Analyzing VGG-19's Bias in Facial Beauty Prediction: Preference for Feminine Features

Nuno Fernandes[1], Sandra Soares[2], Joana Arantes[3]
School of Psychology, University of Minho, Braga, Portugal[1, 3]
Department of Education and Psychology, University of Aveiro, Aveiro, Portugal[2]

***Abstract*—From an evolutionary perspective, sexual dimorphism has been linked to perceived attractiveness, with masculine traits preferred in men and feminine traits in women. Moreover, symmetry is a strong predictor of facial attractiveness across both sexes. Recent advancements in the field of artificial intelligence have enabled algorithms to accurately predict facial attractiveness. This study aims to investigate whether these algorithms accurately replicate human judgments of attractiveness. We hypothesized that sexually dimorphic manipulations (masculinized men and feminized women) (H1), as well as symmetrized versions (H2), would elicit higher attractiveness ratings from a facial beauty prediction algorithm. Employing transfer learning, we trained six deep-learning models using four facial databases with attractiveness ratings (*n* = 6848). The top-performing model, VGG-19, demonstrated a high prediction correlation of .86 on the test set. Surprisingly, our findings revealed an interaction effect between sex and sexual dimorphism. Feminized versions of both men's and women's faces obtained higher attractiveness ratings than their masculinized counterparts. For symmetry, our results indicated that symmetrized faces were perceived as more attractive, albeit exclusively among women. These findings offer novel insights into the understanding of facial attractiveness from both algorithmic and human behavioral perspectives.**

Keywords—Deep learning; facial attractiveness; sexual dimorphism; symmetry; VGG-19

I. INTRODUCTION

What makes a pretty face? The study of facial attractiveness is a multidisciplinary field that draws on knowledge from diverse disciplines, such as psychology, sociology, anthropology, and computer science [1-3]. Researchers in these disciplines employ diverse methods and approaches to explore the factors that contribute to facial attractiveness and its impact on social interactions and relationships.

In psychology, researchers investigate how people perceive and evaluate facial attractiveness, as well as the cognitive and neural mechanisms underlying this process [3]. They also explore the relationship between facial attractiveness and social cognition, such as the formation of impressions and romantic relationships [4-5]. Studies have shown that humans prefer associating with, dating, and mating with individuals considered facially attractive [6-8]. Additionally, attractive people tend to be perceived as more successful, enjoyable, and intelligent than unattractive people [4]. Specifically, research on human preferences for sexually dimorphic faces has garnered significant attention [9-10]. Sexual dimorphism in facial attractiveness refers to the physical differences in facial features that are considered more attractive in men and women. These differences in facial attractiveness are thought to be influenced by both evolutionary and cultural factors [see [11] for a comparison across five populations].

In sociology and anthropology, researchers explore how facial attractiveness and sexual dimorphism are related to cultural norms and values and how they shape social interactions regarding social status, power, and privilege [1]. Cross-cultural studies suggest a high consensus in facial attractiveness judgments across different populations [12-13]. However, research also indicates that cultural factors can shape perceptions of facial attractiveness, with different societies having distinct standards for what is considered attractive [14].

In computer science, computer vision researchers study how to create algorithms and models to predict facial attractiveness [2]. Explaining the functioning of deep learning models has recently gathered increasing attention from researchers and the wider public, including regulators and politicians [15-16]. While attractiveness prediction is of great use for the human-computer interaction field, it also raises significant moral and ethical questions. One of the primary concerns is whether beauty prediction algorithms reveal human-like psychophysical biases [17-18]. The study in [17] found that machine judgments of attractiveness displayed a preference for averaged faces and symmetrical faces, mirroring human judgments of attractiveness [19-20]. Additionally, the prediction of facial attractiveness holds clinical significance, particularly in the realm of plastic surgery [21-23]. To gain a comprehensive understanding of the complex issue of facial attractiveness, advancements in facial beauty prediction algorithms offer valuable insights. Furthermore, incorporating input from cognitive scientists in experimental design can enhance our understanding, especially in uncovering the decision-making processes behind algorithmic outputs.

A. *Sexual Dimorphism and Attractiveness*

From an evolutionary standpoint, males and females have evolved unique differences in their secondary sexual characteristics over time [24]. The development of more pronounced sexual dimorphic phenotypes – feminine traits for women and masculine traits for men – is considered attractive [9, 25], because it implies the inheritance of advantageous genes favorable to offspring survival or reproductive success [26-27].

Evolutionary theories suggest that preferences for sexually dimorphic traits are associated with specific mating functions. For example, men may possess facial characteristics to indicate strength and dominance, such as an enlarged brow ridge, a thicker jawline, and a wider face [28]. Dominant men are considered more attractive by women [29-30], as they often achieve higher social status, thereby enhancing their capacity to provide essential resources for reproduction [31]. Additionally, masculinity is desirable as it signals higher testosterone levels [14, 32]. Since testosterone is immunosuppressive, compromising the body's ability to fight infections [33], only men with high genetic quality and robust immune systems can afford to invest in secondary sexual traits [34-35]. Therefore, men with more masculine traits are often perceived as healthier [36-37], and more attractive to potential mates [38-40].

Conversely, women may signal youth and fertility [41]. Feminine facial traits, such as a smaller jawline and a narrower face [26], indicate higher levels of estrogen, which correlate with superior reproductive qualities [42] and mating desirability [9, 25]. For instance, research has shown that both men and women rated photographs of women's faces captured during the fertile window of the menstrual cycle (late follicular) as more appealing compared to photographs taken during the non-fertile (luteal) phase [43]. Similarly, [44] explored the impact of facial shape transformations towards late follicular and luteal prototypes on male perceptions of attractiveness. Their findings revealed a distinct preference for faces resembling the late follicular phase, suggesting that subtle shape differences can sway men's preferences depending on a woman's menstrual cycle phase. During the late follicular period, a peak of estrogen and luteinizing hormone (LH) leads to ovulation and is followed by a rise in progesterone in the luteal phase [45-47]. This increased attractiveness may serve as an adaptive mechanism to enhance a female's perceived value in the mating pool during the phase of the cycle when the likelihood of conception is at its peak [43]. Ultimately, the preference for women's feminine facial features remains robust across the literature [10, 48-52].

B. *Algorithms and Attractiveness*

Early approaches to quantifying facial attractiveness were grounded in the notion that specific ratios and proportions of facial features are more appealing [17, 53-54] For example, the "golden ratio" or "phi" (1.618) has been used to measure facial proportions and attractiveness [56]. The idea is that if specific facial ratios, such as the distance between the eyes and the distance between the mouth and eyes, conform to phi, the face is more attractive. However, contemporary computer-based methods, including machine learning and deep learning algorithms, have shown more promising outcomes.

Machine learning offers a method that can be used to predict facial attractiveness by analyzing patterns in predefined facial features known to be associated with attractiveness [17, 53-54]. These models assess multiple hand-crafted features, such as symmetry, geometric ratios, and distances between landmarks. Another approach uses deep learning algorithms, such as Convolutional Neural Networks (CNNs), which also predict facial attractiveness. One advantage of CNNs is that they do not require pre-defined facial attributes, as they can extract useful features directly from raw images. Consequently, these models have achieved a relatively high correlation with the actual judgments of attractiveness [56-58].

