Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Facial Beauty Prediction Based on Deep Learning: A Review

Wahab Khalaf Arabo [1,2], Adnan Mohsin Abdulazeez [3]

wahab.arabo@uoz.edu.krd[1], wahab.arabo@auas.edu.krd[2], adnan.mohsin@dpu.edu.krd[3]

[1] Akre University of Applied Sciences – Technical College of Informatics-Akre- Department of Information Technology;
[2] Department of Computer Science, College of Science, University of Zakho, Duhok 42001, Iraq;
[3] Technical College of Engineering-Duhok, Duhok Polytechnic University, Duhok 42001, Iraq;

Article Information

Submitted : 31 Jan 2024
Reviewed: 2 Feb 2024
Accepted : 10 Feb 2024

Keywords

Beauty Prediction, Deep Learning, Facial Attractiveness, Convolutional Neural Networks.

Abstract

This review delves into Facial Beauty Prediction (FBP) using deep learning, specifically focusing on convolutional neural networks (CNNs). It synthesizes recent advancements in the field, examining diverse methodologies and key datasets like SCUT-FBP and SCUT-FBP5500. The review identifies trends in FBP research, including the evolution of deep learning models and the challenges of dataset biases and cultural specificity. The paper concludes by emphasizing the need for more inclusive and balanced datasets and suggests future research directions to enhance model fairness and address ethical implications.

A. Introduction

The concept of beauty has been a subject of fascination and contemplation throughout human history, influencing societal norms, cultural perceptions, and individual self-esteem. The perception of beauty, whether applied to a face or any comparable subject, varies among individuals, with diverse values assigned to beauty levels or ranks [1], [2]. In recent times, the intersection of technology and aesthetics has given rise to the intriguing field of beauty prediction. Beauty prediction involves the application of computational methods to assess and quantify perceived beauty, with a focus on understanding the intricate factors that contribute to this subjective notion [3]. This evolution towards a more analytical and data-driven approach to beauty has paved the way for the incorporation of advanced technologies like deep learning [4].

In the realm of artificial intelligence (AI), beauty prediction, particularly in the context of facial analysis, has emerged as a captivating area of research with potential applications in computer vision and image analysis. The development of algorithms capable of discerning and quantifying beauty from facial imagery paves the way for more intuitive and human-like machine interactions. Moreover, the application of beauty prediction extends to the marketing industry, where understanding consumer perceptions of beauty can significantly impact advertising strategies and product development [5]. Deep learning (DL), with its capacity to automatically learn hierarchical representations from data, has garnered increasing interest as a potent tool for beauty prediction tasks [6]. DL is increasingly utilized in various sectors such as healthcare for medical diagnostics and treatment planning [7], agriculture for crop yield predictions and operation of autonomous machinery[8], vocal AI and natural language processing for enhanced voice recognition and translation services[9], autonomous vehicle navigation [10], fraud detection[11], generative modelling [12], mobile multimedia [13], software engineering [14], big data [15], climate change analysis, and so on.

The growing interest in using DL for beauty prediction is rooted in the technology's remarkable success in various image-related tasks [16]. DL models, particularly convolutional neural networks (CNNs), have demonstrated unparalleled capabilities in recognizing complex patterns and features within images [17]. Leveraging these capabilities, researchers and practitioners aim to decode the intricate facets of facial aesthetics and overall beauty. This interest is further fuelled by the potential for creating applications that not only capture cultural and societal beauty ideals but also adapt to individual preferences, thereby personalizing the concept of beauty in computational systems [18].

As deep learning techniques continue to evolve and mature, the exploration of facial beauty prediction presents an exciting frontier with multifaceted implications. Beyond technological advancements, the incorporation of deep learning into facial beauty prediction raises profound questions about the nature of beauty, cultural influences on aesthetic preferences, and the ethical considerations surrounding algorithmic assessments of human appearance [19], [20] .

This review aims to delve into the burgeoning field of beauty prediction based on deep learning, examining its foundations, current state-of-the-art techniques, applications across diverse domains, and the challenges and opportunities that lie ahead. The rest of this paper is organized as follows: section 2 explores various steps in FBP field. Section 3 discusses the datasets employed, highlighting their roles and limitations. Section 4 includes a literature review and analysis, offering insights into the advancements and challenges in applying deep learning to facial beauty prediction. Finally, section 5 summarizes key findings and proposes future research directions.

B. Facial Beauty Prediction

Facial beauty predication and attractiveness computations form a cutting-edge research domain, utilizing automatic techniques to quantify facial beauty by analysing facial features and images. This innovative field seeks to establish a quantitative connection between perceived facial attractiveness and specific facial traits, necessitating the amalgamation of artificial intelligence, image processing, and pattern recognition. The entire process unfolds through a series of interrelated steps. It commences with the acquisition of a diverse database, incorporating data from public face databases, internet resources, and photographs. Subsequent pre-processing steps ensure the uniformity of face data through rectification, noise removal, and normalization[21], [22].

Following the initial data preparation, the construction of a reliable beauty score database becomes paramount. This step addresses challenges related to the scarcity of universally accepted true beauty scores by involving human raters in the evaluation process. Feature extraction and selection come next, identifying crucial attributes that guide the development of prediction models. These features, which can be local or holistic, include geometric, texture, and appearance-based attributes. The subsequent step involves the utilization of various techniques such as statistical methods, machine learning, and deep learning to build prediction models. Techniques like the k-nearest neighbour algorithm, support vector machine, and deep convolutional neural network are commonly employed. The final step in this comprehensive process is model validation, a critical phase where human-rated scores are rigorously compared with model-generated scores to ensure precision and reliability in predicting facial attractiveness [3], [19].

In the practical implementation of facial beauty prediction, two key elements emerge as crucial for the success and reliability of the models: datasets, and techniques. The elements are discussed in next sections.

C. FBP Datasets

In this section, we explore the common datasets utilized in the Facial Beauty Prediction (FBP) field. These datasets are crucial for developing and evaluating algorithms designed to assess facial attractiveness. Each dataset offers unique characteristics and challenges, contributing to a comprehensive understanding of facial beauty standards and perceptions.

1. **Multi-Modality Beauty (M2B) dataset**[23] is a comprehensive dataset specifically constructed for the study of female attractiveness across various modalities. It includes facial images, dressing images, and voice snippets of 1,240 females. This dataset is distinctive in its approach as it divides the data into two ethnic groups: Westerners and Easterners, with each group comprising 620 individuals. The uniqueness of the M2B dataset lies in its multi-modality approach, where previous datasets in attractiveness studies primarily focused on single modalities, such as only facial images.

