Insights into Automated Attractiveness Evaluation from 2D Facial Images: A Comprehensive Review

Ali Ibrahem
Department of Information Technology
Akre University for Applied Science DPU, Iraq
ali.hikmat@auas.edu.krd

Jwan Saeed
Department of Information Technology
Duhok Polytechnic University, Iraq
jwan.najeeb@dpu.edu.krd

Adnan Abdulazeez
Technical College of Engineering,
Duhok Polytechnic University, Iraq
adnan.mohsin@dpu.edu.krd

Abstract: *Predicting facial beauty in computer vision is a relatively emerging research area with diverse applications. However, Facial Beauty Prediction (FBP) has various challenging possibilities due to the lack of a universally accepted assessment procedure for facial beauty, the scarcity of sufficient available databases and computational models, and the way of extracting discriminative features to quantify facial attractiveness. This paper comprehensively reviews the 2D facial image beauty analysis and prediction research that utilizes computation, machine, and deep learning techniques. It introduces the limitations and makes a critical analysis of this research area. Beauty hypotheses, feature extraction, evaluation methods, and FBP benchmark datasets that can help measure the effectiveness of the automatic attractiveness assessment approaches are discussed and analyzed. In addition, this paper tries to figure out the answer to the most debatable question that says, "which face organs contribute to facial beauty and to what extent?". It also highlights concerns and challenges in the FBP domain that can provide a foundation for future work and further development in automatic facial beauty estimation and evaluation.*

Keywords: *Facial beauty prediction, facial attractiveness assessment, beauty hypotheses, beauty computation model.*

Received July 5, 2024; accepted October 10, 2024
https://doi.org/10.34028/iajit/22/1/7

1. Introduction

The impact of facial appearance on human life is significant, as evidenced by a plethora of studies revealing that individuals deemed attractive enjoy greater prospects of securing esteemed employment opportunities, attaining personal satisfaction, and reaping additional social benefits in their daily interactions [43, 106]. Researchers from various professions have contributed to the study of face attractiveness, including psychology [91], evolutionary biology [27], plastic surgery [105], and computer science. However, the ultimate qualitative and quantitative idea of facial beauty remains elusive [123]. Research has discovered that people of all ethnicities, socioeconomic classes, ages, and genders agree on face beauty perception [59]. Based on this conclusion, computer scientists have developed landmark-based and data-driven techniques to analyse facial attractiveness. Consequently, the concept of facial image attractiveness appears to be a universal phenomenon that machines can learn [2, 49]. The primary objective of Facial Beauty Perception (FBP) is to automatically assess facial attractiveness in accordance with human perception. This analysis system encompasses six key phases: image acquisition for dataset creation coupled with image rating, pre-processing, feature extraction and selection, facial beauty modelling, validation of the model, and ultimately, the development of the application. Despite the growing interest in FBP research, progress in this domain has been sluggish due to the lack of ample and dedicated datasets for FBP and the limited scale of FBP investigations [29]. In 1990, Langlois and Roggman [59] conducted a pioneering computer-aided photographic systems beauty evaluation study to assess the objective association between mathematical averageness and facial appeal. In 2001 Aarabi *et al*. [1] introduced the first autonomous scoring system for categorizing facial attractiveness in 2001. Prior methodologies mainly relied on heuristic rules, which rely on the physical quantification of facial beauty through the spatial interrelations between various facial features, and treat FBP as a machine learning classification, regression, or ranking problem. Both regression and classification models are crucial in predictive analytics, particularly in machine learning and Artificial Intelligence (AI). While ranking can be efficient with small datasets [73].

In most investigations, FBP is conventionally treated as a fully supervised task. In early attempts to evaluate facial attractiveness, rudimentary machine-learning techniques, and hand-crafted feature engineering were utilized. The latter approach typically relied on geometric ratios and inter-landmark distances as input features. However, such methods exhibited suboptimal predictive performance and necessitated considerable human effort in labelling facial landmarks for each image within a given dataset. Nonetheless, notable strides have since been achieved in automated facial attractiveness evaluation, largely stemming from the advent of deep learning techniques. Machine-based assessing facial beauty like a human predictor, is an emerging and challenging topic. It has relatively limited resources in general, and there is a scarcity of providing a broad review on this search area compared to other face-related research such as face recognition, facial expression, emotion recognition, etc. Gunes [39] has investigated various aspects of beauty traits with their perceptions and computation. Meanwhile, Laurentini and Bottino [61] provided human sciences and medicine findings supporting computer beauty analysis. They surveyed many studies on face attractiveness analysis. However, these studies were only based on hand-crafted features and traditional machine learning. Liu *et al*. [77] reviewed the progress in facial attractiveness computation based on the adversity of aspects. More recently, Saeed and Abdulazeez [96] expounded upon the latest developments in Facial Beauty Prediction (FBP) via Deep Convolutional Neural Networks (DCNNs) in their investigation of the FBP models.

This paper reviews 2D facial image beauty analysis and prediction studies, beauty hypotheses, methodologies, FBP benchmarks, and challenges. It also attempts to answer the most controversial question, “which face organs contribute to facial beauty and to what extent?” To summarize the main contributions of this paper, they can be listed as follows:

- It delves into the significance of beauty notions and hypotheses in computer vision and shows valuable insights into the intersection of computer science and aesthetics.
- It outlines the main stages of constructing facial beauty-related task models and presents the analysis of commonly used face datasets in FBP research.
- Addressing the challenging question of how different facial organs contribute to facial beauty assessment.
- Identifying open issues and challenges in the computation of facial attractiveness via computer vision.

The remainder of the paper is structured as follows: Section 2 illustrates the beauty hypotheses and the tangible applications of FBP. Section 3 presents a computer-based facial beauty and attractiveness analysis system with six fundamental stages. The importance of facial components in beauty assessment is discussed in section 4. Challenges and open issues of facial beauty assessment in computer vision are outlined in section 5. Finally, section 6 presents the conclusion.

2. Facial Attractiveness Theories and Practical Implications

What exactly is beauty? For ages, psychologists, artists, and philosophers have grappled with this subject. The well-known adage "Beauty is in the eye of the beholder" implies that individual beauty is subjective and unexpected according to our understanding of a person's background and culture. However, it has been proven in several studies that there is a significant cross-cultural agreement in the beauty judgment of the face [18]. People worldwide use comparable criteria to judge attractiveness, and there are specific common characteristics concerning attractiveness; in other words, the perception of facial beauty is not determined by specific people but is a global norm. Several ideal aspects of beauty have been proposed throughout history, primarily by developing canons of face shapes and distances between selected facial landmarks in meaningful and salient locations. Over the past decades, several research investigations have mainly focused on primary characteristics of face beauty, such as the golden ratio, neoclassical canons, averageness, symmetry, and the existence of sexual dimorphism. With the advancements in pattern analysis and machine learning technologies, face attractiveness study is no longer restricted to subjective psychological cognition; face beauty can be predicted and enhanced using machine learning approaches. The most familiar beauty hypotheses are listed in the next subsections.

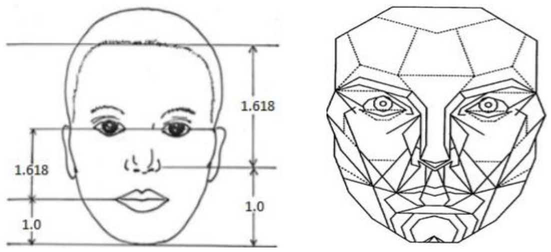
2.1. Golden Ratio

The Greeks were the first to know about the golden ratio, also known as Phi or divine proportion, and they used it to create their most famous piece of art. For instance, the Golden Ratio is present in both the Mona Lisa and the Last Supper paintings by Leonardo da Vinci [107]. In the realm of mathematics, two quantities are said to be in the golden ratio when the ratio of these quantities is identical to the ratio of their sum to the larger of the two [81]. Algebraically expressed as:

$$\frac{a}{b} = \frac{a+b}{b} = \varphi \tag{1}$$

For quantities *a* and *b* where $a>b>0$:

It is an irrational number, nearly equal to 1.618 [122]. It is said that attractive faces have facial proportions close to the golden ratio [24]. The concept of a universal standard of beauty based on the golden ratio is shown in Figure 1-a). Consequently, Marquardt devised a Phi face mask that encapsulated the perfect proportions based on the golden ratio. Marquardt presented an 'ideal' face template. A notion that drew some favour from the plastic surgery profession [92] (see Figure 1-b)). Pallett *et al*. [85] presented a new golden ratio that considers the face attractive when the eye–to–mouth distance on the face is 36% of the face length, and the interocular distance is 46% of the face width. Recently, Tong *et al*. [108] and Liang *et al*. [71] utilized these new golden rations to show that when no explicitly labeled facial features for beauty are provided, a Deep Neural Network (DNN) model can learn potential ratios from face images using just category annotation. However, there has been much dispute over this assumption and its aesthetic characteristics. The usage of this proportion in art and architecture in general and in FBP has been praised [24, 38, 40] and criticized. It has been argued that the aesthetic concept of beauty is elusive and unlikely to be reduced to a few basic ratios [87].



a) Face Phi ratio. b) Marquardt mask.

Figure 1. Facial proportion of face Phi ratio and Marquardt mask [92].

2.2. Neoclassical Canons

Since the Renaissance era, neoclassical canons have been posited by painters as guiding principles for the representation of an aesthetically pleasing face. The central premise of these canons revolves around the notion that attractive facial features should conform to predetermined and specific ratios. A compendium of these concepts is comprehensively delineated as nine neoclassical canons in [26, 108]. Figure 2 illustrates samples of the anthropometric landmarks employed for measuring neoclassical canons. The incorporation of Neoclassical canons and other related guidelines have extended beyond artistic pursuits, and may finding application in certain plastic face surgery methods, including rhinoplasty [84, 101], and can contributing to the advancement of the art.

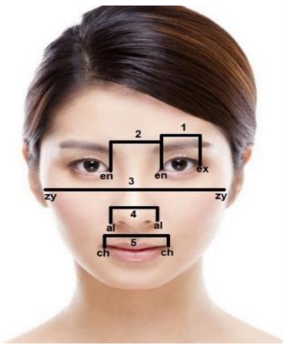
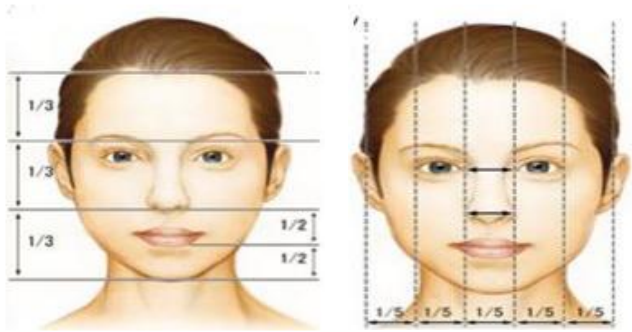


Figure 2. Anthropometric landmarks; al: alare, ch: cheilion, [52].

However, Bozkir *et al*. [10] revealed that the applicability of neoclassical canons was found to be limited across diverse populations. Consequently, their utility in surgical planning is constrained due to the observed variations in facial measurements among different racial and national groups. Furthermore, recent research adopts a more nuanced approach to the study of facial beauty, recognizing the multifaceted nature of attractiveness perceptions [37, 132].

2.3. Facial Vertical Thirds and Horizontal Fifths

This hypothesis seeks to evaluate face height and width. For face height and according to this notion, a harmonious face can be segmented into roughly equal thirds by making horizontal lines throughout the hairline on the forehead, the brows, the base of the nose, and the chin border, as shown in Figure 3-a). Furthermore, the distance between the lips and the chin should be double that between the base of the nose and the lips. While the width of the face is determined by dividing it into equal fifths, as illustrated in Figure 3-b). The width of a fifth of the entire facial width and the intercanthal distance or nasal base width are considered aesthetically pleasing faces [101].



a) Horizontal thirds. b) Vertical fifths.

Figure 3. Facial horizontal thirds and vertical fifths [128].

2.4. Averageness

The average face describes how a face is like the other faces in a population. In studies of physical attractiveness, averageness refers to the concept of beauty derived from blending the facial characteristics of individuals of the same gender and roughly similar age [75] as shown in Figure 4. The popular belief in the early and mid-twentieth centuries was that averageness in average (composite) faces was the most important component in determining the attractiveness of a face. This assertion has lately been supported by digital technologies, which show that the most beautiful facial images were formed by combining several diverse faces. Studies showed that the more traits that were near to average, the more beautiful they are, and a face shape that is extremely away from the average should be less attractive [41]. Meanwhile, the extremely attractive face is not even close to this average [4, 59]. To investigate this hypothesis, Zhang *et al*. [129] quantified the influence of geometrics, and they showed that the face beauty is increased when it is close to the ravaged face, but if the face already has a relatively small distance to average shape, then deforming it to average shape will not significantly effect on the face beauty enhancement. The Weighted Average (WA) hypothesis is proposed in [13], which says that the WA of two female face geometric features is more beautiful than the less attractive one between them, and it is used in face beautification as well. However, WA hypnosis has only been validated for female faces, ignoring color and texture features. The question of quantifying and standardizing these qualities regarding facial beauty remains unsolved. Thus, although considerable computer-based research has been done to test the averageness hypothesis in facial beauty analysis, estimation, and assessment, it is still widely debated.