Furthermore, leveraging transfer learning eliminates the need to develop an algorithm from the ground up, as it allows for the use of pre-existing models trained on large datasets, such as architectures from the Visual Geometry Group Network (VGG; [59]) or residual networks (ResNet) ([60]). Although these models were initially designed for image classification tasks, they can be fine-tuned for facial beauty prediction [61-63]. The fine-tuning process involves adjusting the model's parameters to better fit the new task, typically by training the final layers of the model on the new dataset. Once the model is fine-tuned, it can predict the attractiveness of new facial images. This approach has been widely used in attractiveness prediction tasks as it saves computational time and resources while improving performance [61-63].

The benchmark for training and evaluating facial beauty prediction algorithms is the SCUT-FBP5500 [64], with cutting-edge models achieving a Pearson correlation of .93 with human judgments of attractiveness [57]. However, despite the large number of images (n = 5500), attractiveness ratings were performed only by Asian raters. Therefore, this could be a potential limitation for generalizing the beauty prediction algorithms. Additionally, the high accuracy observed could be attributed to greater internal consistency in attractiveness judgments within the Asian population [10].

C. *The Present Study*

Despite significant advancements in facial beauty prediction algorithms, a gap remains in the literature regarding how these algorithms make their decisions. The main objective of this study is to investigate whether beauty prediction algorithms replicate human biases toward sexually dimorphic facial traits. We hypothesize that sexually dimorphic versions of male (masculinized) and female (feminized) facial images will be perceived as more attractive (H1). Moreover, we will manipulate symmetry, which is a well-established predictor of facial attractiveness [19,65-67]. We expect that symmetrized facial images of both men and women will be rated as more attractive (H2). Our secondary aim is to provide a benchmark for facial attractiveness prediction using multiple databases. This novel approach aims to enhance our understanding of how facial beauty prediction algorithms assess attractiveness and whether they support the sexual dimorphism and symmetry hypotheses.

II. METHOD

A. *Materials*

1) Databases: For the transfer learning phase, four databases were used: SCUT-FBP5500 [64]; Chicago Face Database (CFD) [68]; Karolinska Directed Emotional Faces (KDEF) [69]; and FACES [70], resulting in a total of 6,848 images. Attractiveness ratings across all datasets were normalized to a scale from 0 to 1. An additional database, the Face Research Lab London (FRL-London; [71]), was used to create sexually dimorphic and symmetrized versions of faces, which were then evaluated for facial attractiveness by the trained model. Following common preprocessing practices for image data in deep learning tasks [59-60] pixel values were rescaled to the range of (0, 1) by dividing each pixel intensity value by 255. This normalization step ensures that the input features are appropriately scaled and facilitates efficient training of machine learning models, particularly neural networks, by mitigating issues such as vanishing or exploding gradients [72-73].

The SCUT-FBP5500 [64] dataset consists of 5500 frontal face images, including 2000 Asian females, 2000 Asian males, 750 Caucasian females, and 750 Caucasian males. Most of the faces have neutral expressions and simple backgrounds. The faces in the dataset were labeled with attractiveness ratings from 1 to 5 by 60 Asian raters aged 15 to 60.

The Chicago Face Database (CFD) [68] consists of 970 facial images displaying neutral, happy, threatening, and fearful expressions. The 158 volunteers who photographed included 37 Black males, 48 Black females, 36 White males, and 37 White females, ranging in age from 18 to 40 years. A sample of 1,087 raters made subjective ratings of the image's attractiveness on a 1–7 Likert scale, with each subject rating only 15 faces. Participants included 552 females, 308 males, and 227 who preferred not to disclose their gender, with an average age of 26.75 ($SD_{age} = 10.54$). They came from diverse racial backgrounds: 516 White, 117 Asian, 74 Black, 72 biracial or multiracial, 57 Latino, 18 other, and 233 who did not report their race.

The Karolinska Directed Emotional Faces (KDEF) [69] is one of the most widely used databases of human facial expressions. [74] presented subjective attractiveness ratings for a subset of 210 pictures from 70 Caucasian amateur actors (35 women and 35 men, aged between 20 and 30 years) with different facial expressions: angry, happy, and neutral. The sample of raters included 155 students from Portuguese universities (83.20% female; $M_{age} = 23.73$; $SD_{age} = 7.24$). Each participant rated the facial attractiveness of 36 pictures on a 1-7 Likert scale.

The FACES [70] is a set of facial images from 171 women and men categorized into three age groups: young ($n = 58$; age range: 19–31; $M_{age} = 24.3$; $SD_{age} = 3.5$), middle-aged ($n = 56$; age range: 39-55; $M_{age} = 49.0$; $SD_{age} = 3.9$), and older ($n = 57$; age range: 69-80; $M_{age} = 73.2$; $SD_{age} = 2.8$). Each individual displays six facial expressions: neutrality, sadness, disgust, fear, anger, and happiness, resulting in a total of 1026 pictures. Facial attractiveness evaluations were performed by 154 raters, all of whom were Caucasian and German. Each participant was randomly assigned to either set A or set B, which consisted of identical pictures of the same subjects and facial expressions, with only minor differences (e.g., head inclination angle). For the present work, only set A was considered. Each image was rated by a minimum of 8 and a maximum of 14 raters per age group by gender. Therefore, six attractiveness ratings were provided in the metadata of the original work. In our case, we used the average of these ratings on a scale from 1 to 100.

The Face Research Lab London (FRL-London) was chosen for the sexual dimorphic transformations [71], which consists of 102 neutral front faces (male = 52; female = 50; $M_{age} = 26.9$; $SD_{age} = 7.07$) from different ethnicities (69 White, 13 Black, and 20 Asian). We used only 100 photos from this database due to problems manipulating two male images (codes 005_03 Asian and 114_03 Black), resulting in an equal number of male and female facial images. This dataset included pre-delineated face shape templates of 189 coordinates for each image. Attractiveness ratings (on a 1-7 scale) were made by 2,513 people (age range: 17-90).

2) Algorithms: A set of six facial recognition models was used to determine which of these models would be the most adequate for facial attractiveness prediction. The selected models were: InceptionResNetV2 [75], MobileNetV2 [76], EfficientNetV2B0 [77], ResNet50 [60], Xception [78], and VGG-19 [59].

B. Procedure

The procedure was divided into three phases. Initially, we employed a transfer learning methodology across a range of models, followed by fine-tuning the one that yielded the most favorable results. In the second phase, we applied sexual dimorphic and symmetrical transformations of the images from the FRL-London dataset for further evaluation. Finally, we utilized the fine-tuned model to predict the facial attractiveness of the original, sexually dimorphic, and symmetrized versions of both female and male facial photos.

1) Transfer learning: We selected a set of six commonly used architectures with pre-trained weights for image classification problems from the Keras Applications Module: tf.keras.applications [79]. Then, we applied transfer learning to each model to predict facial attractiveness, similar to previous studies [61-63]). This involved freezing all layers except the last one, and transforming the initial classification problem into a regression problem by incorporating a final dense layer.

The models were trained for 100 epochs with patience set to 30 epochs, using the Adam optimizer with a learning rate of 0.001 and weight decay of 0.004. Adam is an "algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments" [80]. A train/validation/test split of 60/20/20 size was used, resulting in 4110/1369/1369 images for the corresponding set. The model with the best results was the VGG-19 with a Pearson's correlation of .76 ($RMSE = 0.094$; $MAE = 0.07$).