   The attractiveness scores within the M2B dataset are annotated by human subjects, employing a scoring system that typically ranges from 1 to 10. This scoring is based on absolute value ratings, a popular method in such studies, where a user rates a single image or voice snippet on a scale of 1 to 10. This form of rating is crucial as it requires each image or snippet to be rated by many users to achieve a representative distribution of scores. The dataset's design aims to capture a more holistic view of attractiveness, considering that attractiveness is a multi-faceted concept influenced by various factors, including facial features, fashion, and voice.

   Furthermore, the M2B dataset, with its unique focus on multi-modality cues and diverse cultural representation, marks a significant advancement in the field of attractiveness research, offering new avenues for exploring how beauty is perceived and evaluated across different cultures and modalities.

2. **SCUT-FBP dataset** [24] is a specialized collection comprising 500 portraits of Asian females, designed for research in automatic facial beauty perception. It provides a benchmark for evaluating facial attractiveness prediction methods. Each image in the dataset is a high-resolution, front-facing portrait with a neutral expression and a simple background, rated for attractiveness. The scores, ranging from 1 to 5, were given by an average of 70 ratters per picture through a web-based platform, guaranteeing a wide-ranging and varied evaluation. The dataset had been used in many fields such as analysis of image and video [25] , study label distribution [26], [27]. The dataset has been instrumental in advancing understanding in computer vision and psychology, particularly in the study of facial attractiveness. Samples of this dataset is presented in Figure 1.

Figure 1. Samples of SCUT-FBP dataset [24]

3. **Large-Scale Asian Female Beauty (LSFBD) dataset** [28], this dataset represents a significant advancement in the field of facial beauty analysis, addressing the challenges posed by the subjective nature of beauty evaluation. It comprises a substantial collection of 20,000 labelled images, evenly distributed between unconstrained male and female subjects. This extensive dataset is carefully curated to minimize biases and external influences, with each image validated through a well-designed rating system that incorporates average scores and standard deviations. The beauty scores are assigned on a discrete integer scale ranging from 1 (extremely unattractive) to 5 (most attractive), providing a broad spectrum for beauty analysis. The LSFBD stands out for its size and the meticulous approach to rating, offering a more comprehensive and nuanced understanding of facial beauty.

   The rating process of the LSFBD involved approximately 200 volunteers, predominantly teachers and students aged between 20 and 35 years, reflecting a diverse and relevant demographic for contemporary beauty standards. The images were divided into groups, with each group being rated by an equal number of male and female ratters, ensuring a balanced perspective in the evaluations. This methodical approach to data collection and rating not only enhances the reliability of the dataset but also sets a new standard in the field for analysing and understanding facial beauty. Figure 2 contains some samples of this dataset.

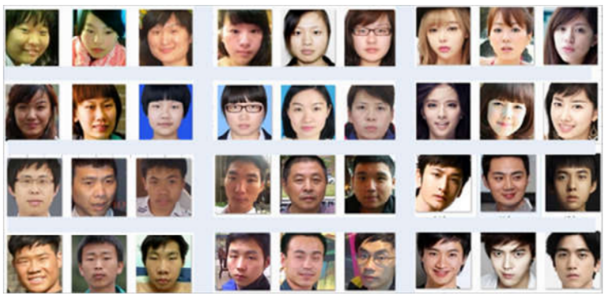

Figure 2. Samples of LSFBD[28]

4. **SCUT-FBP5500** dataset [29] is a comprehensive and diverse collection specifically designed for the study and prediction of facial beauty. It comprises 5500 images of frontal, unclouded faces with neutral expressions, capturing a wide age range from 15 to 60 years. The dataset is meticulously categorized into four distinct subsets based on race and gender, encompassing 2000 Asian females, 2000 Asian males, 750 Caucasian females, and 750 Caucasian males. This diverse demographic representation aids in creating a more holistic and inclusive approach to facial beauty analysis.
   The dataset's primary purpose is to facilitate the development and evaluation of computational models for facial beauty prediction. The images were labelled with beauty ratings between 1 and 5. By offering a mix of images from different races and genders, the SCUT-FBP5500 dataset stands as a valuable resource for researchers and developers in fields such as cosmetic recommendation, aesthetic analysis, and facial recognition technology, enabling them to build and test algorithms that are more representative of diverse populations. Figure 3 presents some samples of this dataset.

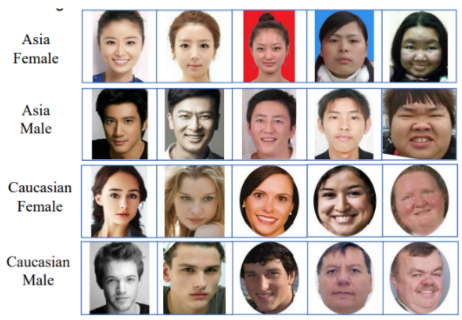

Figure3. Samples of SCUT-FBP5500

5. **ECCV HotOrNot Dataset**, it presents a significant challenge for researchers in the domain of facial beauty prediction due to its unique composition and inherent complexities. Comprising 2056 facial images sourced from the internet, this dataset introduces a diverse range of real-world conditions such as varied postures, cluttered backgrounds, inconsistent illumination, low-resolution images, and unaligned faces. These factors collectively contribute to the difficulty of accurately predicting facial beauty, thereby making this dataset an invaluable benchmark for testing and improving the robustness of ML algorithms. Each image within the dataset is meticulously labelled with a specific score, facilitating a comprehensive approach to training and evaluation. The dataset is thoughtfully partitioned into five distinct training and test sets, with the final evaluation metric being the average of the results obtained from a 5-fold cross-validation process. This methodological approach ensures a more reliable and generalized

assessment of algorithmic performance, underscoring the dataset's significance in advancing computer vision and machine learning [19], [30]. This dataset had utilized in experiments on transferring rich features [31] , designing image recognition CAPTCHAs [32]

6. **Geometric Facial Beauty** (GFB) dataset[33], the (GFB) dataset is a unique collection aimed at analysing facial attractiveness through geometric features. It includes 4905 male and 4510 female samples, divided into labelled and unlabelled categories. Labelled samples encompass 195 attractive and 2005 unattractive male faces, and 191 attractive and 1869 unattractive female faces. Additionally, it contains 2705 unlabelled male and 2450 female samples, with a significant portion sourced from the Shanghai database. This dataset, with its focus on geometric characteristics, provides a valuable resource for developing machine learning models in facial beauty assessment.
7. **Gray dataset** [34], this collection contains 2,056 images of frontal female faces, aged 18-40, with few restrictions on ethnicity, lighting, pose, or expression. Most images were cropped from low-quality photos taken by cell-phone cameras. The dataset is significantly larger than previous studies, being 20 times larger. The unique aspect of this dataset is the minimal restrictions imposed, adding to the challenge of the learning task. For labelling, a pairwise rating method was used, where ratters chose the more attractive face from a pair, to eventually develop a global absolute score for each image. This approach aims to overcome the limitations of previous studies that used smaller, more controlled datasets
8. **CelebA dataset** [35] is a large collection, composed of more than 200,000 images of celebrities, each with dimensions of 178×218. It is predominantly utilized for the classification of facial beauty. The dataset's significant inclusion of natural images of both males and females, along with 40 binary labels related to beauty, contributes to its utility. Moreover, its binary categorization of facial attractiveness into 'attractive' and 'unattractive' labels simplifies the computational process. This dataset is been used by many researchers for different purposes such as study: the data consistency and accuracy [36], unusual effectiveness with GAN [37] , imager generation [38] .