Figure 4. The general concept of face averageness [47].

2.5. Symmetry

Face symmetry is considered an essential aspect of beauty, and it is accepted that human faces display large quantities of directional asymmetry and anti-symmetry in their skeletal and soft tissue components [102]. Its direct analogy to health was evolutionary advantageous [11]. Moreover, symmetry investigations can significantly aid in planning operations and evaluating various surgical methods in surgery [8]. It is believed that the more symmetrical a person's face is, the more attractive they are perceived to be. Both left-left and right-right symmetrical faces are used in research and aesthetics to examine the influence of facial symmetry on beauty perceptions and to explore how small differences in facial symmetry can impact the perception of an individual's appearance as shown in Figure 5. However, there are conflicting views about the importance of symmetry in the perception of FBP. Some researchers advocate the positive impact of symmetry in facial beauty assessments [95, 111]. In contrast, others proved that symmetry has been less effective than other proportional characteristics [44].

a) Original face. b) Left symmetry c) Right symmetry.

Figure 5. Facial morphing of original face, left symmetry and right symmetry [66].

2.6. Sexual Dimorphism

Sexual dimorphism refers to the differences in physical characteristics between males and females of the same species. In the context of FBP, sexual dimorphism can play a role in how attractiveness is perceived. Research has shown that people tend to find facial features that are associated with masculinity (in males) or femininity (in females) to be attractive. For example, men with strong, chiselled jawlines and prominent brow ridges may be considered more attractive, while women with softer, more rounded facial features may be considered more attractive [64, 111]. Figure 6 shows the depictions of facial images demonstrating varying degrees of attractiveness, with the top row depicting more attractive faces and the bottom row depicting less attractive ones. The images on the left side of the left group portray masculine male faces, while those on the right side of the left group display feminine male faces. In contrast, the left side of the right group features feminine female faces, and the right side of the right group showcases masculine female faces.

Figure 6. Sexual dimorphism’s influence on facial attractiveness samples [45].

2.7. Facial Beauty Prediction Aapplications

Facial attractiveness has been involved in various intriguing applications as follows:

- It provides insight into psychology to explain human behavior and evolutionary biology, revealing evolutionary directions and the associated biological benefits [122]. Several studies have found that face attractiveness influences behavioral responses as a highly visible social cue. It has also been discovered that beautiful faces evoke different brain activation than unattractive or neutral looks [56]. Many websites, or attractiveness ranking tools, are already available. Choosing the finest photos for family albums, social networks, and online dating are some examples of this. Millions of people use online dating services daily to make one of the most significant decisions of their lives: finding a partner [9, 11].
- Face enhancement and beautification is another wide range application of FBP. Thus, Virtual face beautification, like cosmetic surgery in the physical world, is a developing field that involves automated retouching and unblemished photographs for advertising, magazine covers, motion movies, and special effects, as well as screening applicants for specialized occupations such as entertainment and modeling, where beauty is needed face enhancement and beatification, can involve facial skin beautification, average facial beautification, facial shape beautification [68], and beautification based on style or makeup transfer [78]. Furthermore, superior results could be gained when hybridizing these types of face improvement and beautification. Since face cosmetics and hairstyles are now required by most

modern women, makeup and hairstyle recommendation models are constructed based on the notion of FBP. Facial aesthetic assessment could be used as an automated evaluation tool, the cosmetic industry tool for analyzing the efficacy of cosmetic assistance. The techniques of simulating the hairstyle and physical makeup of these items and synthesizing their effects have inspired research studies [83].

- FBP contributes to the medical industry as well. Medical imaging and computer-assisted surgical planning are essential components of the preoperative workup because examining alternative operating techniques virtually can minimize operation time and expense. Furthermore, it facilitates more consistent and optimized surgery [54]. Beauty analysis systems of plastic, reconstructive surgery, and orthodontics capable of evaluating various simulations or even recommending how to improve the attractiveness of a particular patient would considerably improve planning effectiveness. These computer vision-based systems can enhance clinical outcomes, save operation time, and save money. Therefore, a considerable number of publications on face beauty and the influence of the face components on facial attractiveness have been published in cosmetic surgery [28, 105] and Orthodontics [55, 89], as well as utilizing a variety of FBP models that help in assessing facial beauty before and after the surgery.

3. Computer-Based Analysis System for Facial Image Beauty and Attractiveness

Facial beauty analysis is a relatively new topic in multimedia and biometrics [130]. With the increasing use of digital cameras, images invade all areas of everyday life, contributing to significant growth in objectively and accurately judging facial attractiveness analysis and applying it in a diverse related research area [88]. The facial image beauty analysis system consists of six main stages, beginning with image acquisition for database construction, then image scoring (labeling), pre-processing, feature extraction and selection, facial beauty modeling, validating this model and ultimately developing applications that predict the score of beauty. These six fundamental steps are illustrated in Figure 7 and discussed in the next subsections.

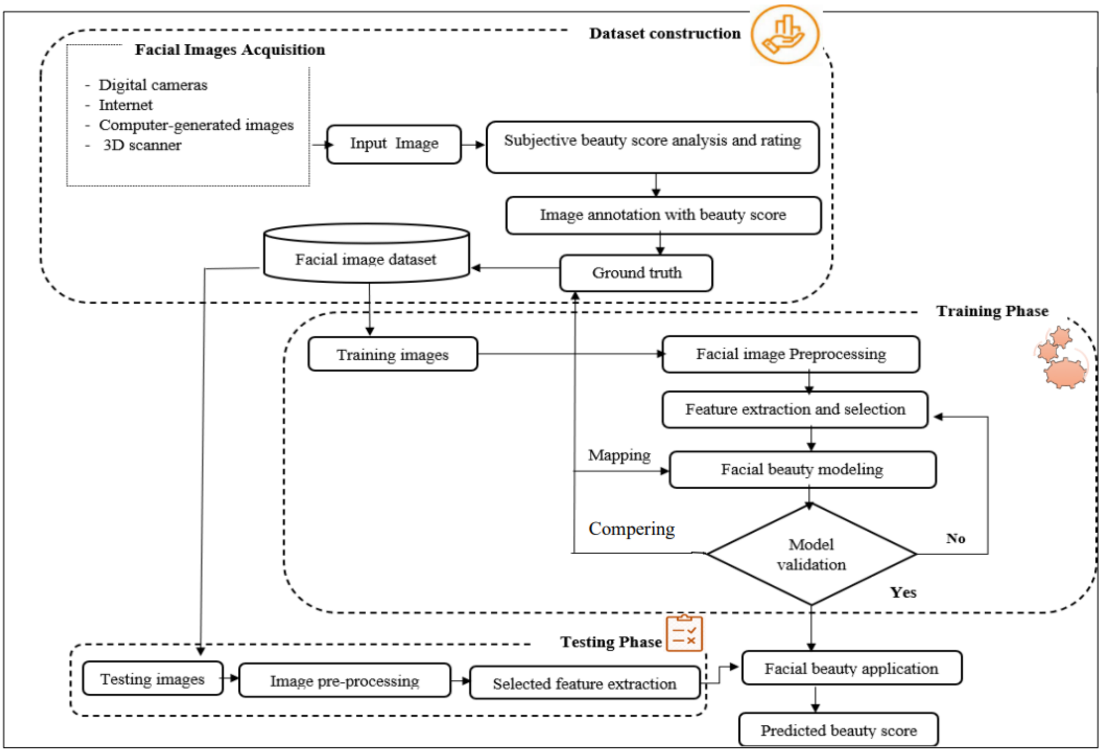


Figure 7. The fundamental general steps of facial images beauty prediction.

3.1. Facial Beauty Dataset and Image Rating

Although studies have found that facial beauty is a universal notion that can be learned via machine and deep learning, establishing a beauty standard remains challenging due to differing individual judgments of facial beauty. Thus, FBP has less public authoritative data. Therefore, constructing a database on a wide scale is challenging [29]. A considerable amount of diversity in attractiveness levels is necessary for the facial image beauty research dataset. Therefore, choosing which face database to use is crucial for the model's performance. The facial images could be gained from the Internet, digital cameras or 3D scanner, facial image databases, and computer-generated images. Moreover, these images need to be rated in various ways. Then, images are labeled with their beauty score to create the ground truth for the sack of the learning process and model validity.

3.1.1. Facial Images Databases

Typically, the database should involve sufficient facial images of all beauty levels and genders, ages, races, poses, and expressions. Furthermore, the dataset should include both extremes of facial attractiveness: highly attractive and very unattractive faces, to have a good representation of the concept of beauty [82]. However, due to privacy concerns, the majority of datasets are not publicly accessible. Also, landmarks are critical for evaluating the geometry cue, which is lacking in some benchmarks [1]. The majority of existing datasets are relatively small [23, 38, 82] and have neutral (postures, expression, attractiveness), insufficiently diverse, and most of them are limited to female facial beauty [5, 94, 96].

Subject's gender is another worth mentioning aspect since most facial datasets are dedicated to females. This might justify that beauty in females is simpler to analyze and compute than in the face of a male. Accordingly, many researchers employ images of either females or both genders. In contrast, using male images only, as in [3] for FBP is still rare.

As far as we know, no currently available face databases meet all conditions above. There are just a few attractive faces in current databases, and they tend to be from a small number of ethnic groups because beautiful faces are relatively rare among the population. Existing public face databases built for other face-related tasks, such as facial expression recognition datasets, cannot be used directly in facial beauty research. Evaluations based on these datasets may introduce significant differences [70] because such datasets may not meet FBP dataset conditions. For instance, it does not consider the diversity of attractiveness levels, especially for both extremely (attractive and unattractive). However, due to the scarcity of such dedicated public face datasets, especially in the early stages, authors utilized face recognition datasets in [103, 80, 33], and facial expression datasets in [121] for FBP purposes. Face posture and expression are two essential aspects that may restrict the dataset and model performance when they have extreme variation. That is why most existing FBP datasets have a frontal face and neutral expression.

3.1.2. The Existing FBP Datasets

In spite of the fact that facial images datasets dedicated to FBP and its related tasks are scarce, some well-known face datasets are utilized mainly for this purpose as a benchmark (see Table 1). Gray *et al*. [37] performed the first attempt to tackle the issue of facial feature landmark localization and imposing stringent limits on training samples by building a considerable FBP database of 2056 females known as HotorNot. It is a challenging dataset because the images are under unconstrained conditions on the background, expression, position, lighting, race and age without the use of landmarks, as well as the problem of adequately predicting the locations of landmark features. The same notion was utilized for the construction of a benchmark called Large-Scale Asian Facial Beauty Database (LSAFBD) [126], which contained 20,000 labeled images of gender and 80,000 unlabelled ones. However, both of these datasets are based on apparent features only and not publicly accessible. Nguyen *et al*. [84] built Multi-Modality Beauty (M2B), including the attractiveness of dressing and voice besides female facial image beauty. Chen *et al*. [13] developed a database called beauty 799 with 390 famous female beautiful faces and 409 common faces to prove their proposed WA of facial attractiveness.

The most top common and available FBP databases are SCUT-FBP [115] and SCUT-FBP5500 [69], as they applied in many facial beauty studies with good results to some extent in most of them. However, SCUT-FBP contains only 500 Asian female faces, limiting the performance of the data needed model for FBP. On the other hand, the South China University of Technology-Facial Beauty Prediction-FBP5500 (SCUT-FBP5500) dataset is relatively large in size and has been utilized in a variety of recent publications. However, it has limitations in terms of light, blurriness, and position may affect the attractiveness prediction model's effectiveness [64]. Moreover, Tong *et al*. [108] found that the distribution of the SCUT-FBP5500 dataset's beauty scores is imbalanced, which may influence the correlation analysis. Another public dataset known as CelebA [79] is a massive face database containing over 200,000 famous images with 40 attribute annotations; attractiveness is one of them. This dataset contains images with a wide range of poses and backgrounds that form a kind of challenge. Furthermore, it is employed basically in a variety of computer vision applications such as face recognition, face detection, facial component localization, face editing and synthesis are all computer vision tasks. CelebA dataset was frequently used in facial beauty classification as it has only two face beauty scores (attractive and unattractive) and this might be easier in computation terms (process) but may not be fair in assessing beauty.

This is because the beauty score range has a significant impact on the beauty evaluation fairness to be more realistic, conducting a robust ground truth which is the most important element in both learning and model validation processes.