The VGG-19 belongs to a series of deep neural networks from the Visual Geometry Group Network (VGG), also including VGG-11, VGG-13, and VGG-16 ([59]). These networks share their structure by having several blocks of convolutional layers connected to a final block consisted of three fully connected layers. The VGGNet has been trained on over one million images in 1000 classes [59]. Specifically, VGG-19 has five blocks of convolution layers and one last block of three fully connected layers. A 3x3 Max Pooling layer connects all blocks. The first (size 64x64x64) and second (size 128x128x128) blocks have two convolutional layers, while the third (size 256x256x256), fourth (size 512x512x512) and fifth (size 512x512x512) blocks have four convolutional layers. Since VGG-19 showed the best performance for predicting facial attractiveness compared to the other architectures, we fine-tuned this model.

Therefore, we unfroze the 3-fully connected layers, whereas in the previous phase only the last regression layer was trainable. We used the same hyperparameters and train/validation/test sets. The model's performance on the test set increased from a Pearson's correlation of .76 to .84 (*RMSE* = 0.12; *MAE* = 0.10). Finally, we fine-tuned the model with the best parameters by unfreezing all the model's layers. We set the training for 10 epochs (patience of 5), using Adam optimizer with a learning rate of 0.00001 and weight decay of 0.004, to prevent overfitting [see [81] for a revision on hyperparameter optimization for fine-tuning]. A final correlation of .86 ($p < .001$) was achieved between the model's predictions and the test set (*RMSE* = 0.10; *MAE* = 0.08).

2) Image preprocessing: To achieve sexual dimorphic and symmetric transformations, we used the FRL-London [71] database. All image manipulations were performed using computer vision techniques provided by the Psychomorph software [82] and its web-based version, WebMorph [83]. Following previous research [52, 84-85], sexually dimorphic versions of each image were created using a 50% spatial linear transformation towards either a female avatar (feminized version) or a male avatar (masculinized version). The avatars (average faces) were generated by averaging 30 male faces (male avatar) and 30 female faces (female avatar). Additional information regarding theses methods can be found in [86]. Symmetrized versions of both the original and manipulated photos (feminized and masculinized faces) were created by mirror reversing the image and then combining one side of the original non-symmetrized photo with the opposite side of the reversed photo [67, 87].

3) Data analysis: The algorithm's prediction of attractiveness was used as the dependent variable in a linear mixed-effects model (LME) with three fixed main effects: sexual dimorphism manipulation (masculinized/feminized), symmetry (original/symmetric), and sex (male/female), along with the three-way interaction term. The photo's code was included as a random factor with varying intercepts to account for individual differences in attractiveness, as the model predicted attractiveness for multiple versions (sexually dimorphic and symmetric transformations) of the same individual [see [88] for a detailed explanation on LME models]. A post-hoc analysis using Bonferroni adjustment was performed for the observed significant effects. All statistical analyses were conducted using R Statistical Software (v4.3.1; [89]) using lme4 (v.1.1.34; [90]) and ggplot2 (v.3.4.4.9000; [91]) for plotting. The data and reproducible code are publicly available at [92].

III. Results

The LME (estimated using REML) was composed of both main effects (sexual dimorphism manipulation, symmetry, and sex) and its three-way interaction term (Fig. 1). The model included the image's code as a random factor. The model's total explanatory power was substantial (conditional $R^2 = .93$), and the part related to the fixed effects alone (marginal R^2) was equal to .33. The model showed a statistically significant main effect of sexual dimorphism, $F(2, 490) = 64.02$, $p < .001$, symmetry, $F(2, 490) = 32.07$, $p < .001$, and sex, $F(1,98) = 49.74$, $p < .001$. For the interaction terms, the two-way interaction of sexual dimorphism by symmetry manipulation was shown to be statistically significant, $F(2,490) = 5.40$, $p = .005$, while the two-way interaction of sexual dimorphism and sex, $F(2,490) = 2.73$, $p = .066$, or the triple interaction of sexual dimorphism*symmetry*sex, $F(2,490) = 0.91$, $p = .40$, showed no significant results.

Post-hoc analysis using Bonferroni correction showed that feminized faces were rated as more attractive for both male and female images (H1). In the non-symmetric condition, feminized females were judged as more attractive than masculinized females, $t(490) = 4.41$, $p < .001$, 95% CI [0.01, 0.02], $d = 0.88$, as well as in the symmetric condition, $t(490) = 5.44$, $p < .001$, 95% CI [0.01 0.02], $d = 1.09$. Also, feminized males in the non-symmetric condition were rated as more attractive than masculinized ones, $t(490) = 5.47$, $p < .001$, 95% CI [0.01, 0.02], $d = 1.09$, as well as in the symmetric condition, $t(490) = 6.34$, $p < .001$, 95% CI [0.01 0.03], $d = 1.27$. Additionally, feminized manipulations were perceived as more attractive than the original faces for all conditions, except for females in the non-symmetrized version, $t(490) = 0.77$, $p = 1$, 95% CI [-0.01, 0.01], $d = 0.15$. Similarly, no differences were found between the masculinized versions and the original faces, except for that condition in which masculinized female photos were judged as less attractive than the original ones in the non-symmetrized versions, $t(490) = -3.64$, $p = .007$, 95% CI[-0.02, -0.01], $d = -0.73$.

For the symmetry main effect that was observed (H2), post-hocs revelated that symmetrized versions were only more attractive for feminized female faces, $t(490) = 4.43$, $p < .001$, 95% CI [-0.02, -0.01], $d = 0.89$, and masculinized female faces $t(490) = -3.41$, $p = .02$, 95% CI [-0.02, -0.01], $d = 0.58$, in comparison with the non-symmetrized ones. No differences were observed for males or between original photos.

Finally, supporting the sex main effect, in the non-symmetrized condition, feminized females were perceived as more attractive than feminized males, $t(118) = 6.12$, $p < .001$, 95% CI [0.03, -0.08], $d = 3.71$, with the same occurring for masculinized versions, $t(118) = 6.47$, $p < .001$, 95% CI [0.03, 0.08], $d = 3.92$, and original faces, $t(118) = 7.28$, $p < .001$, 95% CI [0.04, 0.09], $d = 4.42$. In the symmetrized condition, the same pattern was found, with feminized female photos being rated as more attractive than feminized male photos, $t(118) = 6.61$, $p < .001$, 95% CI [0.03, 0.08], $d = 4.01$, as well as masculinized versions, $t(118) = 6.90$, $p < .001$, 95% CI

[0.03, -0.09], $d = 4.19$, and original ones, $t(118) = 6.98$, $p < .001$, 95% CI [0.03, 0.09], $d = 4.23$.

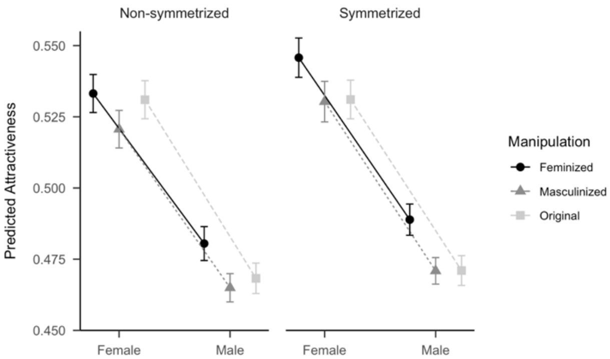

Fig. 1. The attractiveness ratings of VGG-19.