D. Deep Learning in FBP

In the last four years, the field of FBP has garnered substantial attention within the academic landscape, spanning diverse disciplines. Deep learning models have been prominently featured in FBP, attaining a noteworthy status. This section delves into a comprehensive examination of recent studies pertaining to FBP through the utilization of Deep Learning models. The first subsection provides a chronological synthesis of recent investigations, elucidating the evolving landscape, while the second subsection articulates our analytical perspectives and facilitates a discussion on the subject matter.

- **Literature Review**

Zhai et al. [39] designed a multiscale CNN model, called BeautyNet, for unconstrained facial beauty prediction. The model is composed of 11 convolution layers involving 2671456 parameters. They adopted max-feature-map (MFM) activation function instead of the common one, such as ReLU. The new activation function helped to obtain more sparse gradients and compact representation simultaneously, leading to lightened model and faster training. BeautyNet overcomes the overfitting issues and datasets limitations using transfer learning concepts, where the net first run to optimize parameters on CASIA-WebFace and LFW databases. The female iamges of LSFBD dataset were classified by BeautyNet recording 64.84% accuracy rate and 80.20 Pearson's correlation coefficient

Vahdati and Suen [40] provided a new framework for analysing the attractiveness of female faces based on transferring the learning of Ms-Celeb-1M and VGGFace2 models, which have been trained on relatively similar datasets for recognizing female faces. They combine the results of several base models to predict the attractiveness of a face, these models are based on Lasso, SVR (linear) and Ridge regression methods. SCUT-FBP and SCUT-FBP 5500 benchmark datasets are included in this study, the achieved results on these datasets 0.89 and 0.91 respectively in prediction correlations term.

In [41], Chen and Deng proposed a novel deep architecture called Deep Adaptive Label Distribution Learning (DALDL). In DALDL, the discrete label distribution of possible ratings was used rather than single label to supervise the learning process of facial attractiveness prediction. This distribution was automatically updated during the training process. To handle imbalances in the sample distribution within the dataset, DALDL utilizes a weighted Euclidean loss function. This function assigns different weights to different labels based on the number of samples, ensuring a more balanced and fair learning process. Both SCUT-FBP and SCUT-FBP5500, 5-fold cross validation are used with architecture VGG_DALDL_WeightedLoss, scoring better results than competitors in term of PC, MAE, and RMSE.

GAN et al. [42] utilized a multi-input, multi-task beauty network, also known as 2M BeautyNet, for the purpose of predicting facial beauty through transfer learning. This method employs auxiliary data from related tasks to address the primary task, effectively minimizing overfitting risks. In their framework, the secondary or auxiliary task involved gender recognition, while the primary focus was on beauty prediction. The use of multi-task training in this context is beneficial for enhancing the performance of Facial Beauty Prediction (FBP), as it involves the automatic learning of multi-task loss weights. Additionally, the original Softmax classifier in their system was replaced with a random forest algorithm for improved results. The LSFBD and SCUT-FBP5500 are used in this study with the accuracy of FBP up to 68.23%.

Another deep learning approach for predicting facial beauty presented [6] by Cao ae al., their focus was on creating a model that aligns with human opinions on attractiveness. Key innovations include a residual-in-residual (RIR) structure for deeper network learning and a spatial-wise and channel-wise attention (SCA) mechanism to enhance feature representation. The model outperforms existing CNN-based methods in aligning with human assessments of beauty. The research also discussed the influence of network design and attention mechanisms on performance, comparing the proposed method with established approaches like AlexNet and ResNeXt using SCUT-FBP5500 dataset. The study emphasizes the importance of efficient information transmission pathways and feature correlation understanding in deep learning models for facial beauty prediction.

Zhai et al. [43] discussed an innovative approach for facial beauty prediction (FBP) using artificial intelligence. The method involves a fast-training FBP model based on local feature fusion and a broad learning system (BLS). The technique employs two-dimensional principal component analysis (2DPCA) to reduce image dimensionality, thus minimizing redundancy. The local feature fusion method is used to extract advanced features, reducing the impact of unstable illumination, individual differences, and various postures. The fusion of local texture features and BLS significantly improves operational efficiency and precision of the FBP model. The proposed approach is tested on a large-scale Asian female beauty database, demonstrating its effectiveness in achieving high accuracy and speed in FBP tasks.

Zhai et al. [44] presented a novel CNN model for predicting Asian female facial beauty. This model, enhanced by Softmax-MSE loss function and a double activation layer, utilized the Large-Scale Asian Female Beauty Dataset (LSAFBD) of 20,000 images and the CASIA-WebFace dataset for pre-training. Achieving a 64.85% rank-1 recognition rate and a 0.8829 Pearson correlation coefficient, the model outperforms traditional methods, despite its reliance on large-scale datasets and complex feature extraction processes. This work significantly contributes to the field by establishing the LSAFBD and innovatively applying transfer learning and feature fusion techniques.

Xu and Xiang [45] introduced ComboLoss, a new loss function for facial attractiveness analysis using SEResNeXt50 networks. It is evaluated on SCUT-FBP, HotOrNot, and SCUT-FBP5500 datasets, showing improved performance over previous methods. They highlight the importance of this loss function in enhancing the learning process without adding complexity, despite limitations like not exploring advanced discretization methods or additional datasets for pretraining. This research significantly contributes to the field by systematically comparing loss functions and setting new performance benchmarks.

Lebedeva et al.[46] adapted the Facenet Convolutional Neural Network, originally designed for face recognition, to predict facial attractiveness using Support Vector Regression (SVR). They experiments were performed on the Gray and SCUT-FBP5500 datasets, encompassing a diverse range of facial images, underscored its efficacy.

Bougourzi et al. [47] introduced an approach in FBP using deep learning. they presented the REX-INCEP architecture, combining two pre-trained networks for handling complex facial features. The study introduced dynamic robust loss functions like ParamSmoothL1, Huber, and Tukey, enhancing training behaviour. It also employs ensemble regression, integrating different CNN models for improved accuracy. Evaluated on the SCUT-FBP5500 database, this method shows significant advancements over existing FBP methods. This research contributes to the field by offering innovative architecture and dynamic loss functions for more precise FBP.