A multi-ethnic of 2550 male and female facial images in-the-wild dataset called MEBeauty is provided by Saeed and Abdulazeez [96] and it is expected to be one of the most used datasets for the future research of facial aesthetic assessment. However, the personalized aspect of FBP can be affected due to the relatively large number of raters. Lately, Vijayarajan *et al*. [110] introduced a dataset of beauty pageant photos called Celebrity consists of three beauty classes, with 450 images per class. The classes were ordered based on beauty, with the first class representing the most beautiful, followed by the second class with a higher level of beauty, and the third-class denoting beauty.

Most FBP performances are achieved on relatively small facial datasets using traditional machine or shallow network learning approaches [125]. Therefore, when the dataset scale is too small, it frequently leads to overfitting during the model training process, making the prediction model unsuitable for effective prediction. Moreover, datasets built under specific computation constraints would restrict the computational model's performance and flexibility, and it is not easy to compare different models produced from the dataset using different computation methodologies [69]. All existing face datasets are regarded as static features. To target the dynamic FBP, Weng *et al*. [113] built the first Video-based Facial Attractiveness Prediction (VFAP) dataset that includes 1,430 short video clips of facial performance from TikTok. However, some variations could be found in the beauty scores due to using the number of likes, comments, and forwards to indicate attractiveness levels. Moreover, the gender distribution of TikTok's facial performance videos is skewed. As a result, the beauty rankings of male faces are not always as easily explained as those of female ones.

The majority of current research on face attractiveness computation focused on Two-Dimensional (2D) image faces, particularly the 2D frontal aspect of faces. Meanwhile, some early studies, as in [24], have used computer-generated faces that are digitally blended rather than real faces. However, evaluating the attractiveness of real human faces requires considering privacy and ethical issues. The concern with computer-generated facial images is that the expression of hair and facial texture is significantly different than in real facial images. Recent advances in deep learning-based image creation models effectively solve the abovementioned issue. The Generative Adversarial Network (GAN) framework for machine learning [53] is a generator architecture that excels in image creation. Local and global features, such as human faces, are significant and can be used to construct wholly fabricated but realistic and convincing face images [57]. Information such as the height of the cheekbones and nose may help to determine facial beauty. Such information is difficult to obtain with a 2D frontal facial shot. However, relying solely on a 2D frontal face view will omit much information about attractiveness [61]. While using landmarks to characterize forms is fairly frequent, there is no agreement on the exact landmarks that should be utilized for facial attractiveness analysis. Because the frontal and profile views emphasize different aspects of the face, combining them can result in more accurate facial beauty measurements, reinforcing the theory of 2.5D facial beauty [76]. The 3D representation is ultimately the only approach to gaining a complete and accurate representation of face beauty. Despite the increasing use of 3D imaging [42, 112], it remains relatively underutilized in the field of FBP.

Table 1. The description of the FBP dataset benchmarks.

Dataset	Ref.	Year	Size	F/M	Age	LM	Race	Pose	Expression	Score	Raters
HotorNot	[37]	2010	2056	F	18-40	-	Diverse	Almost frontal	Different	10	30
M2B	[84]	2012	1240	F	19-40	-	Western /Eastern	Different	Different	10	40
Beauty 799	[13]	2014	799	F	N/A	98	Diverse	Frontal	Almost Neutral	3	25
SCUT-FBP	[115]	2015	500	F	N/A	84	Asian	Frontal	Neutral	5	70
Celeb A	[79]	2015	200K	F/M	Different	5	Diverse	Different	Different	2	40
LSAFBD	[126]	2016	20000	F/M	Different	5	Asian	Almost frontal	Different	5	200
SCUT-FBP5500	[69]	2018	5500	F/M	15-60	86	Asian /Caucasian	Frontal	Neutral	5	60
MEBeauty	[63]	2021	2550	F/M	Different	-	Diverse	Different	Different	10	300
Celebrity	[110]	2023	1350	F	N/A	-	Diverse	Different	Different	3	N/A

3.1.3. Human Raters and Attractiveness Scores

While attractiveness is subjective, the ratings given to facial images by various people tend to be related. This allowed for constructing datasets containing facial images and human-labeled beauty scores. Identifying a ground truth to assess measures of faces is the most challenging stage in judging facial attractiveness [76, 109], and it has a considerable effect on the model's accuracy and robustness, especially in model training and validation.

Commonly the beauty scores in FBP studies are divided based on beauty level assessment which is varied in different attractiveness scales, usually ranging between 2-10 scores to allow the raters to show how much they agree or disagree with facial image beauty level. Human participants were asked to judge the attractiveness of the faces in the database, and the scores were produced. Therefore, asking a diverse set of human referees to evaluate a series of facial images in terms of their facial beauty produced unimodal and compact grade histograms, supporting the idea that perception of beauty is universal to some extent [38]. Image can be rated in several ways. The vast majority of studies have used some form of absolute rating, typically a Likert scale in which the person is required to rate their level of agreement with a statement. This form of evaluation requires a considerable number of people to evaluate each image to produce a dispersion of scores that can be averaged to estimate the actual score. However, this approach which uses the average rating is not ideal because each user will have a different background in image rating, and a user's assessment of one image may be influenced by the rating of the prior image, among other aspects.

Additionally, when it comes to controversial faces, the average score is not necessarily a fair reflection of universally agreed preference [46]. Fan *et al*. [25] proposed a different way of image rating. They recast the computation of face attractiveness as a Label Distribution Learning (LDL) problem, which can cope with incomplete or insufficient training data.

Another technique is asking a user to arrange images based on specified criteria. This strategy is more likely to yield a reliable result. But sorting a large dataset is impractical for users [38]. Pairwise rating is another way to show users two images and ask which one is more appealing [35, 36].

This strategy provides the user with a binary decision that can be faster than an absolute rating [37, 57, 73, 84]. Therefore, the pairwise rating was regarded in [121] and then converted to absolute aesthetic scores.

One of the FBP challenges in the research on facial attractiveness computation is a lack of realistic attractiveness labels (scores) and an insufficiently precise face representation [25]. This is because the facial aesthetics in images is substantially adjustable, and it has been proved that non-permanent features significantly influence the rating score. For example, makeup, eyeglasses, expression lighting, and photo quality strongly influence attractiveness [16, 17, 22]. Moreover, an average rating may ignore some individual preferences of attractiveness. While men and women agree on beauty generally [60], Schmid *et al*. [103] and Fan *et al*. [24] found that male raters tend to offer slightly higher ratings than females.

Regarding individual preference for beauty assessment, some studies take beauty prediction based on personal preference. Thus, the image's beauty score is given by a single rater and reflects his/her perception of beauty solely. Although significant results have been obtained in the appraisal of general attractiveness [62, 114], efforts to investigate personal beauty preferences are scarce and have some ethnic preference bias.

Developing ground truth to evaluate facial attractiveness rating is crucial as it considerably affects the model's accuracy and robustness. For instance, the number of images, beauty score levels, the number of raters, their diversity in both age and ethnicity, and the rating technique are all crucial aspects that contribute to the efficiency of the dataset and then model performance.

3.2. Facial Image Pre-Processing

Pre-processing is crucial after gaining an image because data pre-processing increases the accuracy of learning algorithms. Many image pre-processing methodologies have existed, as well as their performance on images taken in various environments [64, 58]. Pre-processing of data involved noise removal, face detection, face alignment, face cropping, and resizing. Face detection is a crucial stage that serves as the foundation for extracting facial features and performing beauty analysis. Landmarks are locations on a plane of a facial image that is morphometrically significant. Thus, landmark localization is essential in evaluating facial geometry to locate important face landmarks [37]. The accuracy of facial landmark tracking is affected by several parameters, including lighting conditions, camera motion, subject movement, and head direction [111].

Face alignment and cropping procedures have the potential to increase performance significantly. Additionally, posture and facial expression may influence one's perception of attractiveness [120]. Normalization is required when the value of features varies greatly between its maximum and minimum values. Normalization aids in scaling feature values to lie inside a given range [15]. Data augmentation techniques, such as flipping or rotation, are also used in this stage to increase the training sample size [25, 31].

3.3. Facial Image Beauty Representation

Facial beauty feature representations can be categorized into three main types:

1. Feature-based, which involves geometric and/or appearance features approaches.
2. A holistic approach where the entire image is used to predicate the facial beauty.
3. Hybrid approaches that use local features and the entire image.

These approaches are used for face attractiveness representation derived from the facial image to analyze and develop FBP models, as illustrated in Figure 8.

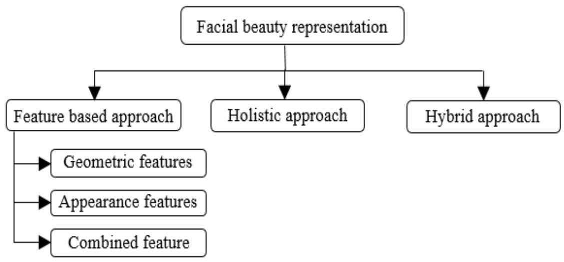


Figure 8. Facial image beauty-representation diagram.

Geometric features can be obtained from face landmark positions, distances between landmarks, ratios of these distances, angles, and inclinations [80, 76]. Most computer science research on finding attractive facial traits is geometric or landmark feature based [103]. The landmarks can be detected either manually [72] or automatically to represent the face in a variety of many different studies. The main aspect of evaluating facial geometry is the correct localization of landmarks on the vital facial regions, such as the face contour, eyebrows, eyes, nose, mouth, and chin [77]. Computing facial geometric distances and ratio vectors between them to extract significant facial features is accurate and resistant to lighting and background noises. In addition, geometric facial features based on significant data evidence have the greatest impact on facial attractiveness and should be treated first during facial surgery [105]. Increasing landmark numbers leads to more accurate face representation. However, this will require more computation complexity.

Psychological studies indicate that facial detail, color, smoothness, and lightness are significant for facial beauty perception [116]. Regarding computer insight, apparent features consider the local or overall appearances of faces as research items that are not limited to the simple amount or proportion relationships or costly manual landmarks of facial features to analyze beautiful traits [126]. A diversity of hand-crafted feature descriptors such as Gabor features, Local Binary Pattern (LPB), Histograms of Oriented Gradients (HOG), and Scale-Invariant Feature Transform (SIFT) could be utilized for facial appearance. This is because a human's judgment of facial beauty is influenced by facial appearance factors such as skin texture, color, gender, race, and age. For instance, due to sexual dimorphism, male and female faces are considerably dissimilar in shape and skin texture [74, 86]. However, representing facial features based on each type (geometric/appearance) alone might not be appropriate enough [23, 134]. Therefore, combining geometric and appearance features could give a better facial beauty representation.

On the other hand, gaining features from the entire image is a holistic method, usually with a Convolutional Neural Network (CNN). Holistic ways can incorporate all feature representations rather than taking a specific feature type and focusing on local discrete face parts in feature-based ways to produce more accurate predictions [116]. Consequently, hybridizing the feature above representation will include more information in the detail face layer and improve performance.

3.4. Feature Extraction and Selection

Generally, FBP was considered a classification, regression, or ranking problem. Regression and classification models are essential in predictive analytics. Most research regards face beauty prediction to be a fully supervised task. Few studies have used a semi-supervised technique, as in [21]. Assessing face beauty as a regression task with continuous values may provide a wider range of beauty evaluation scales than the discrete number or categorical attractiveness scores of the classification problem. Meanwhile, tackling FBP as a ranking can yield good results when dealing with a small dataset. Accurate feature extraction is essential for quantifying facial beauty reliably [7]. Many methods exist for extracting facial features from images. This could be done either based on hand-crafted features and traditional machine learning ways or by utilizing the embedded extraction of deep learning techniques. Recently DCNN and transfer learning notion show the efficiency in extracting informative facial beauty features in some studies and as a predictor of attractiveness score in others.

3.4.1. Traditional Methods

Hand-crafted features, such as facial shape and geometry features obtained from the landmarks and the distances and ratios between them. Researchers often build sets of landmarks using heuristics and putative rules of beauty such as the averageness, symmetry, golden ratio, the facial fifths and thirds, and neoclassic canons to extract such geometric features as in [1, 13, 24, 38, 40, 44, 80, 103]. The majority of researchers used a static facial image to model and predict facial attractiveness. Kalayci *et al*. [50] was the first study that combined 3-static and 7-dynamic geometric features from a video clip to boost the performance instead of using these feature sets alone. While Wei *et al*. [111] used Google Face API to extract geometrical features from 2D color video streams to create mobile software for FBP and to help plastic surgeons plan facial reconstruction surgery. More recently, capturing dynamic attractiveness features from facial appearance and landmarks were proposed by Weng *et al*. [113]. Thus, using a video clip that contains information about the mobility and dynamic behavior of the face provides a broader understanding and significant information in the facial attractiveness analyses. Generally, rule-driven features tend to lack generality when tested on various face datasets, and they may lose much of the feature information that distinguishes beautiful faces. Although landmark features and ratios appear to be correlated with facial attractiveness, extracting geometric characteristics requires a significant amount of effort, and it is yet unclear to what extent human brains use these features to form their notion of facial beauty.