IV. Discussion

The main aim of the present study was to investigate whether beauty prediction algorithms replicate human judgments concerning facial sexually dimorphic traits. We hypothesized that sexually dimorphic versions of male (masculinized) and female (feminized) facial pictures would be rated as more attractive (H1). Also, we expected that symmetrized versions of both male and female facial pictures would be perceived as more attractive (H2). Additionally, we aimed to provide a benchmark for facial attractiveness prediction using multiple databases.

The fine-tuned model was used to predict the attractiveness of symmetrized and non-symmetrized versions of masculinized, feminized, and original frontal facial images with neutral expressions from the FRL-London database. Surprisingly, our results showed that feminized manipulations were rated as more attractive compared to the masculinized versions for both male and female faces. Thus, we partially confirmed our first hypothesis (H1), as we expected a sexual dimorphism main effect, with feminized versions being preferred for females, but not for males. Moreover, despite the observed main effect of symmetry, post-hoc analyses suggested that symmetry increased facial attractiveness specifically for feminized and masculinized versions of female faces. Thus, the data also partially supported our second hypothesis (H2). Additionally, across all conditions, females were consistently rated as more attractive than males.

We achieved a Pearson correlation of .86 with the fine-tuned VGG-19 on the test set across four databases. Although slightly lower than state-of-the-art results (e.g., [57] with $r = .93$), our approach addresses a key limitation by reducing bias from relying solely on the SCUT-FBP5500 dataset. This offers a valuable benchmark for facial beauty prediction, encouraging future research to explore and improve on these findings.

A. *Theoretical Implications*

Firstly, our work contributes to understanding how facial beauty prediction algorithms make decisions. The consistent overall preference for feminine facial traits suggests that beauty prediction algorithms can implicitly learn human attractiveness preferences, like [17] findings. More recently, [18], using the VGG architecture, showed that DNNs use putative ratios (e.g., golden ratio) as essential attributes for predicting facial attractiveness, akin to human judgments. This convergence between algorithmic and human judgments underscores the complex interplay between facial features, attractiveness perception, and evolutionary pressures. Further investigation into the mechanisms underlying both facial attractiveness prediction and sexual dimorphism is therefore warranted.

1) Sexual dimorphism: Regarding our first hypothesis (H1), and despite our initial assumptions, feminized versions of both male and female faces were rated as more attractive. The preference for femininity in women is well-supported in the existing body of literature [10, 48, 50, 93-94]. However, there are mixed findings concerning the effect of sexual dimorphism on male attractiveness. Some research reports a stronger preference for more masculine male faces [95-100], while other studies suggest a preference for more feminine faces [10,37,97,101-104], or report no differences between masculinized and feminized versions of male faces [50,105-106].

A possible explanation for these diverse findings may be associated with the level of attractiveness from both the rater and the photo. For instance, [107] found that when judging attractive faces, both male and female participants favored masculine male faces over feminine ones. Conversely, in the less attractive condition, they preferred feminine male faces to masculine ones. Furthermore, [108] found that women's preferences for femininity in men's faces decreased after viewing images of highly attractive men. Since the predicted attractiveness in our results may be considered average, this could explain the lack of preference for masculinity. Moreover, there is an increased preference for masculinity among women who perceive themselves as attractive [102]. Furthermore, women perceived as having relatively low facial attractiveness by others, showed a preference for more "feminine" male faces when selecting partners for long-term relationships compared to short-term relationships [85].

Furthermore, enhancing masculine facial features may increase perceptions of dominance and negative attributes, such as coldness or dishonesty [10]. These aspects are considered significant by women within the context of relationships and paternal investment [10]. Subsequent research has unveiled that women prefer less masculine male faces in long-term relationships under specific conditions of environmental harshness, such as resource scarcity or elevated stress levels [39,109]. These findings align with the notion that men with more favorable genes (higher testosterone levels) invest less in the relationship, as they can afford to choose their mate [109].

However, it is important to note that this aspect remains ambiguous within the existing body of literature. In a recent study, women exhibited a preference for masculinized faces over feminized ones [100]. Notably, this preference for masculine male faces was more pronounced when assessing potential co-parenting partners compared to short-term relationships [100]. Supporting these findings, [110] demonstrated that facial attractiveness, instead of facial masculinity influenced perceptions of paternal involvement.

Finally, hormonal changes occurring during the menstrual cycle may alter women's preference for sexually dimorphic cues. In women, the preference for masculinity is higher during the late follicular phase of the menstrual cycle (when fertility is high and progesterone level low) [111]. Similarly, [112] found that women tend to prefer masculinity in male faces around ovulation, while they prefer feminized male faces during the rest of the cycle. This likely indicates a balanced trade-off between the attraction to males who appear to ensure reproductive success and males who are perceived as capable and nurturing fathers.

2) Symmetry: Concerning our second hypothesis (H2), our results showed that the facial beauty prediction algorithm found symmetrized faces more attractive, but only for sexually dimorphic versions of female faces. Similar to our findings, machine learning models trained to predict facial attractiveness based on geometric facial features have been found to implicitly learn human psychophysical biases towards symmetric faces [53,113]. However, it is important to note that the specific characteristics that make symmetric faces more attractive, independent of their symmetry, are not definitively defined.

Since our results showed a preference for symmetric sex-dimorphic traits exclusively in female faces, insights from prior studies [97,114] suggesting an intercorrelation between symmetry and sexual dimorphism may shed light on why symmetrized male faces were not found more attractive than non-symmetrized versions. In a cross-cultural study, perceived sexual dimorphism demonstrated a correlation with symmetry measurements [97]. Additionally, research found that women's preferences for symmetry were positively correlated with preferences for masculinity in male faces and that men's preferences for symmetry were positively correlated with preferences for femininity in female faces [97]. Furthermore, [115] found that facial symmetry was associated with masculine facial cues in males, such as increased lower face length and cheekbone prominence [32,116]. Symmetry could potentially enhance the perception of masculinity or be interpreted by the algorithm as a cue of masculinity in male faces, which might explain why the effect was observed only for female faces.

Despite the critical role of symmetry in the perception of attractiveness, evidence suggests that sexual dimorphic cues may be more influential than symmetry in attractiveness evaluations, which supports our findings. For instance, [117] found that participants focused more on sexual dimorphic cues than on symmetry when selecting a partner for a romantic relationship. Furthermore, [118] altered symmetry within a face while keeping the mean size of facial features constant, revealing that symmetric faces were perceived as less attractive. Additionally, [119] found that symmetry enhances facial attractiveness by increasing perceived normality, suggesting that attractiveness perception may involve additional inferences about a face. Likewise, using a random forest machine learning algorithm to predict facial attractiveness, [20] found that shape averageness, dimorphism, and skin texture symmetry emerged as valuable features capable of yielding relatively precise predictions. However, it was observed that shape symmetry did not contribute significantly to the predictive accuracy.

