthodontics*, 14, 125–139.
19. Rothe, R., Timofte, R. and van Gool, L. (2016) Some like it hot—visual guidance for preference prediction. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5553–5561.
20. Mathias, M., Benenson, R., Pedersoli, M. and van Gool, L. (2014) Face detection without bells and whistles. In Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds), *Computer Vision—ECCV 2014. ECCV 2014. Lecture Notes in Computer Science*. Springer International Publishing, Cham, Switzerland, Vol. 8692, pp. 720–735.
21. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261–2269.
22. Rothe, R., Timofte, R. and Van Gool, L. (2018) Deep expectation of real and apparent age from a single image without facial landmarks. *International Journal of Computer Vision*, 126, 144–157.
23. Russakovsky, O., *et al.* (2015) ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115, 211–252.
24. Agustsson, E., Timofte, R., Escalera, S., Baro, X., Guyon, I. and Rothe, R. (2017) Apparent and real age estimation in still images with deep residual regressors on APPA-REAL database. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). doi: 10.1109/FG.2017.20
25. Yami, E.A.A., Kuijpers-Jagtman, A.M. and van 't Hof, M.A. (1998) Assessment of dental and facial aesthetics in adolescents. *European Journal of Orthodontics*, 20, 399–405.
26. Stephen, I.D. and McKeegan, A.M. (2010) Lip colour affects perceived sex typicality and attractiveness of human faces. *Perception*, 39, 1104–1110.
27. Terry, R.L. and Kroger, D.L. (1976) Effects of eye correctives on ratings of attractiveness. *Perceptual and Motor Skills*, 42, 562–562.
28. Benson, P.E., Javidi, H. and DiBiase, A.T. (2015) What is the value of orthodontic treatment?. *British Dental Journal*, 218, 185–190.
29. Johal, A., Alyaqoobi, I., Patel, R. and Cox, S. (2015) The impact of orthodontic treatment on quality of life and self-esteem in adult patients. *European Journal of Orthodontics*, 37, 233–237.
30. Bernabé, E., Sheiham, A., Tsakos, G. and Messias de Oliveira, C. (2008) The impact of orthodontic treatment on the quality of life in adolescents: a case–control study. *European Journal of Orthodontics*, 30, 515–520.
31. Tatarunaite, E., Playle, R., Hood, K., Shaw, W. and Richmond, S. (2005) Facial attractiveness: a longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 127, 676–682.
32. Tagai, K., Ohtaka, H. and Nittono, H. (2016) Faces with light makeup are better recognized than faces with heavy makeup. *Frontiers in Psychology*, 7, 226.
33. McKelvie, S.J. (1997) Perception of faces with and without spectacles. *Perceptual and Motor Skills*, 84, 497–498.
34. Brown, M.J., Henriquez, E. and Groscup, J. (2008) The effects of eyeglasses and race on juror decisions involving a violent crime. *American Journal of Forensic Psychology*, 26, 25–43.