Skin texture, colors, smoothness, and other non-permanent features, such as expressions and makeup, all affect the facial beauty's appearance. They are derived from the face image and employed as attributes in constructing attractiveness analysis and evaluation models. For instance, Hue, Saturation, and Value (HSV) and canny edge detection were used by Choudhary and Gandhi [15] to identify the color and the smoothness of the skin, respectively. LBP was used by Diamant *et al.* [19] to extract skin texture, and local phase quantization (LPQ), and Local Monotonic Pattern (LMP) were employed as well to extract local texture features followed by 2DPCA for dimensionality reduction. Meanwhile, a multi-scale K-means algorithm was utilized by Gan *et al*. [30] and Gan *et al*. [32] for the texture feature extraction. Texture features impose fewer constraints on face postures than geometrical features. However, details such as the spatial relationship between nearby structures (facial geometric) are neglected when apparent features are used only. That is why most researchers use a combination of both of them.

Eisenthal *et al*. [23] showed that the performance of landmark-based methods outperforms appearance-based methods and that combining the two improves performance. Thus, different features, including a combination of low and high features (multi-feature fusion), were applied in [12, 14, 15, 16, 82, 134]. Resulting in high-dimensional features. Principle Component Analysis (PCA) is one of the most widely utilized ways to minimize the size of the features [3, 57].

3.4.2. Deep Learning Methods

Learned features, on the other side, deep learning networks can be used for the extraction of facial beauty-related features and/or predicting facial beauty scores (see Table 2).

Table 2. The summary of facial beauty computational models based on deep learning approaches.

	Problem type	Dataset/#Images	Gender	Performance	
[29]	Classification	LSAFBD	F	ACC.	63.5%
[108]	Classification	4512	F/M	ACC.	98.1%
[70]	Regression	SCUT-FBP	F	PC	0.83
[33]	Classification	1600	F/M	ACC.	98.6%
[121]	Classification	JAFFE Face warehouse	F/M	ACC.	74.1%
[125]	Regression	LSAFBD SCUT-FBP	F/M	PC	0.88 0.92
[63]	Regression	MEBeauty SCUT-FBP5500	F/M F/M	PC PC	0.74 0.88.5
[25]	Regression	SCUT-FBP	F	PC	0.92
[37]	Regression	2056	F	PC	0.46
[97]	Regression	35	F/M	PC	0.61
[64]	Regression	HotOrNot	F/M	PC PC	0.49
[120]	Regression	SCUT-FBP HotOrNot	F F	PC PC	0.85 0.46
[31]	Classification	LSAFBD SCUT-FBP5500 CelebA	F/M F/M F/M	ACC.	68%
[72]	Ranking	HOTorNOT SCUT- FBP	F F	PC PC	0.46 0.81
[116]	Regression	SCUT-FBP	F	PC	0.88
[74]	Regression	SCUT-FBP5500	F/M	PC	0.90
[21]	Regression	SCUT-FBP5500 M2B SCUT-FBP	F/M	PC PC PC	86.60 48.05 84.64
[6]	Classification	CelebA	F/M	ACC.	82.5%
[67]	Ranking	FS-1500a LPH d	F/M F/M	ACC. ACC.	89.60% 84.80%
[117]	Regression	SCUT-FBP	F	PC	0.87
[34]	Regression	SCUT-FBP	F	PC	0.92
[104]	Classification	SCUT-FBP5500 CelebA.	F/M F/M	PC ACC.	0.92 85.6%
[100]	Classification	SCUT-FBP SCUT-FBP5500 CelebA	F F/M F/M	F1-Score F1-Score F1-Score	82.6% 78.6% 81.8%
[98]	Regression	SCUT-FBP MEBeauty SCUT-FBP5500	F F/M F/M	PC PC PC	0.879 0.888 0.886
[99]	Regression	SCUT-FBP MEBeauty SCUT-FBP5500	F F/M F/M	PC PC PC	0.91 0.93 0.926
[119]	Classification Regression	SCUT-FBP HotOrNot	F F	PC PC	0.87 0.48
[118]	Regression	SCUT-FBP5500 SCUT-FBP	F/M	PC	0.87 0.89
[97]	Classification	CelebA	F/M	ACC.	82.8%

- F: Female; M: Male; ACC: Accuracy; PC: Person Correlation, SCUT-FBP: South China University of Technology-Facial Beauty Prediction.

Lebedeva *et al*. [64] were pioneers in building FBP models without landmarks under unconstrained conditions based on texture features using a hierarchical feed-forward model to tackle the issue of facial feature landmark localization and imposing stringent limits on training samples. While apparent features were extracted by adaptive deconvolutional network AND were proposed in [33], that extracts significant information in an unsupervised way for FB evaluation. A CNN with a variety of network architectures takes the entire image as an input for the sack of facial beauty representation to extract informative and descriptive features as in [6, 25, 62, 67, 72, 70] and the efficiency of the inception model to extract multi-scale features of facial images for FBP has been proved in [29]. It was discovered that cascaded fine-tuning is an important strategy for combining feature information for face attractiveness prediction, as in [73, 116, 117] to enhance FBP's performance.

A pre-trained CNN can offer more than excellent results in computer vision [20], and due to the scarcity of annotated face images, training a DCNN directly may result in significant overfitting. Transfer learning has recently received much attention as it allows fine-tuning from a pre-trained model or employing the learned neural network as a feature extractor to complete the tasks. Such pre-trained models were used for face-related tasks. For instance, DenseNet in [63] pre-trained on ImageNet. Face verification based on VGG16 [31, 120], and VGG-face identification [21] is used for high-level feature extraction to predict facial attractiveness. A hybrid approach in [34] that simultaneously performed the FBP and the localization of facial landmarks based on appearance and geometric features extraction employing a multi-task deep learning approach. Meanwhile, pixel-wise labeling masks presented in [104] were fed to MobileNetV2 for learning the semantic representation of high-level features.

3.5. Predicting Facial Attractiveness Methods

Research has shown that machine has the capability to learn how to assess facial beauty. Studies conducted between 2001 and 2010 predominantly depended on statistical and shallow machine learning techniques to automatically predict the facial beauty score. Subsequently, the use of deep learning has become dominant due to its superior performance over traditional machine learning methods. More detailed information about facial image estimation approaches is illustrated in the next subsections.

3.5.1. Machine Learning Techniques

The derived features from facial images and the ground truth are utilized for training data samples. The training set should be sufficiently large and diversified to provide the classifier with adequate training [38]. The development of pattern analysis and machine learning technology has expanded the scope of study on facial attractiveness beyond subjective psychological cognition. Machine learning techniques can predict and improve facial beauty classification, regression, and ranking. Different machine learning algorithms and statistical techniques have been developed to map the representative aspects of a beauty score. This includes K-Nearest Neighbors (KNN) [1, 48], C4.5 [38, 40], Support Vector Machines (SVM) [30, 32, 33, 80, 134], Random Forest (RF) [62, 31] and Artificial Neural Networks (ANN) [67, 99]. Regression analysis includes Linear Regression (LR) [16, 23, 24, 103]. Support Vector Regression (SVR) [12, 13, 14, 64, 70] and Bayesian regression [120] as well. However, using statistical and traditional machine learning techniques for facial image aesthetic assessment becomes less effective with advanced deep learning approaches.

3.5.2. Deep Learning Techniques

Due to the difficulty of modeling beauty from fundamental principles, it is a perfect target for data-driven techniques, including deep learning [19]. Deep learning is extremely useful for computer vision applications. Stack layers are used to extract the features automatically. CNNs have been brought to sharp attention by many researchers in recent years as a new machine learning research technique [65]. DCNNs outperform hand-crafted descriptors in feature extraction and prediction. However, to complete a task, it may be necessary to create alternative network architectures and train DNNs roughly from scratch, which requires significant computational effort. A light CNN called Facial Image Attractiveness Classification Network (FIAC-Net) was proposed by Saeed *et al*. [100] to classify the attractiveness of facial images into four different beauty classes utilizing SCUT-FBP and SCUT-FBP5500 datasets. The FIAC-Net was also implemented on CelebA dataset to assess whether the facial image is attractive or unattractive, with a significant reduction in the number of the learnable parameters, and it provided competitive results.

Since deep learning techniques necessitate large amounts of data, which frequently scars, the information processing technology office of the defense advanced research projects agency’s department put out a new transfer learning task in 2005 as the capacity to recognize and adapt knowledge and abilities gained in prior tasks to new tasks [125]. According to this definition, transfer learning is extracting knowledge from one or more source tasks and applying it to a target task to tackle the overfitting issue caused due to data scarcity. In addition, transfer learning can help in saving training time. Therefore, transferring the knowledge of the pre-trained CNNs and Network fine-tuning in [116, 117] is preferable to training from scratch since the pre-trained models already have a large amount of task-related data, as in [21, 63, 108, 125]. To get the merits of training a network from scratch and the utilization of the pre-trained deep networks, Saeed *et al*. [98] proposed an ensemble regression model that includes three CNNs, one trained from scratch on the task of FBP, and two well-known (AlexNet, VGG16Net) were fine-tuned and retrained on estimating the facial aesthetic score. In the same context, a new ensemble loss function has been proposed by Romm [99] through combining L1, l2 and log-cosh loss functions. It contributes in enhancing the beauty prediction performance by leveraging the strengths of each individual loss function, as evaluated across various pretrained networks utilizing three FBP benchmarks.

Another aspect that impacts the deep learning process is the number of hidden layers. A deeper network is supposed to give a better outcome. Consequently, a deep residual network was used in [6, 25, 110, 119, 121] for FBP. However, deep architectures generally require more computational complexity. The complexity and the depth of neural network structure require many parameters and high dimensions that may lead to significant time consumption. Therefore, Zhai *et al*. [127] used a Broad Learning System (BLS) and fused local features with two-dimensional PCA for faster FBP training. Their proposed method could be considered an alternative to deep learning, considering the trade-off between training time and accuracy.

In addition to the classification and regression problems, FBP could be a ranking problem. Beauty deep ranking for a pair of facial images was implemented in [67, 72]. Ranking facial beauty can work relatively well with small datasets. Facial parsing masks with co-attention learning were adopted in [104] to enhance FBP.

Some studies tried simultaneously estimating the facial attractiveness level with distinct but related tasks. Multi-Task Learning (MTL) is a technique for concurrently learning multiple related tasks utilizing a common representation. It has demonstrated the potential to leverage the synergy between distinct but related tasks such as the prediction of gender, race and/or landmark localization in addition to the FBP as the main task as in [31, 34, 118], which in turn can improve the model accuracy.

3.6. Model Validation and Performance Evaluation

The performances vary depending on the variety of utilized datasets, as well as the variety of features and the learning methods. These models can estimate the facial attractiveness score as a regression problem and can be evaluated using Pearson Correlation (PC) or sometimes reefed as (r). It assesses the strength of the linear relationship between predicted beauty scores and actual human ratings. A high PC coefficient, close to +1, indicates that the model's predictions align closely with human judgments, suggesting that the model effectively captures the features influencing perceived facial attractiveness. While the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) measure the error rate among the target and the predicted scores [98]. Similarly, metrics such as the *F1-Score* and accuracy are employed to assess the efficiency of FBP based on the classification task [99, 100]. Greater *PC, F1-Score, Accuracy*, and lower *MAE* and *RMSE* values indicate higher prediction performance. These metrics are expressed as follows:

Given *N* testing images set:

$$PC = \frac{\sum_{i=1}^{N}(y_i - \bar{y})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{N}(p_i - \bar{p})^2}} \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - p_i| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - p_i)^2} \tag{4}$$

$$F1 - Score = \frac{2 * (precision * recall)}{(precision + recall)} \tag{5}$$

$$Accuracy = \frac{No.\,correct\ predictions}{Total\ No.\,predictions} * 100 \tag{6}$$

The label of the ground truth represented by y_i and p_i denotes the predicted score of the I^{th} image. The average of all labels in the ground truth is symbolized as $\bar{y}$ while $\bar{p}$ referred to the average of the predicted scores. Regarding classification tasks, precision is the proportion of true positives (correctly classified positive instances) out of all positive predictions made by the model [119]. It measures the accuracy of the positive predictions. Meanwhile, the recall is the proportion of true positives out of all actual positive instances in the dataset. It measures the ability of the model to identify positive instances correctly. Dividing the successfully predicted cases over the total number of forecasts yields performance accuracy [97].

4. The Importance of Facial Components and Automatic Beauty Assessment

There is a significant relationship between facial attractiveness and facial features. The attractiveness of the face and facial features are inextricably linked. The cultural effect on beauty judgment is extremely significant, as evidenced by the various canons created to evaluate female attractiveness over time. The Greeks liked an oval facial for men and women alike, with a narrow forehead to emphasize the hair. On the other hand, the ancient Egyptians thought a broad forehead and well-defined mandibles attractive [109]. There was evidence of a preference for wider foreheads and the absence of wrinkles throughout the Middle Ages, even if the criterion was not judged ugly at the time, as evidenced by modern positive remarks on grey hair [93]. However, cultural variations can also be seen within the same period.