3) Sex: Our findings indicate that the facial beauty prediction algorithm rated women as more attractive than men. This observation aligns with previous research that utilized neural networks with facial landmarks as input data [120], as well as studies based on human assessments of attractiveness [121-122]. Considering the sex effect observed, it could be argued that the algorithm's general preference for feminization may stem from an inclination towards feminine traits, regardless of the sex of the face being evaluated. However, several studies have highlighted the impressive capability of convolutional neural networks in sex recognition tasks, e.g., [123-124].

B. Limitations / Further Studies

There are some limitations mainly associated with the databases that warrant consideration. Novel research has revealed that beauty prediction algorithms trained using SCUT-FBP5500 images exhibit poor prediction accuracy when assessing faces from other databases [125]. There may be a cultural effect in our study, as the largest database consisted solely of Asian raters, potentially influencing preferences for masculinity in male faces. Research suggests that a preference for more masculine men is more pronounced in countries with higher economic and national health indices [103]. To prevent the model from overfitting to the characteristics of the SCUT-FBP5500 dataset, we included additional datasets with standardized frontal facial images.

Furthermore, there is an interrelationship between attractiveness and emotional expression [74,126,128]. Human raters consistently perceive attractive faces as more appealing [75,126], while positive emotions are more readily recognized in attractive faces [126-127]. While emotion recognition algorithms have advanced [128], they haven't yet integrated attractiveness ratings, and beauty prediction models typically overlook emotionality. Future research should explore whether algorithms replicate human biases regarding attractiveness and emotions, and examine their interaction with sexual dimorphism.

Additionally, integrating multiple databases with varying rating scales for facial beauty prediction can pose challenges, as noted by [125]. [129] addresses the challenge of standardizing Likert scales, noting that the mean may shift depending on the scale's point range. To address this issue, [129] suggests a rescaling system based on transitional probabilities. However, given that the scales from the selected datasets range from 1-5 to 1-100, we opted for a straightforward normalization to the 0-1 range. Nonetheless, using ordinal attractiveness judgments as scalar intervals presents a rescaling challenge. The interval properties of Likert scales are inherent to the data, influenced by observers' mean responses rather than the labels themselves [130]. Thus, we advocate for a unified rating system for attractiveness datasets to enhance predictive models.

In this study, VGG-19 was trained using attractiveness ratings from both male and female evaluators who assessed photos of individuals from both the same and opposite sexes. From an evolutionary perspective, the sexual dimorphism hypothesis posits that individuals seek high-quality partners for reproduction. Therefore, the same-sex attractiveness judgments in these datasets may represent a potential limitation. However, as previously discussed, [107] demonstrated congruent attractiveness preferences among men and women, indicating a preference for masculinity in attractive male faces, whereas feminine traits were preferred in less attractive male faces. Additionally, earlier research, such as that by [131], did not find differences between same-sex and opposite-sex attractiveness preferences. Supporting the preference for sexual dimorphism, [81] reported that male participants rated masculinized versions of male faces as more attractive than feminine faces. Moreover, the self-rated sex-typicality appears to moderate this effect. For instance, individuals who perceive themselves as possessing more exaggerated sex-typical traits tend to exhibit stronger preferences for exaggerated sex-typical shape cues in faces of the same sex [132]. Further studies may explore whether same-sex or opposite-sex ratings could explain the observable effects. Therefore, researchers should consider incorporating this factor when developing future databases of facial attractiveness.

Furthermore, sexual orientation may influence preferences for sex-dimorphic face shapes. Research indicates that homosexual men tend to have a stronger preference for masculinity in male faces compared to other demographic groups [133]. However, there is evidence suggesting that sexual position moderates facial masculinity preferences [134-135]. Individuals' self-labeling as "top" or "bottom", as well as their preferences for their partner's sexual position, are indicative of their preferences for masculinity. Generally, individuals seeking a "top" as a romantic partner prefer more masculine faces than those seeking "bottoms" [134-135]. Additionally, homosexual women exhibited stronger preferences for masculinity in female faces compared to heterosexual women [133]. Thus, sexual orientation may be a crucial factor to consider in subsequent models. To account for this, we suggest that authors incorporate sexual orientation when developing databases of facial attractiveness and make this information available to other researchers.

V. Conclusion

The results of the present study align with previous findings in the psychological literature, which indicate a preference for femininity in both sexes. Moreover, we supported prior findings about a preference for symmetric faces. However, our results reveal an interplay between sexual dimorphism and symmetry, as the preference for symmetric faces was observed only in dimorphic faces, not in the original versions. Furthermore, our results provide a benchmark for developing beauty prediction algorithms using multiple databases. Future research could build on these findings by investigating additional cues of facial attractiveness. Overall, our findings suggest that beauty prediction algorithms can implicitly learn and replicate human biases regarding facial attractiveness.

Acknowledgment

This study was conducted at the Psychology Research Centre (CIPsi), School of Psychology, University of Minho, supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (Ref.: UIDB/PSI/01662/2020). This work was also supported by the Portuguese Foundation for Science and Technology (FCT) through the individual research grant 2021.05008.BD.

References

[1] Frevert, T. K., & Walker, L. S. (2014). Physical attractiveness and social status. *Sociology Compass, 8*(3), 313–323.

[2] 2d, J. N., & Abdulazeez, A. M. (2021). Facial beauty prediction and analysis based on deep convolutional neural network: A Review. *Journal of Soft Computing and Data Mining, 2*(1), 1–12.

[3] Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*(12), 452-460.

[4] Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*(3), 285–290.

[5] Talamas, S. N., Mavor, K. I., & Perrett, D. I. (2016). Blinded by beauty: Attractiveness bias and accurate perceptions of academic performance. *PLoS ONE 11*(2): e0148284.

[6] Buss, D. M. (1994). *The evolution of desire: Strategies of human mating*. Basic Books.

[7] Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success?. *Evolution and Human Behavior, 26*(2), 186-201.

[8] Riggio R., & Woll S. (1984). The role of non-verbal and physical attractiveness in the selection of dating partners. *Journal of Social and Personal Relationships, 1*(3), 347–357.

[9] Barber, N. (1995). The evolutionary psychology of physical attractiveness: Sexual selection and human morphology. *Ethology and Sociobiology, 16*(5), 395–424.

[10] Perrett, D., Lee, K., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., … & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887.

[11] Fiala, V., Třebický, V., Pazhoohi, F., Leongómez, J. D., Tureček, P., Saribay, S., … & Kleisner, K. (2021). Facial attractiveness and preference of sexual dimorphism: A comparison across five populations. *Evolutionary Human Sciences, 3*, E38.

[12] Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C.-H. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology, 68*(2), 261–279.

[13] Jones, D., & Hill, K. (1993). Criteria of facial attractiveness in five populations. *Human Nature, 4*, 271-296.

[14] Penton-Voak, I. S., Jacobson, A., & Trivers, R. (2004). Populational differences in attractiveness judgements of male and female faces: Comparing British and Jamaican samples. *Evolution and Human Behavior, 25*(6), 355-370.

[15] Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *Data Mining and Knowledge Discovery*, *11*(5), e1424.

[16] Samek, W., & Müller, K. R. (2019). Towards Explainable Artificial Intelligence. In: W. Samek, G. Montavon, A. Vedaldi, L. Hansen, K. R. Müller (Eds.) *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*(pp. 5-22). Springer.