Luthfi et al.[48] presented a study on implementing FBP on mobile devices using Convolutional Neural Networks (CNNs). It emphasizes the use of MobileNetV2 for its efficiency and performance, validated on the SCUT-FBP5500 dataset, which contains 5500 facial images rated for beauty. The model demonstrates a good balance between accuracy and file size, making it viable for mobile applications. Its contribution lies in advancing FBP applications for mobile platforms, particularly in makeup and cosmetic industries. They had chosen the MobileNetV2 after comparing its performance with each of EfficientNetB0, VGG-16, ShuffleNet, MobileNet, and Inception.

Saeed et al.[49] introduced a new deep learning model, the FIAC-Net, for facial image attractiveness assessment. It aimed to objectively model beauty using deep convolutional neural networks (DCNNs) and addresses challenges like dataset scarcity and computational demands. They proposed a lightweight DCNN architecture with fewer parameters, enabling it to function effectively on devices with limited hardware capabilities. The experiments were performed by using the CelebA, SCUT-FBP, and SCUT-FBP5500 datasets, employing data augmentation and soft labels for training. The FIAC-Net demonstrates significant improvements in facial attractiveness classification over existing methods, particularly in terms of computational efficiency and adaptability to different datasets.

Saeed et al.[50] proposed to ensemble three regression-loss functions, namely L1, L2, and Log-cosh, and subsequently averaging them to create a new composite cost function. They incorporating their new loss function with three pretrained CNNs, namely AlexNet, VGG16-Net, and FIACNet. Furthermore, they tested these pretrained models on three FBP benchmarks, i.e., SCUT-FBP, SCUT-FBP5500, and MEBeauty. Notably, their results demonstrated exceptional performance when incorporating the newly loss function with FIAC-Net, showing remarkable outcomes across all datasets.

Yang et al. [51] aimed to develop a model for assessing facial attractiveness. It combined transfer learning, convolutional neural networks (CNNs), Xception, and attention mechanisms using datasets from Chang Gung Memorial Hospital and SCUT-FBP5500. The model used MAPE and RMSE for evaluation. It achieved a best RMSE of 0.50 and 18.5% average error in MAPE. A web page was developed to visualize the predictive model. The research offers a systematic approach to facial beauty assessment, potentially aiding in medical plastic surgery and related fields.

Gan et al. [52] introduced a new approach to FBP by integrating transfer learning with a Broad Learning System (BLS). They proposed a method combining EfficientNets-based transfer learning for feature extraction and BLS for rapid model training. Two strategies, E-BLS and ER-BLS, are developed, showing improvements in accuracy and training speed compared to existing methods. The approach demonstrates applicability in pattern recognition, object detection, and image classification. Extensive experiments validate the effectiveness and efficiency of these methods on both SCUT-FBP 5500 and LSAFBD (LSFBD) datasets.

Saeed et al. [53] developed an ensemble approach using three Deep Convolutional Neural Networks - Beauty-Reg-Net, VGG16, and AlexNet - for automatic FBP. The final regression value is derived by averaging the scores from each of the sub-models, ensuring equal contribution from each model to the ensemble decision. This method enhances the model's performance by combining the strengths of individual networks. The study employs three dedicated FBP datasets: SCUT-FBP, SCUT-FBP 5500, and MEBeauty, to validate the effectiveness of the proposed model. Both online and offline data augmentation were employed to overcome misbalancing in datasets, as they reflect the natural skewness in facial beauty data.

Zejmo et al. [54] explored facial attractiveness using three machine learning models. It employs the SCUT-FBP5500 dataset for training, and the Face Research Lab London Set for testing, featuring 102 adult faces. The first model evaluates dominant background colors in photographs, employing clustering to form training vectors. The second model applies deep learning neural networks, including MobileNetV2, VGG19, ResNet50V2, and Xception, trained on the ImageNet dataset to extract facial features for attractiveness prediction. The third model assesses facial proportions by measuring distances between key facial points, using these measurements in algorithms like RandomForestRegressor and LinearSVR for attractiveness predictions. The regression results from the three models are combined to determine the final attractiveness score of a face.

Bougouzi et al. [55] presented an innovative approach for facial beauty estimation using deep learning. They introduced a dual CNN architecture, 2B-IncRex, and dynamic robust loss functions (ParamSmoothL1, Huber, Tukey) to handle complex facial features. Additionally, they proposed an ensemble regression combining five different regressors. This method, tested on SCUT-FBP5500 and KDEF-PT datasets, demonstrates superior accuracy and robustness in face beauty assessment, outperforming existing methods. In the study, the effectiveness of dynamic robust losses in creating more flexible and accurate estimators for facial beauty prediction are been highlighted.

- **Literature Analysis**

Through analyzing the studies in previous section, we notice that the landscape of facial beauty prediction research has undergone notable shifts in both model architectures and methodological approaches. In the early stages, researchers predominantly centered their efforts on refining convolutional neural network

(CNN) models. Notably, early models such as BeautyNet were characterized by a focus on architecture enhancements, including the adoption of innovative activation functions like max-feature-map (MFM) to mitigate overfitting issues. Additionally, the integration of transfer learning from established databases like CASIA-WebFace and LFW underscored an initial emphasis on foundational model development and generalization.

As the field progressed chronologically, there emerged a trend towards diversification and sophistication in methodologies. Researchers introduced novel loss functions, such as ComboLoss and dynamic robust losses, aiming to enhance the learning process and improve overall model performance. Ensemble methods gained prominence, with studies leveraging the combination of multiple models or regressors to achieve heightened accuracy. The exploration of attention mechanisms, mobile applications, and innovative architectures, like 2B-IncRex, demonstrated a broader recognition of the need for adaptable and efficient models. This evolution in research direction reflects a transition from foundational model refinement to a more nuanced exploration of advanced techniques, emphasizing computational.

Tables 1 provides a summary of the studies discussed in the preceding section, detailing the models or methods employed, as well as the datasets utilized in each study. Furthermore, it includes a critical analysis of some limitations identified, based on our evaluation.