Many research has been published in social, psychological [131], and medical viewpoints [89, 124, 51] that tried to answer the key question in facial attractiveness, namely, which parts of the face contribute to the overall attractiveness of a face and to what extent? as it is not wholly answered yet. This review tries to figure out which facial component has more impact on the attractiveness of a facial image. This is done according to the surveyed previous related work. This investigation, to our knowledge, has not been done before from FBP perspective as it is an emerging topic, and the resources are relatively few. Researchers in [104] presented that the facial components are more important for attractiveness prediction than the background. Consequently, it is clear from Table 3 that the shapes and look of the eyes is the most impact element in FBP followed by the shape of mouth and lips in a second degree of importance in assessing facial beauty.

Table 3. The impact of facial components on FBP.

Reference	Year	Eyes	Mouth/Lips	Face shape	Skin	Hairstyle	Eyebrow	Nose	Chin
[98]	2023	√	√			√			
[111]	2021	√							
[108]	2021				√		√		√
[29]	2020	√	√		√		√		√
[134]	2020			√	√				
[109]	2020		√						
[46]	2020	√	√						
[118]	2019	√				√			
[104]	2019	√						√	
[133]	2019	√		√			√		
[69]	2018			√					
[119]	2018		√	√				√	
[120]	2018					√			
[116]	2015	√	√		√				
[129]	2011			√					

Face shape also significantly influences facial attractiveness, consistent with earlier psychological studies. In addition, Zhao *et al*. [134] discovered that faces characterized by oval and heart-shaped have more beauty scores than other face forms. Faces with square shapes ranked the lowest on average regarding beautiful faces. Since facial skin can reflect one's age, wrinkles, and overall health, it affects how people perceive facial attractiveness. Similarly, the eyebrow and the hairstyle both have a comparable degree of importance with the facial skin when evaluating the aesthetic of a face. Zhao *et al*. [133] explored the relationship between facial beauty and its features by combining the global face shape and local geometric features of the eye and brow with computer big data analysis on the image of 300 Asian females and utilizing machine learning methods for FBP. They revealed that high-bending brows with small eyes and a round face, low-bending brows with big eyes, and a longer face would result in a higher attractiveness score for the facial image. Nose and chin could have approximately the same effect on the FBP, which might be relatively less effective on facial beauty assessment compared to the other facial components. This may reflect the limitation of 2D facial images in capturing some informative beauty-related features.

Lin *et al*. [74] have shown that other factors, such as gender and race, might have a more potent impact on facial attractiveness than geometric ratios. Meanwhile, Rhazi *et al*. [90] found that the main gender difference is detected in the forehead and chin portions. They also found that the chin and forehead in men tend to be stronger and larger than in women's. On the other hand, the greatest variations by age occur in areas such as the brows, nose, and chin. The brows decrease from high to low, making the eyes appear smaller. Similarly, the nasal tip gradually descends, causing nose growth, and the chin devolves the same way as the nose and brows.

Moreover, it was found by Komori *et al*. [57] that the height of the forehead form, which is associated with child schema, and the height of the cheekbone contour, which is related to sexual dimorphism, both have a significant impact on facial image aesthetic appraisal.

Figure 9 visually presents the contribution percentages of facial components to FBP, based on an extensive investigation of state-of-the-art approaches. The pie chart illustrates the relative importance of different facial components in the process of beauty assessment. Notably, the eyes emerge as the most influential factor, contributing significantly with a percentage of 25%. Following closely, the lips/mouth exhibit the second highest degree of importance, accounting for 18% of the overall contribution. Other prominent factors include face shape at 15%, skin at 12%, and hairstyle at 9%. The remaining 21% is distributed among eyebrows at 9%, while the nose and chin each account for 6% of the overall contribution. These findings provide valuable insights into the relative importance of facial components in assessing beauty and offer a foundation for further research and practical applications in the field of facial aesthetics assessment.

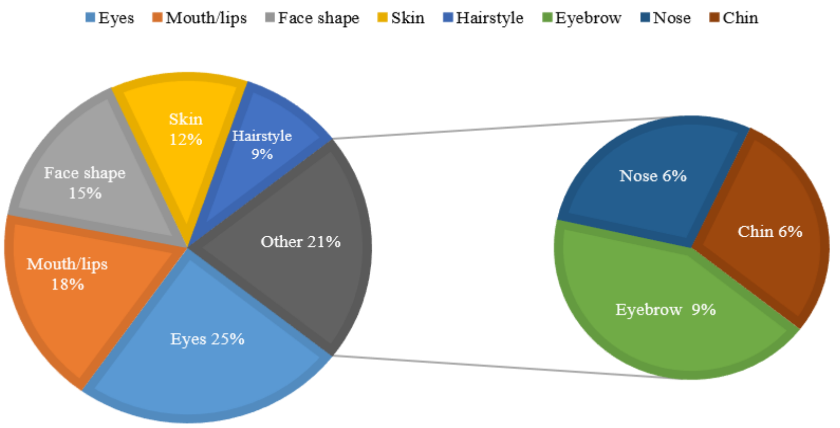


Figure 9. Facial components' contribution percentage to FBP based on the investigated state-of-the-art.

5. Challenges and Open Issues of Facial Images Automatic Beauty assessment

Despite the considerable strides in FBP, several obstacles still impede progress in this nascent research domain. Addressing these open issues and concerns is critical for optimizing the accuracy and generalizability of FBP models. This review seeks to highlight and explore these challenges, which encompass the following:

1. The first and main issue is the lack of sufficient public dedicated face beauty benchmark that satisfies all the required conditions, including the diversity of:
   a) Subject's ethnicity, gender, age, pose, expression, and level of beauty, particularly extremely attractive and unattractive faces due to the dominance of average faces.
   b) Rater's number, ethnicity, gender, age, and social background. One of the most serious challenges with FBP is the lack of universally agreed ground truth.
2. Most existing face databases are relatively small, gender-specific (female), and not publicly accessible. While deep learning approaches require massive data, this will lead to overfitting issues. In addition, the majority of databases are for female faces. This is

because assessing beauty with a male is more challenging than with a female. The public accessibility to the existing FBP databases is a worth mentioning issue that confronts the researchers due to the privacy issues of their authors. Consequently, predicting the attractiveness model becomes easier when restricting any of all the aspects above in 1 and 2. However, this would affect the accuracy of the model and the generalizability.

3. Most existing face databases are relatively small, gender-specific (female), and not publicly accessible. While deep learning approaches require massive data, this will lead to overfitting issues. In addition, the majority of databases are for female faces. This is because assessing beauty with a male is more challenging than with a female. The public accessibility to the existing FBP databases is a worth mentioning issue that confronts the researchers due to the privacy issues of their authors. Consequently, predicting the attractiveness model becomes easier when restricting any of all the aspects above in 1 and 2. However, this would affect:

a) The accuracy of the model and the generalizability.
b) How to represent the face and the role of beauty-related features in this representation and extraction are vital factors that should be considered. The majority of existing work is based on 2-dimensional images. However, many informative beauty-related features could not be captured based on 2D images. For instance, lateral face view contributes significantly to beauty assessment. Therefore, only 3D representations can represent the human face accurately. However, only a few studies utilized 3D facial images with some limitations, such as a small number of images.
c) The scarcity of realistic beauty scores or labels is due to the lack of an accurate facial depiction that captures the key features of attractiveness. This is due to various positions, facial expressions, low resolution, and lighting issues. For instance, some smiley faces are more appealing, and when two images of the same face are taken, the smiling image is often more beautiful. Thus, smiling faces are perceived as more attractive than neutral faces. Thereby eliminating the attractiveness gap between faces with various expressions allows realistic comparison and study of beauty and a more objective appraisal of facial beauty that is dominated by different features or elements and little impacted by other or superficial aspects. Moreover, it has been argued that the average rating may not consider the individual preference for attractiveness. Therefore, appropriate image analysis and rating methods significantly affect ground truth building. Although both men and women agree on the concept of attractiveness in general, studies have shown that some rating differences might occur between both genders. Where males provide slightly higher ratings than female raters. Consequently, developing an appropriate model for precisely predicting facial beauty levels remains challenging.
d) Many recent studies have used pre-trained DCNNs on various face-related social traits to estimate how people will react to new facial images based on these attributes. However, conflicting findings have been discovered when the performance of previously trained models has been compared [72]. Furthermore, it is unclear to what extent pre-trained DCNN attributes have accounted for the same or distinct variance in social evaluations based on facial features. The effectiveness of this strategy in generalizing out of sample across face databases and human raters has remained a barrier, which is an increasing concern in modern learning algorithms for practical applications.
e) Until now, there has been no complete answer to the question that asks, "Which face components and to what extent contribute to facial attractiveness?" This paper tries to figure out the answer to this question in terms of computer vision based on the reviewed FBP literature that finds the shape and the look of the eyes as one of the most effective elements in evaluating facial aesthetics.

6. Conclusions

Artists, philosophers, and scientists have been attempting to unravel the mystery of beauty for centuries. While evaluating facial attractiveness is a complex process that varies among individuals and over time, studies have found that people use consistent criteria. Computer vision can be utilized to create a robust attractiveness prediction model, despite the fact that facial beauty perception is connected to personal preference. Facial beauty evaluation has many medical, social, and psychological applications. However, due to a lack of sufficient benchmarks dedicated to facial attractiveness evaluation, there are limitations to developing an ideal assessment. This, in turn, requires diversity in several crucial aspects and techniques that affect facial beauty representation and assessment. The following points shed light on the key findings and implications for future FBP-based research.

AI and deep learning use regression and classification models in predictive analytics. Furthermore, quantifying facial beauty images as a regression task with continuous values may provide greater beauty judgment levels than discrete numbers or categorical attractiveness scores. Meanwhile, FBP ranking works well with small data.

The process of estimating the level of beauty in a facial image from a machine perspective is limited by the lack of sufficient, dedicated benchmarks and the way of representing the aesthetic-related features. This, in turn, requires diversity in several crucial aspects and techniques that affect facial beauty representation and assessment.

Early FBP research developed features based on heuristics and putative rules (averageness, symmetry, neoclassical and golden ratio). Most of these facial beauty principles involve ratios and distances between geometric facial features. These concepts and hand-crafted qualities are not universal. It also requires laborious face landmark annotation. Moreover, traditional descriptors were utilized in several research studies for local region appearance features or as a holistic image. Combining geometric and appearance features through multi-feature fusion and dimensionality reduction may improve results. With advanced deep learning neural networks, beauty feature extraction and prediction using statistical and conventional machine learning methods becomes less effective.

Deep data-driven techniques like CNNs have shown their efficiency in estimating facial image attractiveness as extractors and predictors. Better performance can be achieved with a deeper CNN structure, larger input images, and smaller convolution kernels. Deep learning requires large amounts of data. Thus, overfitting can result from inadequate data. Many effective ways exist to solve this problem. For instance, transfer learning and fine-tuning are used to adjust the parameters of a pre-trained network to acquire a better result in an acceptable time. Data augmentation is another method of increasing training samples to handle the overfitting issue. In addition, the use of dropout as a strategy to combat overfitting in neural networks is becoming increasingly popular.

This review analyses FBP literature to answer the most debatable question: “which facial features affect beauty assessment, and to what extent?”. Most of the literature that has been studied emphasizes the importance of the eyes and mouth (lips) when evaluating a face's attractiveness. However, it is commonly recognized that the nose affects facial beauty. According to FBP-reviewed studies, the nose has less impact on the facial beauty assessments based on the machine perspective. This is because 2D images cannot properly capture some informative features, such as the height of the nose. Thus, utilizing facial images with three dimensions can assist in representing facial features more accurately.

It has been concluded that automatic FBP combines both objective and subjective elements. While objective and data-driven features are used for analysis and scoring, the concept of the beauty itself in the construction of face dataset is inherently subjective. The models aim to provide objective beauty scores, but they may still reflect certain subjective biases present in the data or the cultural context in which they were developed. This makes the researchers strive to minimize subjectivity and improve model generalizability.