[17] Kagian, A., Dror, G., Leyvand, T., Meilijson, I., Cohen-Or, D., & Ruppin, E. (2008). A machine learning predictor of facial attractiveness revealing human-like psychophysical biases. *Vision Research, 48*(2), 235-243.

[18] Tong, S., Liang, X., Kumada, T., & Iwaki, S. (2021). Putative ratios of facial attractiveness in a deep neural network. *Vision Research, 178*, 86-99.

[19] Grammer, K., & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, *108*(3), 233–242.

[20] Jones, A. L., & Jaeger, B. (2019). Biological bases of beauty revisited: The effect of symmetry, averageness, and sexual dimorphism on female facial attractiveness. *Symmetry, 11*(2), 279.

[21] Kanevsky, J., Corban, J., Gaster, R., Kanevsky, A., Lin, S., & Gilardino, M. (2016). Big data and machine learning in plastic surgery: A new frontier in surgical innovation. *Plastic Reconstructive Surgery, 137*(5), 890e-897e.

[22] Štěpánek, L., Měšťák, J., & Kasal, P. (2019). *Machine-learning at the service of plastic surgery: a case study evaluating facial attractiveness and emotions using R language*. 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), Leipzig, Germany, 2019, pp. 107-112,

[23] Štěpánek, L., Kasal, P., & Měšťák, J. (2019). *Evaluation of Facial Attractiveness after Undergoing Rhinoplasty Using Tree-based and Regression Methods*. 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, pp. 1-4.

[24] Darwin, C. (1871). The Descent of Man, and Selection in Relation to Sex. John Murray.

[25] Symons, D. (1995). Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. In P. R. Abramson & S. D. Pinkerton (Eds.), Sexual nature, sexual culture (pp. 80–119). The University of Chicago Press.

[26] Little A. C., Jones B. C., & DeBruine L. M. (2011). Facial attractiveness: evolutionary based research. *Philosophical transactions of the Royal Society of London*, 366(1571), 1638-59.

[27] Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*(3), 157–175.

[28] Scott, I. M. L., Clark, A. P., Boothroyd, L. G., & Penton-Voak, I. S. (2013). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology 24*(3), 579–589.

[29] Neave, N., Laing, S., Fink, B., & Manning, J. T. (2003). Second to Fourth Digit Ratio, Testosterone and Perceived Male Dominance. *Proceedings: Biological Sciences, 270*(1529), 2167–2172.

[30] Townsend, J. M., & Roberts, L. W. (1993). Gender differences in mate preference among law students: Divergence and convergence of criteria. *The Journal of Psychology: Interdisciplinary and Applied, 127*(5), 507–528.

[31] Geary, D. C., Vigil, J., & Byrd-Craven, J. (2004). Evolution of human mate choice. *The Journal of Sex Research, 41*(1), 27-42.

[32] Pound. N., Penton-Voak, I. S., & Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society of London, B,* 276(1654), 153-9.

[33] Foo, Y. Z, Nakagawa, S., Rhodes, G., & Simmons, L. W. (2017). The effects of sex hormones on immune function: a meta-analysis. *Biological Reviews, 92*(1), 551-571.

[34] Folstad, I., & Karter, A. J. (1992). Parasites, Bright Males, and the Immunocompetence Handicap. *The American Naturalist, 139*(3), 603–622.

[35] Hamilton, W. D., & Zuk, M. (1982). Heritable true fitness and bright birds: A role for parasites? *Science, 218*(4570), 384–387.

[36] Dixson, B. J., & Brooks, R. C. (2013). The role of facial hair in women's perceptions of men's attractiveness, health, masculinity and parenting abilities. *Evolution and Human Behavior, 34*(3), 236-241.

[37] Rhodes G., Chan J., Zebrowitz, L. A., & Simmons, L. W. (2003). Does sexual dimorphism in human faces signal health? *Proceedings of the Royal Society of London, Biological Sciences, 270*(suppl_1), S93–S95.

[38] Little, A. C., Jones, B. C., & Burriss, R. P. (2007). Preferences for masculinity in male bodies change across the menstrual cycle. *Hormones and Behavior, 51*(5), 633-639.

[39] Little, A. C., Cohen, D. L., Jones, B. C., & Belsky, J. (2007). Human preferences for facial masculinity change with relationship type and environmental harshness. *Behavioral Ecology and Sociobiology, 61*, 967–973 (2007).

[40] Rennels, J. L., Bronstad, P. M., & Langlois, J. H. (2008). Are attractive men's faces masculine or feminine? The importance of type of facial stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 884–893.

[41] Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*(5), 154–158.

[42] Foo, Y., Simmons, L., & Rhodes, G. (2017). Predictors of facial attractiveness and health in humans. *Scientific Reports, 7*(39731).

[43] Roberts, S. C., Havlicek, J., Flegr, J., Hruskova, M., Little, A. C., Jones, B. C., … & Petrie, M. (2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(suppl_5).

[44] Bobst, C., & Lobmaier, J. S. (2012). Men's preference for the ovulating female is triggered by subtle face shape differences. *Hormones and Behavior, 62*(4), 413-417.

[45] Itriyeva, K. (2022). The normal menstrual cycle. *Current Problems in Pediatric and Adolescent Health Care*, *52*(5), 101183.

[46] Mihm, M., Gangooly, S., & Muttukrishna, S. (2011). The normal menstrual cycle in women. *Animal Reproduction Science*, *124*(3), 229–236.

[47] Reed, B. G., & Carr, B. R. (2000). The Normal Menstrual Cycle and the Control of Ovulation. In K. R. Feingold, B. Anawalt, M. R. Blackman, A. Boyce, G. Chrousos, E. Corpas, … & D. P. Wilson (Eds.), *Endotext*. MDText.com, Inc.

[48] Fraccaro, P. J., Feinberg, D. R., DeBruine, L. M., Little, A. C., Watkins, C. D., & Jones, B. C. (2010). Correlated male preferences for femininity in female faces and voices. *Evolutionary Psychology, 8*(3), 447–461.

[49] Jones, D., Brace, C. L., Jankowiak, W., Laland, K. N., Musselmann, L. E., Langlois, J. H., & Symons, D. (1995). Sexual selection, physical attractiveness, and facial neoteny: Cross-cultural evidence and implications. *Current Anthropology, 36*(5), 723–748.

[50] Kleisner, K., Tureček, P., Saribay, S. A., Pavlovič, O., Leongómez, J. D., Roberts, S. C., … & Varella, M. A. C. (2024). Distinctiveness and femininity, rather than symmetry and masculinity, affect facial attractiveness across the world. *Evolution and Human Behavior, 45*(1), 82-90.

[51] Moore, F. R., Law Smith, M. J., Taylor, V., & Perrett, D. I. (2011). Sexual dimorphism in the female face is a cue to health and social status but not age. *Personality and Individual Differences, 50*(7), 1068-1073.

[52] Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, 368, 239–242.

[53] Sutić, D., Brešković, I., Huić, R., & Jukić, I. (2010). "Automatic evaluation of facial attractiveness," The 33rd International Convention MIPRO, Opatija, Croatia, 1339-1342.