Table 1. Summary of the recent studies

<table>
<tr><th>Ref</th><th>Model / Method summary</th><th>Limitation (if noted)</th><th>Datasets</th><th colspan="3">Results</th></tr>
<tr><td rowspan="3">Zhai et al. [39]</td><td rowspan="3">Name: BeautyNet,
11 Con. Layers, 2671456 parameters
New activation function MFM
Use Transfer learning methodology</td><td rowspan="3">Limited data usage so it accurately cannot represent global beauty standards</td><td rowspan="3">female part of LSFBD</td><td>+ Transfer</td><td>ACC</td><td>PC</td></tr>
<tr><td>No</td><td>64.84</td><td>80.20</td></tr>
<tr><td>Yes</td><td>67.48</td><td>83.54</td></tr>
<tr><td rowspan="4">Vahdati & Suen [40]</td><td rowspan="4">Transfer learning of Ms-Celeb-1M and VGGFace2 models
Use combination of Lasso, SVR (linear) and Ridge regression methods to predict final result</td><td rowspan="4">Depend only on transfer learning to compute facial attractiveness.</td><td rowspan="2">SCUT-FBP</td><td>PC</td><td>MAE</td><td>RMSE</td></tr>
<tr><td>0.8898</td><td>0.2409</td><td>0.3105</td></tr>
<tr><td>SCUT-FBP 5500 Asian females</td><td>0.9141</td><td>0.2196</td><td>0.2895</td></tr>
<tr><td>SCUT-FBP 5500 Caucasian females</td><td>0.9112</td><td>0.2304</td><td>0.2951</td></tr>
<tr><td rowspan="4">Chen & Deng [41]</td><td rowspan="4">Name: DALDL
VGG+DALDL+WeightedLoss
Use discrete label distribution
balanced and fair learning process using WeightedLoss</td><td rowspan="4">Complexity of Implementation,
The initial label distribution relies on certain assumptions</td><td rowspan="2">SCUT-FBP</td><td>PC</td><td>MAE</td><td>RMSE</td></tr>
<tr><td>0.903</td><td>0.227</td><td>0.312</td></tr>
<tr><td>SCUT-FBP 5500</td><td>0.915</td><td>0.210</td><td>0.278</td></tr>
<tr><td>CelebA</td><td colspan="3">Accuracy: 83.9 %</td></tr>
<tr><td>GAN et al. [42]</td><td>Name: 2M BeautyNet
two networks, 1st for gender recognition, and 2nd for beauty prediction</td><td>It needs a more robust and powerful multi-input multi-task network
some changes in dataset</td><td>SCUT-FBP 5500</td><td colspan="3">Accuracy: 68.23%</td></tr>
</table>

	Transfer learning used in each net is based on VGG16 14798023 parameters	lead to unfair comparison					
Cao ae al [6]	residual-in-residual (RIR) structure for deeper network learning and a spatial-wise and channel-wise attention (SCA) mechanism to enhance feature representation	Students within a specific cultural context evaluate training labels, and there is a lack of consensus among them. The selected trained images feature individuals from both Asian and Caucasian backgrounds, potentially introducing bias in terms of diversity	SCUT-FBP 5500	PC	MAE	RMSE	
				0.9003	0.2287	0.3014	
Zhai et al. [43]	Employing broad learning and local feature fusion, combined with a texture-based approach and 2DPCA to enhance model accuracy and efficiency in analysing facial characteristics	The cost-sensitive facial beauty dilemma needs to be solved in the future	LSAFBD	Accuracy: 58.97%			
Zhai et al. [44]	Enhanced CNN model by Softmax-MSE loss function and a double activation layer	Suitable for western faces may not apply to the east	LSAFBD	Accuracy	PC		
				64.85%	0.8829		
Xu and Xiang[45]	ComboLoss New loss function on SEResNeXt50 network	The study doesn't deeply explore advanced discretization methods for the attractiveness scores	SCUT-FBP HotOrNot	PC	MAE	RMSE	
				0.9090			
			SCUT-FBP 5500	0.50			
			SCUT-FBP	0.9199	0.2050	0.2704	
Lebedeva	Facenet+ SVR	The exploration of various	SCUT-FBP	Data operation	PC	MAE	RMSE

et al.[46]	Facenet Convolutional Neural Network used for features extraction, then fed extracted features to Support Vector Regression (SVR).	data preprocessing techniques may have led to inconsistencies in predicting facial attractiveness across different datasets	5500	Original Images	0.8834	0.2415	0.3221
				Face Cropping	0.8728	0.2565	0.2319
				Original Photos + Face Alignment	0.8941	0.2365	0.3101
				Face Cropping + Face Alignment	0.8807	0.2455	0.3252
			Gray	Original Images	0.4638	0.1417	0.9467
				Face Cropping	0.4704	1.1337	0.9290
				Original Photos + Face Alignment	0.4665	1.1378	0.9371
				Face Cropping + Face Alignment	0.4732	1.1315	0.9254
Bougourzi et al. [47]	REX-INCEP architecture, combining 2 pre-trained networks different loss functions used to enhance training	the intervals for the parameters of the loss functions are static	SCUT-FBP 5500	PC	MAE	RMSE	
				0.9159	0.2071	0.2739	
Luthfi et al.[48]	Implement MobileNetV2 as FBP	The study acknowledges limitations like the large size of alternative models and occasional face detection failures	SCUT-FBP 5500	PC : 0.7893			
Saeed et al.[49]	Name: FIAC-Net a lightweight DCNN architecture with fewer parameters	the SCUT-FBP and SCUT-FBP5500 datasets are frequently employed in previous studies for regression tasks. This necessitates a fair comparison to accurately evaluate the FIAC-Net's performance against	SCUT-FBP 5500	Accuracy			
				85.9%			
			CelebA	82%			
			SCUT-FBP	89%			

		existing methods					
Saeed et al.[50]	Ensemble L1, L2, and Log-cosh into on loss function Tested with AlexNet, VGG16-Net, and FIAC-Net	The three loss functions have same weights, it may be better to use dynamic weight for each, extracting such weights can be challenge and add valuable contribution	SCUT-FBP 5500	Model	PC	MAE	RMSE
				AlexNet	0.9140886	0.2448064	0.3096576
				VGG16-Net	0.93092	0.2183198	0.2791592
				FIAC-NET	0.9305098	0.2028174	0.2614154
			SCUT-FBP	AlexNet	0.903758	0.26348	0.348266
				VGG16-Net	0.905851	0.222954	0.292028
				FIAC-NET	0.9100582	0.185949	0.259156
			MEBeauty	AlexNet	0.8976712	0.476394	0.6097078
				VGG16-Net	0.907883	0.512113	0.636318
				FIAC-NET	0.925977	0.426317	0.536646
Yang et al. [51]	Combined transfer learning, CNNs, Xception, and attention mechanisms	Using SCUT-FBP5500 with pre-trained CNN not enough to build discissions in medical plastic surgery. The combination of several techniques increased the number of parameters leading to mor complexity	SCUT-FBP 5500	MAPE		RMSE	
				18.5% average error		0.50	
Gan et al. [52]	EfficientNets-based transfer learning with BLS Two strategies named, E-BLS and ER-BLS	Dependence on Hyperparameters Due to the combination of multiple models the learning time is increased	SCUT-FBP 5500	Accuracy			
				73.13%			
			LSAFBD	62.13%			
Saeed et al. [53]	Ensemble approach using Beauty-Reg-Net, VGG16, and AlexNet	The study's approach of equally weighting outputs from different, variably complex models in an ensemble could be a limitation. This method may not accurately reflect	SCUT-FBP 5500	PC	MAE	RMSE	
				0.886	0.242	0.320	
			SCUT-FBP	0.879	0.226	0.33	