References

[1] Aarabi P., Hughes D., Mohajer K., and Emami M., “The Automatic Measurement of Facial Beauty,” *in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, e-Systems and e-Man for Cybernetics in Cyberspace*, Tucson, pp. 2644-2647, 2001. https://ieeexplore.ieee.org/abstract/document/972963

[2] Abrahams T. and Bein D., “Analyzing Beauty by Building Custom Profiles Using Machine Learning,” *in Proceedings of the IEEE 9th Annual Computing and Communication Workshop and Conference*, Las Vegas, pp. 0372-0376, 2019. https://ieeexplore.ieee.org/abstract/document/8666528

[3] Ahmadimehr S. and Moridani M., “Evaluating Facial Attractiveness through Proportions Analysis Based on Geometric Features,” *Journal of Image Processing and Pattern Recognition Progress*, vol. 7, no. 2, pp. 20-26, 2020. file:///C:/Users/user/Downloads/2186-10383-1-PB.pdf

[4] Alley T. and Cunningham M., “Article Commentary: Averaged Faces are Attractive, but very Attractive Faces are not Average,” *Psychological Science*, vol. 2, no. 2, pp. 123-125, 1991. https://doi.org/10.1111/j.1467-9280.1991.tb00113.x

[5] Altwaijry H. and Belongie S., “Relative Ranking of Facial Attractiveness,” *in Proceedings of the IEEE Workshop on Applications of Computer Vision*, Florida, pp. 117-124, 2013. https://ieeexplore.ieee.org/abstract/document/6475008

[6] Anderson R., Gema A., Suharjito., and Isa S., “Facial Attractiveness Classification Using Deep Learning,” *in Proceedings of the Indonesian Association for Pattern Recognition International Conference*, Jakarta, pp. 34-38, 2018. DOI:10.1109/INAPR.2018.8627004

[7] Ayeche F. and Alti A., “Novel Descriptors for Effective Recognition of Face and Facial Expressions,” *Revue d'Intelligence Artificielle*, vol. 34, no. 5, pp. 521-530, 2020. https://doi.org/10.18280/ria.340501

[8] Berssenbrugge P., Berlin N., Kebeck G., Runte C., Jung S., Kleinheinz J., and Dirksen D., “2D and 3D Analysis Methods of Facial Asymmetry in Comparison,” *Journal of Cranio-Maxillofacial*

Surgery, vol. 42, no. 6, pp. 327-334, 2014. https://doi.org/10.1016/j.jcms.2014.01.028

[9] Bonilla-Zorita G., Griffiths M., and Kuss D., "Online Dating and Problematic Use: A Systematic Review," *International Journal of Mental Health and Addiction*, vol. 19, no. 6, pp. 2245-2278, 2020. https://link.springer.com/article/10.1007/s11469-020-00318-9

[10] Bozkir M., Karakas P., and Oguz O., "Vertical and Horizontal Neoclassical Facial Canons in Turkish Young Adults," *Surgical and Radiologic Anatomy*, vol. 26, no. 3, pp. 212-219, 2004. https://link.springer.com/article/10.1007/s00276-003-0202-2

[11] Brozovsky L. and Petricek V., "Recommender System for Online Dating Service," *arXiv Preprint*, vol. arXiv:cs/0703042, pp.1-12, 2007. https://arxiv.org/pdf/cs/0703042

[12] Chen F., Xiao X., and Zhang D., "Data-Driven Facial Beauty Analysis: Prediction, Retrieval and Manipulation," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 205-216, 2016. https://10.1109/TAFFC.2016.2599534

[13] Chen F., Xu Y., and Zhang D., "A New Hypothesis on Facial Beauty Perception," *ACM Transactions on Applied Perception*, vol. 11, no. 2, pp. 1-20, 2014. https://dl.acm.org/doi/10.1145/2622655

[14] Chen F., Zhang D., Wang C., and Duan X., "Comparison and Fusion of Multiple Types of Features for Image-based Facial Beauty Prediction," *in Proceedings of the Chinese Conference on Biometric Recognition*, Shenzhen, pp. 23-30, 2017. https://link.springer.com/chapter/10.1007/978-3-319-69923-3_3

[15] Choudhary G. and Gandhi T., "Indexing Facial Attractiveness and Well Beings Using Machine Learning," *in Proceedings of the IEEE Region 10 Humanitarian Technology Conference*, Agra, pp. 1-6, 2016. https://ieeexplore.ieee.org/abstract/document/7906813

[16] Dantcheva A. and Dugelay J., "Assessment of Female Facial Beauty Based on Anthropometric, Non-Permanent and Acquisition Characteristics," *Multimedia Tools and Applications*, vol. 74, no. 24, pp. 11331-11355, 2015. https://doi.org/10.1007/s11042-014-2234-5

[17] Dantcheva A. and Dugelay J., "Female Facial Aesthetics Based on Soft Biometrics and Photo-Quality," *in Proceedings of IEEE International Conference for Multimedia and Expo*, Barcelona, pp. 1-6, 2011. https://www.eurecom.fr/publication/3368

[18] Deng Z. and Zhang Z., "Event-Related Complexity Analysis and its Application in the Detection of Facial Attractiveness," *International Journal of Neural Systems*, vol. 24, no. 7, pp. 1450026, 2014. DOI:10.1142/S0129065714500269

[19] Diamant N., Zadok D., Baskin C., Schwartz E., and Bronstein A., "Beholder-GAN: Generation and Beautification of Facial Images with Conditioning on their Beauty Level," *in Proceedings of the IEEE International Conference on Image Processing*, Taipei, pp. 739-743, 2019. https://ieeexplore.ieee.org/abstract/document/8803807

[20] Dornaika F., Moujahid A., Wang K., and Feng X., "Efficient Deep Discriminant Embedding: Application to Face Beauty Prediction and Classification," *Engineering Applications of Artificial Intelligence*, vol. 95, pp. 103831, 2020. https://doi.org/10.1016/j.engappai.2020.103831

[21] Dornaika F., Wang K., Arganda-Carreras I., Elorza A., and Moujahid A., "Toward Graph-based Semi-Supervised Face Beauty Prediction," *Expert Systems with Applications*, vol. 142, pp. 112990, 2020. https://doi.org/10.1016/j.eswa.2019.112990

[22] Eggerstedt M., Rhee J., Urban M., Mangahas A., Smith R., and Revenaugh P., "Beauty is in the Eye of the Follower: Facial Aesthetics in the Age of Social Media," *American Journal of Otolaryngology*, vol. 41, no. 6, pp. 102643, 2020. https://doi.org/10.1016/j.amjoto.2020.102643

[23] Eisenthal Y., Dror G., and Ruppin E., "Facial Attractiveness: Beauty and the Machine," *Neural Computation*, vol. 18, no. 1, pp. 119-142, 2006. DOI:10.1162/089976606774841602

[24] Fan J., Chau K., Wan X., Zhai L., and Lau E., "Prediction of Facial Attractiveness from Facial Proportions," *Pattern Recognition*, vol. 45, no. 6, pp. 2326-2334, 2012. DOI:10.1016/j.patcog.2011.11.024

[25] Fan Y., Liu S., Li B., Guo Z., Samal A., Wan J., and Li S., "Label Distribution-Based Facial Attractiveness Computation by Deep Residual Learning," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2196-2208, 2017. DOI:10.1109/TMM.2017.2780762

[26] Farkas L., Hreczko T., Kolar J., and Munro I., "Vertical and Horizontal Proportions of the Face in Young Adult North American Caucasians: Revision of Neoclassical Canons," *Plastic and Reconstructive Surgery*, vol. 75, no. 3, pp. 328-338, 1985. https://pubmed.ncbi.nlm.nih.gov/3883374/

[27] Fink B. and Neave N., "The Biology of Facial Beauty," *International Journal of Cosmetic Science*, vol. 27, no. 6, pp. 317-325, 2005. https://onlinelibrary.wiley.com/doi/10.1111/j.1467-2494.2005.00286.x

[28] Gabrick K., Wu R., Singh A., Bartlett S., Taylor J., and Persing J., "Assessing Facial Asymmetry in Postoperative Patients with Unilateral Coronal Craniosynostosis," *Journal of Craniofacial Surgery*, vol. 31, no. 4, pp. 1000-1005, 2020. DOI:10.1097/SCS.0000000000006355

[29] Gan J., Jiang K., Tan H., and He G., "Facial Beauty Prediction Based on Lighted Deep Convolution Neural Network with Feature Extraction Strengthened," *Chinese Journal of Electronics*, vol. 29, no. 2, pp. 312-321, 2020. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cje.2020.01.009

[30] Gan J., Wang B., and Xu Y., "A Novel Method for Predicting Facial Beauty under Unconstrained Condition," *in Proceedings of the 8th International Conference on Image and Graphics*, Tianjin, pp. 350-360, 2015. https://link.springer.com/chapter/10.1007/978-3-319-21978-3_31

[31] Gan J., Xiang L., Zhai Y., Mai C., He G., and Zeng J., "2M BeautyNet: Facial Beauty Prediction Based on Multi-Task Transfer Learning," *IEEE Access*, vol. 8, pp. 20245-20256, 2020. DOI:10.1109/ACCESS.2020.2968837

[32] Gan J., Zhai Y., and Wang B., "Unconstrained Facial Beauty Prediction Based on Multi-Scale K-Means," *Chinese Journal of Electronics*, vol. 26, no. 3, pp. 548-556, 2017. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cje.2016.10.020

[33] Gan J., Zhou L., and Zhai Y., "A Study for Facial Beauty Prediction Model," *in Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition*, Guangzhou, pp. 8-13, 2015. https://ieeexplore.ieee.org/abstract/document/7295918

[34] Gao L., Li W., Huang Z., Huang D., and Wang Y., "Automatic Facial Attractiveness Prediction by Deep Multi-Task Learning," *in Proceedings of the 24th International Conference on Pattern Recognition*, Beijing, pp. 3592-3597, 2018. https://ieeexplore.ieee.org/document/8545033

[35] Gattupalli V., Chandakkar P., and Li B., "A Computational Approach to Relative Aesthetics," *in Proceedings of the 23rd International Conference on Pattern Recognition*, Cancun, pp. 2446-2451, 2016. https://ieeexplore.ieee.org/abstract/document/7900003

[36] Gill D., "Women and Men Integrate Facial Information Differently in Appraising the Beauty of a Face," *Evolution and Human Behavior*, vol. 38, no. 6, pp. 756-760, 2017. https://doi.org/10.1016/j.evolhumbehav.2017.07.001

[37] Gray D., Yu K., Xu W., and Gong Y., "Predicting Facial Beauty without Landmarks," *in Proceedings of the 11th European Conference on Computer Vision*, Crete, pp. 434-447, 2010. https://link.springer.com/chapter/10.1007/978-3-642-15567-3_32

[38] Gunes H. and Piccardi M., "Assessing Facial Beauty through Proportion Analysis by Image Processing and Supervised Learning," *International Journal of Human-Computer Studies*, vol. 64, no. 12, pp. 1184-1199, 2006. https://doi.org/10.1016/j.ijhcs.2006.07.004

[39] Gunes H., "A Survey of Perception and Computation of Human Beauty," *in Proceedings of the Joint ACM Workshop on Human Gesture and Behavior Understanding*, Arizona, pp. 19-24, 2011. https://dl.acm.org/doi/10.1145/2072572.2072580

[40] Gunes H., Piccardi M., and Jan T., "Automated Classification of Female Facial Beauty Using Learning Algorithms," *in Proceedings of the Image and Vision Computing Conference*, New Zealand, pp. 169-173 2002. https://opus.lib.uts.edu.au/handle/10453/7053

[41] Halberstadt J. and Rhodes G., "It's not just Average Faces that are Attractive: Computer-Manipulated Averageness Makes Birds, Fish, and Automobiles Attractive," *Psychonomic Bulletin and Review*, vol. 10, no. 1, pp. 149-156, 2003. DOI: 10.3758/BF03196479

[42] Han W., Chen F., and Sun F., "Can We Extract 3D Biometrics from 2D Images for Facial Beauty Analysis?," *in Proceedings of the 9th International Conference on Computing and Pattern Recognition*, Xiamen, pp. 213-217, 2020. https://dl.acm.org/doi/10.1145/3436369.3436476

[43] Hicks K. and Thomas J., "The Changing Face of Beauty: A Global Assessment of Facial Beauty," *Otolaryngologic Clinics of North America*, vol. 53, no. 2, pp. 185-194, 2020. https://doi.org/10.1016/j.otc.2019.12.005

[44] Hong Y., Nam G., Choi H., Cho J., and Kim I., "A Novel Framework for Assessing Facial Attractiveness Based on Facial Proportions," *Symmetry*, vol. 9, no. 12, pp. 1-12, 2017. DOI:10.3390/sym9120294

[45] Hu Y., Abbasi N., Zhang Y., and Chen H., "The Effect of Target Sex, Sexual Dimorphism, and Facial Attractiveness on Perceptions of Target Attractiveness and Trustworthiness," *Frontiers in Psychology*, vol. 9, pp. 1-8, 2018. https://pmc.ncbi.nlm.nih.gov/articles/PMC6002689/

[46] Ibanez-Berganza M., Amico A., Lancia G., Maggiore F., Monechi B., and Loreto V., "Unsupervised Inference Approach to Facial Attractiveness," *PeerJ*, vol. 8, pp. 1-20, 2020. DOI:10.7717/peerj.10210

[47] Iyengar A., Kulkarni R., and Vidya T., "Koinophilia and Human Facial Attractiveness,"

Resonance, vol. 20, no. 4, pp. 311-319, 2015. https://doi.org/10.1007/s12045-015-0187-2