[54] Zaidel, D. W., & Cohen, J. A. (2005). The face, beauty, and symmetry: Perceiving asymmetry in beautiful faces. *International Journal of Neuroscience, 115*(8), 1165-1173.

[55] Fan, J., Chau, K. P., Wan, X., Zhai, L., & Lau, E. (2012). Prediction of facial attractiveness from facial proportions. *Pattern Recognition, 45*(6), 2326-2334.

[56] Bougourzi, F., Dornaika, F., & Taleb-Ahmed, A. (2022). Deep learning based face beauty prediction via dynamic robust losses and ensemble regression. *Knowledge-Based Systems*, 242.

[57] Gan, J., Xie, X., Zhai, Y., He, G., Mai, C., & Luo, H. (2022). Facial beauty prediction fusing transfer learning and broad learning system. *Soft Computing, 27*, 13391-13404.

[58] Gao, L., Li, W., Huang, Z., Huang, D., & Wang, Y. (2018). *Automatic facial attractiveness prediction by deep multi-task learning*. 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 3592-3597.

[59] Simonyan K., & Zisserman A. (2014). Very deep convolutional networks for large-scale image recognition. *Computer Vision and Pattern Recognition*.

[60] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. *Computer Vision and Pattern Recognition*.

[61] Gan, J., Xiang, L., Zhai, Y., Chaoyun, M., He, G., Zeng, J., … & Piuri, V.-C. (2020). 2M BeautyNet: Facial beauty prediction based on multi-task transfer learning. *IEEE Access, 8*, 20245-20256.

[62] Vahdati, E., & Suen, C. Y. (2019). Female Facial Beauty Analysis Using Transfer Learning and Stacking Ensemble Model. In: Karray, F., Campilho, A., Yu, A. (eds) Image Analysis and Recognition. ICIAR 2019. Lecture Notes in Computer Science, 11663. Springer, Cham.

[63] Zhai, Y., Yu, H., Ying, X., Junying, G., He, C., Wenbo, D., … & Fabio, S. (2020). Asian Female Facial Beauty Prediction Using Deep Neural Networks via Transfer Learning and Multi-Channel Feature Fusion. *IEEE Access, 8*, 56892-56907.

[64] Liang, L., Lin, L., Jin, L., Xie, D., & Li, M. (2018). SCUT-FBP5500: A diverse benchmark dataset for multi-paradigm facial beauty prediction.

[65] Fink, B. Neave, N., Manning, J. T., & Grammer. K. (2006). Facial symmetry and judgements of attractiveness, health and personality. *Personality and Individual Differences, 41*(3), 491-499.

[66] Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Facial symmetry and judgements of apparent health: Support for a “good genes” explanation of the attractiveness–symmetry relationship. *Evolution and Human Behavior, 22*(6), 417-429.

[67] Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior, 20*(5), 295-307.

[68] Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122-1135.

[69] Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces—KDEF (CD ROM). Stockholm: Karolinska Institute, Department of Clinical Neuroscience, Psychology Section.

[70] Ebner, N., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior research Methods, 42*(1), 351-362.

[71] DeBruine, L., & Benedict, J. (2017): Face Research Lab London Set. figshare. Dataset.

[72] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 6*(2), 107-116.

[73] Khan, S., Rahmani, H., Shah, S. A. A., & Bennamoun, M. (2018). CNN Learning. In: A Guide to Convolutional Neural Networks for Computer Vision. Synthesis Lectures on Computer Vision. Springer, Cham.

[74] Garrido, M. V., & Prada, M. (2017). KDEF-PT: Valence, emotional intensity, familiarity and attractiveness ratings of angry, neutral, and happy faces. *Frontiers in Psychology, 8*(2181).

[75] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2016). Inception-v4, Inception-ResNet and the impact of residual connections on learning. *Computer Vision and Pattern Recognition.*

[76] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., … & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *Computer Vision and Pattern Recognition*.

[77] Tan, M., & Le, Q. V. (2021). EfficientNetV2: Smaller models and faster training. *Computer Vision and Pattern Recognition*.

[78] Chollet, F. (2016). Xception: Deep Learning with Depthwise Separable Convolutions. *Computer Vision and Pattern Recognition.*

[79] Chollet, F. (2015). Keras. GitHub repository. https://github.com/fchollet/keras

[80] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *Machine Learning*.

[81] Li, H. Chaudhari, P., Yang, H., Lam, M., Ravichandran, A., Bhotika, R., & Soatto, S. (2020). Rethinking the Hyperparameters for Fine-tuning. *Computer Vision and Pattern Recognition*.

[82] Tiddeman, B., & Perrett, D. (2002). Transformation of dynamic facial image sequences using static 2D prototypes. *The Visual Computer, 18*(4), 218-225.

[83] DeBruine, L. (2018). debruine/webmorph: Beta release 2 (v0.0.0.9001). Zenodo.

[84] DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., … & Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1355–1360.

[85] Penton-Voak, I. S., Little, A. C., Jones, B. C., Burt, D. M., Tiddeman, B. P., & Perrett, D. I. (2003). Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*). *Journal of Comparative Psychology, 117*(3), 264–271.

[86] Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE computer graphics and applications*, *15*(5), 70-76.

[87] Gangestad, S. W., & Thornhill, R. (2003). Facial masculinity and fluctuating asymmetry. *Evolution and Human Behavior, 24*(4), 231–241.

[88] Pinheiro, J., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.

[89] R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[90] Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

[91] Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

[92] Fernandes, N. F. R. (2024). Analyzing VGG-19’s Bias in Facial Beauty Prediction: Preference for Feminine Features. https://doi.org/10.17605/OSF.IO/RW4D3

[93] Marcinkowska, U. M, Kozlov, M. V., Cai, H., Contreras-Garduño, J., Dixson, B. J., Oana, G. A., … & Rantala, M. J. (2014). Cross-cultural variation in men's preference for sexual dimorphism in women's faces. *Biology Letters, 10*(4), 20130850.

[94] O'Connor, J. J., Fraccaro, P. J., Pisanski, K., Tigue, C. C., & Feinberg, D. R. (2013). Men's preferences for women's femininity in dynamic cross-modal stimuli. *PLoS One, 8*(7), e69531.

[95] Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour, 69*(3), 561-568.

[96] Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: evidence for hormone-mediated adaptive design. *Evolution and Human Behavior, 22*(4), 251-267.

[97] Little, A. C., Jones, B. C., Waitt, C., Tiddeman, B. P., Feinberg, D. R., Perrett, D. I., … & Reimchen, T. (2008). Symmetry is related to sexual dimorphism in faces: data across culture and species. *PLoS ONE 3*(5), e2106.

[98] Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature, 399*, 741–742.

[99] Penton-Voak, I. S., & Perrett, D. I. (2001). Male facial attractiveness: Perceived personality and shifting female preferences for male traits across the menstrual cycle. *Advances in the Study of Behavior, 30,* 219 – 260.

[100] Stower, R. E., Lee, A. J., McIntosh, T. L., Sidari, M. J., Sherlock, J. M., & Dixson, B. J. (2020). Mating strategies and the masculinity paradox: how relationship context, relationship status, and sociosexuality shape women's preferences for facial masculinity and beardedness. *Archives of Sexual Behavior,* 49, 809–820.