<table>
<tr><td></td><td></td><td>the varying contributions and accuracies of each model, potentially leading to inconsistencies in the final predictions despite the benefit of diverse model perspectives.</td><td>MEBeauty</td><td colspan="2">0.888</td><td>0.365</td><td>0.6</td></tr>
<tr><td rowspan="4">Zejmo et al. [54]</td><td rowspan="4">Assembling 3 models dependent on: (1) dominant background color, (2) deep learning neural networks, and (3) assesses facial proportions by measuring distances between key facial points</td><td rowspan="4">A notable limitation of the study is the small size of the test dataset. Additionally, the methodology involves summing the outputs of the three models without assigning different weights. This equal contribution approach might not accurately reflect the varying significance or effectiveness of each model in determining facial attractiveness</td><td rowspan="4">training on SCUT-FBP5500 and test on Face Research Lab London dataset</td><td>Model</td><td>R2</td><td>MAE</td><td>RMSE</td></tr>
<tr><td>RandomForestRegressor</td><td>0.233</td><td>2.344</td><td>0.613</td></tr>
<tr><td>LinearSVR</td><td>0.146</td><td>2.433</td><td>0.647</td></tr>
<tr><td>SDGRegressor</td><td>0.182</td><td>2.509</td><td>0.634</td></tr>
<tr><td rowspan="2">Bougouzi et al. [55]</td><td rowspan="2">Dual CNN architecture, 2B-IncRex with dynamic robust loss functions (ParamSmoothL1, Huber, Tukey)</td><td rowspan="2">the intervals for the parameters of the loss functions are static</td><td rowspan="2">SCUT-FBP 5500</td><td colspan="2">PC</td><td>MAE</td><td>RMSE</td></tr>
<tr><td colspan="2">0.9150</td><td>0.2085</td><td>0.2744</td></tr>
</table>

The limitations in facial beauty prediction researches, as detailed in table, are extensive and complex. They encompass challenges like limited and non-diverse datasets, which can skew models' understanding of global beauty standards. The prevalent use of transfer learning, while resourceful, raises concerns about the models' ability to predict beauty independently. The implementation complexities, cultural biases in training data, and dataset variability further complicate the matter. These issues underscore the need for more inclusive, fair, and robust models capable of accurately representing diverse beauty standards. The pursuit of universally applicable methodologies is crucial, especially considering the subjective nature of beauty and the need to ensure fairness across different cultural contexts. This highlights both the advancements in facial beauty prediction and the persistent challenges that need to be addressed to enhance the accuracy, fairness, and generalizability of these models.

On the contrary, it is observed that the SCUT-FBP, SCUT-FBP5500, and MEBeauty datasets have prominently emerged as preferred selections, functioning as principal benchmarks in numerous investigations. However, the imbalanced nature of these datasets presents a noteworthy concern that could potentially compromise the validity of outcomes in a majority of studies. Notably, the preponderance of beauty scores within the [2-2.9] range out of a maximum score of 5 in both SCUT-FBP and SCUT-FBP5500 datasets, and the dominance of beauty scores within the [4.0-4.9] range out of a maximum score of 10 in the MEBeauty dataset introduces a skewness in the distribution of beauty evaluations. This disparity in score distributions warrant careful consideration when interpreting and generalizing the findings derived from these datasets, as it may impact the overall representativeness and generalizability of the models trained and evaluated on such imbalanced data.

In term of result, the table 1 shows that the researchers often employ a regression approach, leveraging datasets where FBP is numerically scored. The primary evaluation metrics used to assess the effectiveness of different models include Pearson Correlation (PC), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). Among the datasets, the SCUT-FBP5500 dataset emerges as a commonly used choice, offering insights into the complexity of beauty prediction.

When evaluating MAE [47] stand out with an exceptional performance, achieving a remarkably low MAE of 0.2071. In contrast, [46] report a higher MAE of 0.2565, primarily attributed to their utilization of Face Cropping in their approach. The Pearson Correlation (PC) scores exhibit a diverse range, spanning from 0.5 to 0.93. Yang et al. [51] achieve the highest PC score using the VGG16-Net model with ensemble cost functions, showcasing a strong correlation between predicted and actual beauty ratings. Conversely, the worst PC score is reported by [45], indicating limitations in the model's predictive power. Root Mean Square Error (RMSE) serves as an indicator of prediction accuracy. [46] surpass in this regard with a notable RMSE score of 0.23, primarily attributed to their innovative use of Face Cropping. However, the worst RMSE score, reaching 0.647, is observed when combining classical machine learning with deep learning, highlighting the complexities of accuracy in such an approach.

Within the SCUT-FBP dataset, notable variations in MAE, PC, and RMSE scores are observed. The best MAE score of 0.185949 is achieved by [50] using the FIAC-NET model, reflecting the precision of their predictions. In contrast, the worst MAE score, 0.26348, is reported by [53] . Turning to Pearson Correlation (PC), the SCUT-FBP dataset presents a spectrum of performance. The highest PC score of 0.9100582 is achieved by [50] using the FIAC-NET model, signifying a robust correlation between predicted and actual beauty ratings. Conversely, the lowest PC score of 0.8898 is reported by [40] .In terms of RMSE within the SCUT-FBP dataset, [46] lead the way with an RMSE of approximately 0.25, attributed to their innovative use of Face Cropping. In contrast, the worst RMSE score is 0.312, reported by [41].

E. Conclusion and Recommendations

This review of facial beauty prediction (FBP) using deep learning techniques has illuminated the dynamic interplay between technological advancement and the nuanced concept of beauty. The field, though still evolving, has made notable strides in employing deep learning for the intricate task of beauty assessment.

Throughout the review, it has become evident that deep learning models, particularly convolutional neural networks (CNNs), have greatly enhanced our ability to analyze and predict facial attractiveness. Innovations such as BeautyNet, DALDL, and various ensemble methods have demonstrated improved accuracy and efficiency in beauty prediction. However, these advancements are not without their limitations. The complexity of model implementation, dependency on specific datasets, and the inherent biases in the training data have emerged as significant challenges. These limitations underscore the need for further refinement in model architectures and learning strategies.

The datasets used in FBP, such as SCUT-FBP, LSFBD, and SCUT-FBP5500, have played a critical role in the development of these models. However, the skewed distribution of beauty scores and the cultural specificity inherent in these datasets highlight the need for more balanced and diverse data collection methods. The pursuit of models that accurately represent global beauty standards and are sensitive to cultural diversity is imperative for the future of FBP research.