[48] Iyer T., Rahul K., Nersisson R., Zhuang Z., Raj A., and Refayee I., "Machine Learning-based Facial Beauty Prediction and Analysis of Frontal Facial Images Using Facial Landmarks and Traditional Image Descriptors," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, 2021. https://doi.org/10.1155/2021/4423407

[49] Kagian A., Dror G., Leyvand T., Cohen-Or D., and Ruppin E., "A Humanlike Predictor of Facial Attractiveness," *in Proceedings of the 19th International Conference on Neural Information Processing Systems*, Canada, pp. 649-656, 2006. https://dl.acm.org/doi/10.5555/2976456.2976538

[50] Kalayci S., Ekenel H., and Gunes H., "Automatic Analysis of Facial Attractiveness from Video," *in Proceedings of the IEEE International Conference on Image Processing*, Paris, pp. 4191-4195, 2014. https://ieeexplore.ieee.org/abstract/document/7025851

[51] Kanavakis G., Halazonetis D., Katsaros C., and Gkantidis N., "Facial Shape Affects Self-Perceived Facial Attractiveness," *PloS ONE*, vol. 16, no. 2, pp. 1-17, 2021. https://doi.org/10.1371/journal.pone.0245557

[52] Karaca Saygili O., Cinar S., Gulcen B., Ozcan E., and Kus I., "The Validity of Eight Neoclassical Facial Canons in the Turkish Adults," *Folia Morphol (Warsz)*, vol. 75, no. 4, pp. 512-517, 2016. https://pubmed.ncbi.nlm.nih.gov/27830884/

[53] Karras T., Laine S., and Aila T., "A Style-Based Generator Architecture for Generative Adversarial Networks," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, California, pp. 4396-4405, 2019. https://ieeexplore.ieee.org/document/8953766

[54] Khalid S., Goldenberg M., Grantcharov T., Taati B., and Rudzicz F., "Evaluation of Deep Learning Models for Identifying Surgical Actions and Measuring Performance," *JAMA Network Open*, vol. 3, no. 3, pp. 201664-201664, 2020. DOI:10.1001/jamanetworkopen.2020.1664

[55] Knoops P., Papaioannou A., Borgh A., Breakey R., Wilson A., Jeelani O., Zafeiriou S., Steinbacher D., Padwa B., Dunaway D., and Schievano S., "A Machine Learning Framework for Automated Diagnosis and Computer-Assisted Planning in Plastic and Reconstructive Surgery," *Scientific Reports*, vol. 9, no. 1, pp. 1-12, 2019. https://doi.org/10.1038/s41598-019-49506-1

[56] Kocsor F., Feldmann A., Bereczkei T., and Kallai J., "Assessing Facial Attractiveness: Individual Decisions and Evolutionary Constraints," *Socioaffective Neuroscience and Psychology*, vol. 3, no. 1, pp. 1-11, 2013. https://www.tandfonline.com/doi/full/10.3402/snp.v3i0.21432

[57] Komori M., Shiroshita K., Nakagami M., Nakamura K., Kobayashi M., and Watanabe K., "Investigation of Facial Preference Using Gaussian Process Preference Learning and Generative Image Model," *in Proceedings of the 20th International Conference on Computer Information Systems and Industrial Management*, Poland, pp. 193-202, 2021. https://link.springer.com/chapter/10.1007/978-3-030-84340-3_15

[58] Kumar Y., Verma S., and Sharma S., "Multi-Pose Facial Expression Recognition Using Hybrid Deep Learning Model with Improved Variant of Gravitational Search Algorithm," *The International Arab Journal of Information Technology*, vol. 19, no. 2, pp. 281-287, 2022. https://doi.org/10.34028/iajit/19/2/15

[59] Langlois J. and Roggman L., "Attractive Faces are only Average," *Psychological Science*, vol. 1, no. 2, pp. 115-121, 1990. https://journals.sagepub.com/doi/abs/10.1111/j.1467-9280.1990.tb00079.x?journalCode=pssa

[60] Langlois J., Kalakanis L., Rubenstein A., Larson A., Hallam M., and Smoot M., "Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review," *Psychological Bulletin*, vol. 126, no. 3, pp. 390-423, 2000. DOI:10.1037//0033-2909.126.3.390

[61] Laurentini A. and Bottino A., "Computer Analysis of Face Beauty: A Survey," *Computer Vision and Image Understanding*, vol. 125, pp. 184-199, 2014. https://doi.org/10.1016/j.cviu.2014.04.006

[62] Lebedeva I., Guo Y., and Ying F., "Deep Facial Features for Personalized Attractiveness Prediction," *in Proceedings of the 13th International Conference on Digital Image Processing*, Singapore, pp. 1-9, 2021. https://doi.org/10.1117/12.2599699

[63] Lebedeva I., Guo Y., and Ying F., "MEBeauty: A Multi-Ethnic Facial Beauty Dataset in-the-Wild," *Neural Computing and Applications*, vol. 34, no. 17, pp. 14169-14183, 2021. https://link.springer.com/article/10.1007/s00521-021-06535-0

[64] Lebedeva I., Guo Y., and Ying F., "Transfer Learning Adaptive Facial Attractiveness Assessment," *Journal of Physics: Conference Series*, vol. 1922, no. 1, pp. 1-11, 2021. DOI:10.1088/1742-6596/1922/1/012004

[65] LeCun Y., Bengio Y., and Hinton G., "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015. https://doi.org/10.1038/nature14539

[66] Lee A., De La Mare J., Moore H., and Umeh P., "Preference for Facial Symmetry Depends on Study Design," *Symmetry*, vol. 13, no. 9, pp. 1-15,

2021. https://www.mdpi.com/2073-8994/13/9/1637

[67] Lei J., Feng Z., Song M., and Tao D., "Which Face is more Attractive?," *in Proceedings of the IEEE International Conference on Image Processing*, Phoenix, pp. 1639-1643, 2016. https://ieeexplore.ieee.org/abstract/document/7532636

[68] Li J., Xiong C., Liu L., Shu X., and Yan S., "Deep Face Beautification," *in Proceedings of the 23rd ACM International Conference on Multimedia*, Brisbane, pp. 793-794, 2015. https://dl.acm.org/doi/10.1145/2733373.2807966

[69] Liang L., Lin L., Jin L., Xie D., and Li M., "SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction," *in Proceedings of the 24th International Conference on Pattern Recognition*, Beijing, pp. 1598-1603, 2018. https://ieeexplore.ieee.org/abstract/document/8546038

[70] Liang L., Xie D., Jin L., Xu J., Li M., and Lin L., "Region-Aware Scattering Convolution Networks for Facial Beauty Prediction," *in Proceedings of the IEEE International Conference on Image Processing*, Beijing, pp. 2861-2865, 2017. https://ieeexplore.ieee.org/abstract/document/8296805

[71] Liang X., Tong S., Kumada T., and Iwaki S., "Golden Ratio: The Attributes of Facial Attractiveness Learned by CNN," *in Proceedings of the IEEE International Conference on Image Processing*, Taipei, pp. 2124-2128, 2019. https://ieeexplore.ieee.org/abstract/document/8803166

[72] Liao Y. and Deng W., "Deep Rank Learning for Facial Attractiveness," *in Proceedings of the 4th Asian Conference on Pattern Recognition*, Nanjing, pp. 565-570, 2017. https://ieeexplore.ieee.org/abstract/document/8575885

[73] Lin L., Liang L., and Jin L., "Regression Guided by Relative Ranking Using Convolutional Neural Network (R^3CNN) for Facial Beauty Prediction," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 122-134, 2019. DOI:10.1109/TAFFC.2019.2933523

[74] Lin L., Liang L., Jin L., and Chen W., "Attribute-Aware Convolutional Neural Networks for Facial Beauty Prediction," *in Proceedings of the 28th International Joint Conference on Artificial Intelligence Main Track*, Macao, pp. 847-853, 2019. https://doi.org/10.24963/ijcai.2019/119

[75] Little A., Jones B., and DeBruine L., "Facial Attractiveness: Evolutionary Based Research," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1571, pp. 1638-1659, 2011. https://pubmed.ncbi.nlm.nih.gov/21536551/

[76] Liu S., Fan Y., Guo Z., Samal A., and Ali A., "A Landmark-Based Data-Driven Approach on 2.5 D Facial Attractiveness Computation," *Neurocomputing*, vol. 238, pp. 168-178, 2017. https://doi.org/10.1016/j.neucom.2017.01.050

[77] Liu S., Fan Y., Samal A., and Guo Z., "Advances in Computational Facial Attractiveness Methods," *Multimedia Tools and Applications*, vol. 75, no. 23, pp. 16633-16663, 2016. https://doi.org/10.1007/s11042-016-3830-3

[78] Liu X., Peng H., Yin M., Chen C., and Li X., "Face Beautification: Beyond Makeup Transfer," *Frontiers in Computer Science*, vol. 4, pp. 1-12, 2022. https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2022.910233/full

[79] Liu Z., Luo P., Wang X., and Tang X., "Deep Learning Face Attributes in the Wild," *in Proceedings of the IEEE International Conference on Computer Vision*, Santiago, pp. 3730-3738, 2015. https://ieeexplore.ieee.org/document/7410782

[80] Mao H., Jin L., and Du M., "Automatic Classification of Chinese Female Facial Beauty Using Support Vector Machine," *in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, San Antonio, pp. 4842-4846, 2009. https://ieeexplore.ieee.org/abstract/document/5346057

[81] Meenai Y. and Abbas M., "Beauty Measuring System based on the Divine Ratio an Exploratory Research," *in Proceedings of the International Conference on User Science and Engineering (i-USEr)*, Shah Alam, pp. 189-192, 2010. DOI:10.1109/IUSER.2010.5716749

[82] Mu Y., "Computational Facial Attractiveness Prediction by Aesthetics-Aware Features," *Neurocomputing*, vol. 99, pp. 59-64, 2013. https://doi.org/10.1016/j.neucom.2012.06.020

[83] Nguyen T. and Liu L., "Smart Mirror: Intelligent Makeup Recommendation and Synthesis," *in Proceedings of the 25th ACM International Conference on Multimedia*, California, pp. 1253-1254, 2017. https://doi.org/10.1145/3123266.3127926

[84] Nguyen T., Liu S., Ni B., Tan J., Rui Y., and Yan S., "Sense Beauty Via Face, Dressing, and/or Voice," *in Proceedings of the 20th ACM international Conference on Multimedia*, Nara, pp. 239-248, 2012. https://dl.acm.org/doi/10.1145/2393347.2393385

[85] Pallett P., Link S., and Lee K., "New "Golden" Ratios for Facial Beauty," *Vision Research*, vol. 50, no. 2, pp. 149-154, 2010. https://doi.org/10.1016/j.visres.2009.11.003

[86] Perrett D., Lee K., Penton-Voak I., Rowland D., Yoshikawa S., Burt D., Henzi S., Castles D., and Akamatsu S., "Effects of Sexual Dimorphism on Facial Attractiveness," *Nature*, vol. 394, pp. 884-887, 1998. https://www.nature.com/articles/29772

[87] Prokopakis E., Polat S., and Cevik Y., "The Golden Ratio in Facial Symmetry," *Rhinology*, vol. 51, no. 1, pp. 18-21, 2013. https://pubmed.ncbi.nlm.nih.gov/23441307/

[88] Redi M., Rasiwasia N., Aggarwal G., and Jaimes A., "The Beauty of Capturing Faces: Rating the Quality of Digital Portraits," *in Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, Ljubljana, pp. 1-8, 2015. https://ieeexplore.ieee.org/abstract/document/7163086

[89] Ren H., Chen X., and Zhang Y., "Correlation between Facial Attractiveness and Facial Components Assessed by Laypersons and Orthodontists," *Journal of Dental Sciences*, vol. 16, no. 1, pp. 431-436, 2021. https://10.1016/j.jds.2020.07.012

[90] Rhazi M., Zarghili A., Majda A., Bouzalmat A., and Oufkir A., "Facial Beauty Analysis by Age and Gender," *International Journal of Intelligent Systems Technologies and Applications*, vol. 18, no. 1-2, pp. 179-203, 2019. https://doi.org/10.1504/IJISTA.2019.097757

[91] Rhodes G., "The Evolutionary Psychology of Facial Beauty," *Annual Review of Psychology*, vol. 57, pp. 199-226, 2006. DOI:10.1146/annurev.psych.57.102904.190208

[92] Rizvi Q., Karawia A., and Kumar S., "Female Facial Beauty Analysis for Assesment of Facial Attractivness," *in Proceedings of the 2nd International Conference on Information Management in the Knowledge Economy*, Chandigarh, pp. 156-160, 2013. https://ieeexplore.ieee.org/abstract/document/6915090

[93] Romm S., "The Changing Face of Beauty," *Aesthetic Plastic Surgery*, vol. 13, no. 2, pp. 91-98, 1989. DOI:10.1007/BF01571475