[101] DeBruine, L. M., Jones, B. C., Smith, F. G. & Little, A. C. (2010). Are attractive men’s faces masculine or feminine? The importance of controlling confounds in face stimuli. *Journal of Experimental Psychology. Human Perception and Performance, 36*(3), 751–758.

[102] Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society of London, Series B, 268*(1462), 39–44.

[103] Marcinkowska, U. M., Rantala, M. J., Lee, A. J.. Kozlov, M. V., Aavik, T., Cai, H., ... & Dixson, B. J. W. (2019). Women’s preferences

for men's facial masculinity are strongest under favorable ecological conditions. *Scientific Reports,* 9(3387).

[104] Nakamura, K., & Watanabe, K. (2020). A new data-driven mathematical model dissociates attractiveness from sexual dimorphism of human faces. *Scientific Reports, 10*(16588).

[105] Cornwell, R. E., Boothroyd, L., Burt, D. M., Feinberg, D. R., Jones, B. C., Little, A. C., … & Perrett, D. I. (2004). Concordant preferences for opposite-sex signals? Human pheromones and facial characteristics. *Proceedings of the Royal Society of London, Series B, 271*, 635–640.

[106] Swaddle, J. P., & Riersen, G. W. (2002). Testosterone increases perceived dominance but not attractiveness in human males. *Proceedings of the Royal Society of London, Series B, 269*(1507), 2285–2289.

[107] Hu, Y., Abbasi, N. H., Zhang, Y., & Chen, H. (2018). The effect of target sex, sexual dimorphism, and facial attractiveness on perceptions of target attractiveness and trustworthiness. *Frontiers in Psychology, 9*.

[108] Welling, L. L. M., Jones, B. C., DeBruine, L. M., Little, A. C., & Smith, F. G. (2008). Exposure to sexually attractive men decreases women's preferences for feminine faces. *Journal of Evolutionary Psychology, 6*(3), 219–230.

[109] Little, A. C., Burriss, R. P., Jones, B. C., DeBruine, L. M., & Caldwell, C. A. (2008). Social influence in human face preference: men and women are influenced more for long-term than short-term attractiveness decisions. *Evolution and Human Behavior, 29*(2), 140-146,

[110] Bartlome, R. I., & Lee, A. J. (2023). Facial attractiveness, but not facial masculinity, is used as a cue to paternal involvement in fathers. *Adaptive Human Behavior and Physiology,* 9, 182–197.

[111] Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, *37*(1), 78–84.

[112] Fink, B., & Neave, N. (2005). The biology of facial beauty. *International Journal Cosmetic Science, 27*(6), *317–325.*

[113] Schölkopf, B., Platt, J., & Hofmann, T. (2007). *A Humanlike predictor of facial attractiveness*. Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference, MIT Press, 649-656.

[114] Tybur, J. M., Fan, L., Jones, B. C., Holzleitner, I. J., Lee, A. J., & DeBruine, L. M. (2022). Re-evaluating the relationship between pathogen avoidance and preferences for facial symmetry and sexual dimorphism: A registered report. *Evolution and Human Behavior, 43*(3), 212-223.

[115] Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry, and cues to good genes. *Proceedings of the Royal Society of London, B, 266,* 1913–1917.

[116] Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior, 25*(4), 229–241.

[117] Mogilski, J. K., & Welling, L. L. M. (2017). The relative importance of sexual dimorphism, fluctuating asymmetry, and color cues to health during evaluation of potential partners' facial photographs. *Human Nature, 28*, 53–75.

[118] Swaddle, J. P., & Cuthill, I. C. (1995). Asymmetry and human facial attractiveness: symmetry may not always be beautiful. *Proceedings of the Royal Society B, 261*, 111–116.

[119] Zheng, R., Ren, D., Xie, C., Pan, J., & Zhou, G. (2021). Normality mediates the effect of symmetry on facial attractiveness. *Acta Psychologica, 217.*

[120] Palumbo, R., Adams, R. B., Hess, U., Kleck, R. E., & Zebrowitz, L. (2017). Age and Gender Differences in Facial Attractiveness, but Not Emotion Resemblance, Contribute to Age and Gender Stereotypes. *Frontiers in Psychology, 8*.

[121] McLellan, B., & McKelvie, S. J. (1993). Effects of age and gender on perceived facial attractiveness. *Canadian Journal of Behavioural Science, 25*(1), 135–142.

[122] Levy, B., Ariely D., Mazar, N., Chi, W., Lukas, S., & Elman, I. (2008). Gender differences in the motivational processing of facial beauty. *Learning and Motivation, 39*(2), 136–145.

[123] Nistor, S. C., Marina, A. -C., Darabant, A. S., & Borza, D. (2017). *Automatic gender recognition for "in the wild" facial images using convolutional neural networks*. 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2017, pp. 287-291.

[124] Nga, C. H., Nguyen, K. -T., Tran, N. C., & Wang, J. -C. (2020). *Transfer learning for gender and age prediction*. 2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan), Taoyuan, Taiwan, pp. 1-2.

[125] Żejmo, A., Gielert, M., Grabski, M., & Kostek, B. (2023). *Assessing the attractiveness of human face based on machine learning*. *Procedia Computer Science, 225*, 1019-1027.

[126] Golle, J., Mast, F. W., & Lobmaier, J. S. (2014). Something to smile about: The interrelationship between attractiveness and emotional expression. *Cognition and Emotion, 28*(2), 298-310.

[127] Lindeberg, S., Craig, B. M., & Lipp, O. V. (2019). You look pretty happy: Attractiveness moderates emotion perception. *Emotion, 19*(6), 1070–1080.

[128] Balasubramanian, B., Diwan, P., Nadar, R., & Bhatia, A. (2019). *Analysis of Facial Emotion Recognition*. *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 945-949.

[129] Lim, H.-E. (2008). The use of different happiness rating scales: Bias and comparison problem?. *Social Indicators Research, 87*(2), 259–267.

[130] Allen, I. E., & Seaman, C. A. (2007). Likert Scales and Data Analyses. *Quality Progress, 40*(7), 64-65.

[131] Simpson, J. A., Gangestad, S. W., & Lerma, M. (1990). Perception of physical attractiveness: Mechanisms involved in the maintenance of romantic relationships. *Journal of Personality and Social Psychology, 59*(6), 1192–1201. https://doi.org/10.1037/0022-3514.59.6.1192

[132] Kandrik, M., & DeBruine, L. M. (2012). Self-rated attractiveness predicts preferences for opposite-sex faces, while self-rated sex-typicality predicts preferences for same-sex faces. *Journal of Evolutionary Psychology, 10*(4), 177–186.

[133] Glassenberg, A. N., Feinberg, D. R., Jones, B. C., Little A. C., & Debruine L. M. (2010). Sex-dimorphic face shape preference in heterosexual and homosexual men and women. *Archives of Sexual Behavior, 39*(6), 1289-1296.

[134] Zhang, J., Zheng, L., & Zheng, Y. (2018). Consistency in preferences for masculinity in faces, bodies, voices, and personality characteristics among homosexual men in China. *Personality and Individual Differences, 134*, 137-142.

[135] Zheng, L. (2021). The dyadic effects of top/bottom sexual self-labels and partner sexual role requirements on facial masculinity preferences among gay and bisexual men in china. *Journal of Sex Research, 58*(1), 122-12.