In conclusion, while deep learning has brought new perspectives to the understanding of facial beauty, the journey towards creating unbiased, universally applicable, and ethically sound FBP models is ongoing. The field stands at a crossroads, with opportunities to delve deeper into the fusion of technology and aesthetics, while also facing the challenge of addressing the ethical and cultural implications of its advancements. Future research should focus on enhancing the diversity and representativeness of datasets, improving the fairness and inclusivity of models, and continually questioning and refining the ethical boundaries of using technology to assess human beauty.

F. References

[1] H. Knight and O. Keith, "Ranking facial attractiveness," Eur J Orthod, vol. 27, no. 4, pp. 340–348, Aug. 2005, doi: 10.1093/ejo/cji042.
[2] M. Hönn and G. Göz, "The Ideal of Facial Beauty: A Review," Journal of Orofacial Orthopedics / Fortschritte der Kieferorthopädie, vol. 68, no. 1, pp. 6–16, Jan. 2007, doi: 10.1007/s00056-007-0604-6.
[3] M. K. Moridani, N. Jamiee, and S. Saghafi, "Human-like evaluation by facial attractiveness intelligent machine," International Journal of Cognitive Computing in Engineering, vol. 4, pp. 160–169, Jun. 2023, doi: 10.1016/j.ijcce.2023.04.001.
[4] D. Laurinavičius, R. Maskeliūnas, and R. Damaševičius, "Improvement of Facial Beauty Prediction Using Artificial Human Faces Generated by Generative Adversarial Network," Cognit Comput, vol. 15, no. 3, pp. 998–1015, May 2023, doi: 10.1007/s12559-023-10117-8.
[5] J. Gan, L. Zhou, and Y. Zhai, "A study for facial beauty prediction model," in 2015 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), IEEE, Jul. 2015, pp. 8–13. doi: 10.1109/ICWAPR.2015.7295918.
[6] K. Cao, K. nam Choi, H. Jung, and L. Duan, "Deep learning for facial beauty prediction," Information (Switzerland), vol. 11, no. 8, Aug. 2020, doi: 10.3390/INFO11080391.
[7] S. Shamshirband, M. Fathi, A. Dehzangi, A. T. Chronopoulos, and H. Alinejad-Rokny, "A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues," J Biomed Inform, vol. 113, p. 103627, Jan. 2021, doi: 10.1016/j.jbi.2020.103627.
[8] I. Attri, L. K. Awasthi, T. P. Sharma, and P. Rathee, "A review of deep learning techniques used in agriculture," Ecol Inform, vol. 77, p. 102217, Nov. 2023, doi: 10.1016/j.ecoinf.2023.102217.
[9] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," Multimed Tools Appl, vol. 82, no. 3, pp. 3713–3744, Jan. 2023, doi: 10.1007/s11042-022-13428-4.
[10] J. Fayyad, M. A. Jaradat, D. Gruyer, and H. Najjaran, "Deep Learning Sensor Fusion for Autonomous Vehicle Perception and Localization: A Review," Sensors, vol. 20, no. 15, p. 4220, Jul. 2020, doi: 10.3390/s20154220.
[11] A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, and A. Imine, "Credit card fraud detection in the era of disruptive technologies: A systematic review," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 1, pp. 145–174, Jan. 2023, doi: 10.1016/j.jksuci.2022.11.008.
[12] S. Bond-Taylor, A. Leach, Y. Long, and C. G. Willcocks, "Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models," IEEE Trans Pattern Anal Mach Intell, vol. 44, no. 11, pp. 7327–7347, Nov. 2022, doi: 10.1109/TPAMI.2021.3116668.
[13] K. Ota, M. S. Dao, V. Mezaris, and F. G. B. De Natale, "Deep Learning for Mobile Multimedia," ACM Transactions on Multimedia Computing,

Communications, and Applications, vol. 13, no. 3s, pp. 1–22, Aug. 2017, doi: 10.1145/3092831.
[14] Y. Yang, X. Xia, D. Lo, and J. Grundy, "A survey on deep learning for software engineering," ACM Computing Surveys (CSUR), vol. 54, no. 10s, pp. 1–73, 2022.
[15] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," Information Fusion, vol. 42, pp. 146–157, 2018.
[16] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," Comput Sci Rev, vol. 40, p. 100379, 2021.
[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks." [Online]. Available: http://code.google.com/p/cuda-convnet/
[18] T. J. Iyer, R. K., R. Nersisson, Z. Zhuang, A. N. Joseph Raj, and I. Refayee, "Machine Learning-Based Facial Beauty Prediction and Analysis of Frontal Facial Images Using Facial Landmarks and Traditional Image Descriptors," Comput Intell Neurosci, vol. 2021, pp. 1–14, Aug. 2021, doi: 10.1155/2021/4423407.
[19] J. N. Saeed and A. M. Abdulazeez, "Facial Beauty Prediction and Analysis based on Deep Convolutional Neural Network: A Review," Journal of Soft Computing and Data Mining, vol. 02, no. 01, Apr. 2021, doi: 10.30880/jscdm.2021.02.01.001.
[20] H. Chen, W. Li, X. Gao, and B. Xiao, "Novel Multi-Feature Fusion Facial Aesthetic Analysis Framework," IEEE Trans Big Data, vol. 9, no. 5, pp. 1302–1320, Oct. 2023, doi: 10.1109/TBDATA.2023.3255582.
[21] T. Gerlach, M. Danner, L. Peng, A. Kaminickas, W. Fei, and M. Rätsch, "Who Loves Virtue as much as He Loves Beauty?: Deep Learning based Estimator for Aesthetics of Portraits," in Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, SCITEPRESS - Science and Technology Publications, 2020, pp. 521–528. doi: 10.5220/0009172905210528.
[22] F. Chen, X. Xiao, and D. Zhang, "Data-Driven Facial Beauty Analysis: Prediction, Retrieval and Manipulation," IEEE Trans Affect Comput, vol. 9, no. 2, pp. 205–216, Apr. 2018, doi: 10.1109/TAFFC.2016.2599534.
[23] T. V. Nguyen, S. Liu, B. Ni, J. Tan, Y. Rui, and S. Yan, "Towards decrypting attractiveness via multi-modality cues," ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 9, no. 4, pp. 1–20, Aug. 2013, doi: 10.1145/2501643.2501650.
[24] D. Xie, L. Liang, L. Jin, J. Xu, and M. Li, "SCUT-FBP: A Benchmark Dataset for Facial Beauty Perception," Nov. 2015.
[25] S. Tyagi and D. Yadav, "A detailed analysis of image and video forgery detection techniques," Vis Comput, vol. 39, no. 3, pp. 813–833, 2023.
[26] J. Wang and X. Geng, "Label Distribution Learning by Exploiting Label Distribution Manifold," IEEE Trans Neural Netw Learn Syst, vol. 34, no. 2, pp. 839–852, Feb. 2023, doi: 10.1109/TNNLS.2021.3103178.
[27] R. Zheng, S. Zhang, L. Liu, Y. Luo, and M. Sun, "Uncertainty in Bayesian deep label distribution learning," Appl Soft Comput, vol. 101, p. 107046, Mar. 2021, doi: 10.1016/j.asoc.2020.107046.