[94] Rothe R., Timofte R., and Van Gool L., "Some Like it Hot-Visual Guidance for Preference Prediction," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, pp. 5553-5561, 2016. https://ieeexplore.ieee.org/document/7780968

[95] Roy H., Dhar S., Dey K., Acharjee S., and Ghosh D., *Advanced Computational and Communication Paradigms*, Springer, 2018, https://link.springer.com/chapter/10.1007/978-981-10-8237-5_73

[96] Saeed J. and Abdulazeez A., "Facial Beauty Prediction and Analysis Based on Deep Convolutional Neural Network: A Review," *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 1-12, 2021. https://doi.org/10.30880/jscdm.2021.02.01.001

[97] Saeed J., Abdulazeez A., and Ibrahim D., "2D Facial Images Attractiveness Assessment Based on Transfer Learning of Deep Convolutional Neural Networks," *in Proceedings of the 4th International Conference on Advanced Science and Engineering*, Zakho, pp. 13-18, 2022. https://ieeexplore.ieee.org/document/10075585

[98] Saeed J., Abdulazeez A., and Ibrahim D., "An *Ensemble* DCNNs-Based Regression Model for Automatic Facial Beauty Prediction and Analyzation," *Traitement du Signal*, vol. 40, no. 1, pp. 55-63, 2023. https://iieta.org/journals/ts/paper/10.18280/ts.400105

[99] Saeed J., Abdulazeez A., and Ibrahim D., "Automatic Facial Aesthetic Prediction Based on Deep Learning with Loss Ensembles," *Applied Sciences*, vol. 13, no. 17, pp. 9728, 2023. https://doi.org/10.3390/app13179728

[100] Saeed J., Abdulazeez A., and Ibrahim D., "FIAC-Net: Facial Image Attractiveness Classification Based on Light Deep Convolutional Neural Network," *in Proceedings of the 2nd International Conference on Computer Science, Engineering and Applications*, Gunupur, pp. 1-6, 2022. DOI:10.1109/ICCSEA54677.2022.9936582

[101] Sarver D. and Jacobson R., "The Aesthetic Dentofacial Analysis," *Clinics in Plastic Surgery*, vol. 34, no. 3, pp. 369-394, 2007. https://pubmed.ncbi.nlm.nih.gov/17692698/

[102] Saxton T., DeBruine L., Jones B., Little A., and Roberts S., "A Longitudinal Study of Adolescents' Judgments of the Attractiveness of Facial Symmetry, Averageness and Sexual Dimorphism," *Journal of Evolutionary Psychology*, vol. 9, no. 1, pp. 43-55, 2011. https://doi.org/10.1556/jep.9.2011.22.1

[103] Schmid K., Marx D., and Samal A., "Computation of a Face Attractiveness Index Based on Neoclassical Canons, Symmetry, and Golden Ratios," *Pattern Recognition*, vol. 41, no. 8, pp. 2710-2717, 2008. https://doi.org/10.1016/j.patcog.2007.11.022

[104] Shi S., Gao F., Meng X., Xu X., and Zhu J., "Improving Facial Attractiveness Prediction Via Co-Attention Learning," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, pp. 4045-4049, 2019. https://ieeexplore.ieee.org/document/8683112

[105] Stepanek L., Kasal P., and Mestak J., "Evaluation of Facial Attractiveness for Purposes of Plastic Surgery Using Machine-Learning Methods and Image Analysis," *in Proceedings of the IEEE 20th*

International Conference on e-Health Networking, Applications and Services (Healthcom), Ostrava, pp. 1-6, 2018. https://ieeexplore.ieee.org/document/8531195

[106] Talamas N., Mavor K., and Perrett D., "Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance," *PloS One*, vol. 11, no. 2, pp. 1-18, 2016. https://doi.org/10.1371/journal.pone.0148284

[107] Thapa G. and Thapa R., "The Relation of Golden Ratio, Mathematics and Aesthetics," *Journal of the Institute of Engineering*, vol. 14, no. 1, pp. 188-199, 2018. https://doi.org/10.3126/jie.v14i1.20084

[108] Tong S., Liang X., Kumada T., and Iwaki S., "Putative Ratios of Facial Attractiveness in a Deep Neural Network," *Vision Research*, vol. 178, pp. 86-99, 2021. https://doi.org/10.1016/j.visres.2020.10.001

[109] Ulrich L., Dugelay J., Vezzetti E., Moos S., and Marcolin F., "Perspective Morphometric Criteria for Facial Beauty and Proportion Assessment," *Applied Sciences*, vol. 10, no. 1, pp. 1-17, 2020. https://doi.org/10.3390/app10010008

[110] Vijayarajan V., Roomi S., Maheswari P., and Rajeae R., "Celebrity: Creation of A Database towards an Automatic Beauty Evaluation System," *in Proceedings of the 8th International Conference on Business and Industrial Research*, Bangkok, pp. 1291-1296, 2023. https://ieeexplore.ieee.org/document/10147452

[111] Wei W., Ho E., McCay K., Damasevicius R., Maskeliunas R., and Esposito A., "Assessing Facial Symmetry and Attractiveness Using Augmented Reality," *Pattern Analysis and Applications*, vol. 25, pp. 635-651, 2021. https://doi.org/10.1007/s10044-021-00975-z

[112] Weinberg S., Raffensperger Z., and Marazita M., "The 3D Facial Norms Database: Part 1. A Web-Based Craniofacial Anthropometric and Image Repository for the Clinical and Research Community," *The Cleft Palate-Craniofacial Journal*, vol. 53, no. 6, pp. e185-e197, 2016. https://doi.org/10.1597/15-199

[113] Weng N., Wang J., Li A., and Wang Y., "Two-Stream Temporal Convolutional Network for Dynamic Facial Attractiveness Prediction," *in Proceedings of the 25th International Conference on Pattern Recognition*, Milan, pp. 10026-10033, 2021. https://ieeexplore.ieee.org/abstract/document/9413333

[114] Whitehill J. and Movellan J., "Personalized Facial Attractiveness Prediction," *in Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition*, Amsterdam, pp. 1-7, 2008. https://ieeexplore.ieee.org/abstract/document/4813332

[115] Xie D., Liang L., Jin L., Xu J., and Li M., "SCUT-FBP: A Benchmark Dataset for Facial Beauty Perception," *in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Hong Kong, pp. 1821-1826, 2015. https://ieeexplore.ieee.org/document/7379451

[116] Xu J., Jin L., Liang L., Feng Z., and Xie D., "A New Humanlike Facial Attractiveness Predictor with Cascaded Fine-Tuning Deep Learning Model," *arXiv Preprint*, vol. arXiv:1511.02465, pp. 1-5, 2015. https://arxiv.org/abs/1511.02465

[117] Xu J., Jin L., Liang L., Feng Z., Xie D., and Mao H., "Facial Attractiveness Prediction Using Psychologically Inspired Convolutional Neural Network (PI-CNN)," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, pp. 1657-1661, 2017. https://ieeexplore.ieee.org/abstract/document/7952438

[118] Xu L., Fan H., and Xiang J., "Hierarchical Multi-Task Network for Race, Gender and Facial Attractiveness Recognition," *in Proceedings of the IEEE International Conference on Image Processing*, Taipei, pp. 3861-3865, 2019. https://ieeexplore.ieee.org/abstract/document/8803614

[119] Xu L., Xiang J., and Yuan X., "CRNet: Classification and Regression Neural Network for Facial Beauty Prediction," *in Proceedings of the Pacific Rim Conference on Multimedia*, Hefei, pp. 661-671, 2018. https://link.springer.com/chapter/10.1007/978-3-030-00764-5_61

[120] Xu L., Xiang J., and Yuan X., "Transferring Rich Deep Features for Facial Beauty Prediction," *arXiv Preprint*, vol. arXiv:1803.07253, pp. 1-6, 2018. https://arxiv.org/pdf/1803.07253

[121] Xu M., Chen F., Li L., Shen C., Lv P., and Zhou B., "Bio-Inspired Deep Attribute Learning towards Facial Aesthetic Prediction," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, 2021. https://ieeexplore.ieee.org/document/8454803

[122] Yalta K., Ozturk S., and Yetkin E., "Golden Ratio and the Heart: A Review of Divine Aesthetics," *International Journal of Cardiology*, vol. 214, pp. 107-112, 2016. DOI:10.1016/j.ijcard.2016.03.166

[123] Young P., *Rhinology and Facial Plastic Surgery*, Springer, 2009. https://link.springer.com/chapter/10.1007/978-3-540-74380-4_78

[124] Young P., Sinha U., Rice D., and Stucker F., "Circles of Prominence: A New Theory on Facial Aesthetics," *Archives of Facial Plastic Surgery*,

vol. 8, no. 4, pp. 263-267, 2006. https://doi.org/10.1001/archfaci.8.4.263

[125] Zhai Y., Huang Y., Xu Y., Gan J., Cao H., and Deng W., "Asian Female Facial Beauty Prediction Using Deep Neural Networks via Transfer Learning and Multi-Channel Feature Fusion," *IEEE Access*, vol. 8, pp. 56892-56907, 2020. https://ieeexplore.ieee.org/document/9035498

[126] Zhai Y., Huang Y., Xu Y., Zeng J., Yu F., and Gan J., "Benchmark of a Large Scale Database for Facial Beauty Prediction," *in Proceedings of the 1st International Conference on Intelligent Information Processing*, Wuhan, pp. 1-5, 2016. https://dl.acm.org/doi/10.1145/3028842.3028863

[127] Zhai Y., Yu C., Qin C., Zhou W., Ke Q., and Gan J., "Facial Beauty Prediction via Local Feature Fusion and Broad Learning System," *IEEE Access*, vol. 8, pp. 218444-218457, 2020. DOI:10.1109/ACCESS.2020.3032515

[128] Zhang B., Xiao X., and Lu G., "Facial Beauty Analysis Based on Features Prediction and Beautification Models," *Pattern Analysis and Applications*, vol. 21, no. 2, pp. 529-542, 2018. https://link.springer.com/article/10.1007/s10044-017-0647-2

[129] Zhang D., Zhao Q., and Chen F., "Quantitative Analysis of Human Facial Beauty Using Geometric Features," *Pattern Recognition*, vol. 44, no. 4, pp. 940-950, 2011. https://doi.org/10.1016/j.patcog.2010.10.013

[130] Zhang L., Zhang D., Sun M., and Chen F., "Facial Beauty Analysis Based on Geometric Feature: toward Attractiveness Assessment Application," *Expert Systems with Applications*, vol. 82, pp. 252-265, 2017. https://doi.org/10.1016/j.eswa.2017.04.021

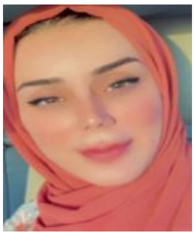
[131] Zhang Y., Wang X., Wang J., Zhang L., and Xiang Y., "Patterns of Eye Movements when Observers Judge Female Facial Attractiveness," *Frontiers in Psychology*, vol. 8, pp. 1-7, 2017. https://doi.org/10.3389/fpsyg.2017.01909

[132] Zhao J., Wu M., Zhou L., Wang X., and Jia J., "Cognitive Psychology-based Artificial Intelligence Review," *Frontiers in Neuroscience*, vol. 16, pp. 1-9, 2022. https://pmc.ncbi.nlm.nih.gov/articles/PMC9582153/

[133] Zhao J., Zhang M., He C., and Zuo K., "Data-Driven Research on the Matching Degree of Eyes, Eyebrows and Face Shapes," *Frontiers in Psychology*, vol. 10, pp. 1-11, 2019. https://doi.org/10.3389/fpsyg.2019.01466

[134] Zhao J., Zhang M., He C., Xie X., and Li J., "A Novel Facial Attractiveness Evaluation System Based on Face Shape, Facial Structure Features and Skin," *Cognitive Neurodynamics*, vol. 14, no. 5, pp. 643-656, 2020. https://link.springer.com/article/10.1007/s11571-020-09591-9

Ali Ibrahim was born in Duhok, Iraq. He Received B.Sc. from the University of Mosul, Iraq in 2006 in Computer Science, M.Sc. in Information Technology from the University of Dr. Baba Sahib, India. His area of interest is Deep Learning, Block Chain and Web Applications.



Jwan Saeed received a B.Sc. Degree in Computer Science from Mosul University in 2002 and the M.Sc. Degree in Computer Science from Duhok University, in 2010. Currently, she has a Ph.D. Degree and her research interests include the Applications of Artificial Intelligence, Computer Vision and Pattern.

Adnan Abdulazeez received a B.Sc. in Electrical and Electronic Engineering from the University of Technology in 1993, M.Sc. in Computer and Control Engineering from the University of Technology in 1998, and PhD. in Computer Engineering from the University of Mosul in 2007, and promoted to Professor in 2014. Dr. Adnan was the president of Duhok Polytechnic University from 2013 to 2022.