[28] Y. Zhai, Y. Huang, Y. Xu, J. Zeng, F. Yu, and J. Gan, "Benchmark of a large scale database for facial beauty prediction," in Proceedings of the 2016 International Conference on Intelligent Information Processing, New York, NY, USA: ACM, Dec. 2016, pp. 1–5. doi: 10.1145/3028842.3028863.
[29] L. Liang, L. Lin, L. Jin, D. Xie, and M. Li, "SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction," Jan. 2018.
[30] L. Xu, J. Xiang, and X. Yuan, "CRNet: Classification and Regression Neural Network for Facial Beauty Prediction," 2018, pp. 661–671. doi: 10.1007/978-3-030-00764-5_61.
[31] L. Xu, J. Xiang, and X. Yuan, "Transferring rich deep features for facial beauty prediction," arXiv preprint arXiv:1803.07253, 2018.
[32] B. B. Zhu et al., "Attacks and design of image recognition CAPTCHAs," in Proceedings of the 17th ACM conference on Computer and communications security, 2010, pp. 187–200.
[33] L. Zhang, D. Zhang, M.-M. Sun, and F.-M. Chen, "Facial beauty analysis based on geometric feature: Toward attractiveness assessment application," Expert Syst Appl, vol. 82, pp. 252–265, Oct. 2017, doi: 10.1016/j.eswa.2017.04.021.
[34] D. Gray, K. Yu, W. Xu, and Y. Gong, "Predicting Facial Beauty without Landmarks," 2010, pp. 434–447. doi: 10.1007/978-3-642-15567-3_32.
[35] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep Learning Face Attributes in the Wild," in 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Dec. 2015, pp. 3730–3738. doi: 10.1109/ICCV.2015.425.
[36] H. Wu, G. Bezold, M. Günther, T. Boult, M. C. King, and K. W. Bowyer, "Consistency and accuracy of celeba attribute values," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 3257–3265.
[37] C.-S. Foo, S. Winkler, K.-H. Yap, G. Piliouras, and V. Chandrasekhar, "The unusual effectiveness of averaging in GAN training," arXiv e-prints, p. arXiv-1806, 2018.
[38] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 5505–5514.
[39] Y. Zhai, H. Cao, W. Deng, J. Gan, V. Piuri, and J. Zeng, "BeautyNet: Joint Multiscale CNN and Transfer Learning Method for Unconstrained Facial Beauty Prediction," Comput Intell Neurosci, vol. 2019, pp. 1–14, Jan. 2019, doi: 10.1155/2019/1910624.
[40] E. Vahdati and C. Y. Suen, "Female Facial Beauty Analysis Using Transfer Learning and Stacking Ensemble Model," 2019, pp. 255–268. doi: 10.1007/978-3-030-27272-2_22.
[41] L. Chen and W. Deng, "Facial Attractiveness Prediction by Deep Adaptive Label Distribution Learning," 2019, pp. 198–206. doi: 10.1007/978-3-030-31456-9_22.
[42] J. Gan et al., "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.

[43] Y. Zhai et al., “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.
[44] Y. Zhai et al., “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, doi: 10.1109/ACCESS.2020.2980248.
[45] L. Xu and J. Xiang, “ComboLoss for Facial Attractiveness Analysis with Squeeze-and-Excitation Networks,” Oct. 2020.
[46] I. Lebedeva, Y. Guo, and F. Ying, “Transfer Learning Adaptive Facial Attractiveness Assessment,” J Phys Conf Ser, vol. 1922, no. 1, p. 012004, May 2021, doi: 10.1088/1742-6596/1922/1/012004.
[47] F. Bougourzi, F. Dornaika, and A. Taleb-Ahmed, “Deep learning based face beauty prediction via dynamic robust losses and ensemble regression,” Knowl Based Syst, vol. 242, Apr. 2022, doi: 10.1016/j.knosys.2022.108246.
[48] M. Luthfi, R. F. Rachmadi, I. K. E. Purnama, and S. M. S. Nugroho, “Mobile Device Facial Beauty Prediction using Convolutional Neural Network as Makeup Reference,” in 2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), IEEE, Nov. 2022, pp. 1–5. doi: 10.1109/CENIM56801.2022.10037321.
[49] J. N. Saeed, A. M. Abdulazeez, and D. A. Ibrahim, “FIAC-Net: Facial Image Attractiveness Classification Based on Light Deep Convolutional Neural Network,” in 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), IEEE, Sep. 2022, pp. 1–6. doi: 10.1109/ICCSEA54677.2022.9936582.
[50] J. N. Saeed, A. M. Abdulazeez, and D. A. Ibrahim, “Automatic Facial Aesthetic Prediction Based on Deep Learning with Loss Ensembles,” Applied Sciences (Switzerland), vol. 13, no. 17, Sep. 2023, doi: 10.3390/app13179728.
[51] C. T. Yang, Y. C. Wang, L. J. Lo, W. C. Chiang, S. K. Kuang, and H. H. Lin, “Implementation of an Attention Mechanism Model for Facial Beauty Assessment Using Transfer Learning,” Diagnostics, vol. 13, no. 7, Apr. 2023, doi: 10.3390/diagnostics13071291.
[52] J. Gan, X. Xie, Y. Zhai, G. He, C. Mai, and H. Luo, “Facial beauty prediction fusing transfer learning and broad learning system,” Soft comput, vol. 27, no. 18, pp. 13391–13404, Sep. 2023, doi: 10.1007/s00500-022-07563-1.
[53] J. N. Saeed, A. M. Abdulazeez, and D. A. Ibrahim, “An Ensemble DCNNs-Based Regression Model for Automatic Facial Beauty Prediction and Analyzation,” Traitement du Signal, vol. 40, no. 1, pp. 55–63, Feb. 2023, doi: 10.18280/ts.400105.
[54] A. Żejmo, M. Gielert, M. Grabski, and B. Kostek, “Assessing the attractiveness of human face based on machine learning,” Procedia Comput Sci, vol. 225, pp. 1019–1027, 2023, doi: 10.1016/j.procs.2023.10.089.
[55] F. Bougourzi, F. Dornaika, N. Barrena, C. Distante, and A. Taleb-Ahmed, “CNN based facial aesthetics analysis through dynamic robust losses and ensemble regression,” Applied Intelligence, vol. 53, no. 9, pp. 10825–10842, May 2023, doi: 10.1007/s10489-022-03943-0.