A Comprehensive Review of Facial Beauty Prediction Using Multi-Task Learning and Facial Attributes

Ali H. Ibrahem[1†] and Adnan M. Abdulazeez[2]

[1]Department of IT, Technical College of Informatics - Akre, Akre University for Applied Sciences, Kurdistan Region – F.R. Iraq

[2]Technical College of Engineering, Duhok Polytechnic University, Kurdistan Region – F.R. Iraq

Abstract—**Beauty multi-task prediction from facial attributes is a multidisciplinary challenge at the intersection of computer vision, machine learning, and psychology. Despite the centrality of beauty in human perception, its subjective nature—shaped by individual, social, and cultural influences—complicates its computational modeling. This review addresses the pressing need to develop robust and fair predictive models for facial beauty assessments by leveraging deep learning techniques. Using facial attributes such as symmetry, skin complexion, and hairstyle, we explore how these features influence perceptions of attractiveness. The study adopts advanced computational methodologies, including convolutional neural networks and multi-task learning frameworks, to capture nuanced facial cues. A comprehensive analysis of publicly available datasets reveals critical gaps in diversity, biases, and ground truth annotation for training effective models. We further examine the methodological challenges in defining and measuring beauty, such as data imbalances and algorithmic fairness. By synthesizing insights from psychology and machine learning, this work highlights the potential of interdisciplinary approaches to enhance the reliability and inclusivity of automated beauty prediction systems.**

Index Terms—**Convolutional neural network, Facial beauty prediction, Facial attractiveness, Human rater.**

I. Introduction

Studies in different fields such as biology, philosophy, psychology, and art have tried to quantify beauty and challenged many aspects of esthetics and facial analysis. Face beauty assessment using computer vision is a relatively novel topic of study with a broad range of applications, including plastic surgery, the cosmetic industry, and facial photographs beautician (Bougourzi, et al., 2023). Facial beauty prediction (FBP) is a multi-paradigm computation problem that aims to develop human-like models for automatically assessing facial attractiveness. Recent studies have explored the use of multi-task learning (MTL) and transfer learning (TL) are robust techniques that can significantly reinforce the performance of FBP models. By leveraging these methods, researchers can build more strong and accurate systems that can capture complex facial attributes and their interplay. Combining these methods allows for better generalization across tasks and datasets, making them particularly useful in scenarios, such as motor imagery signal classification (Xie, et al., 2023). A diagram of the classification model of facial image attractiveness based on the TL approach is shown in Fig. 1. The diagram typically includes some key stages: preprocessing, feature extraction, and classification. In the preprocessing stage, facial images are normalized and augmented to enhance the dataset's diversity and robustness. During the training stage, a pre-trained convolutional neural network (CNN), such as VGG16 or ResNet50, is employed for feature extraction. This pre-trained model, initially trained on a large dataset, such as ImageNet, is fine-tuned on the attractiveness classification task by replacing its top layers with a new fully connected layer specific to the new task. The model learns to map the extracted features to attractiveness scores. In the testing stage, the fine-tuned model evaluates the unseen facial images to predict their attractiveness, demonstrating its generalization capabilities. The performance is then assessed using metrics such as accuracy and precision. The advancement of the model is enhancing predictive accuracy through MTL, which integrates task and feature correlations for effective knowledge sharing and moving beyond the traditional or classical handcrafted features for achieving an end-to-end optimization using deep learning. Moreover, the synthesis highlights the transformative potential of these models across applications in healthcare, cosmetics, social media, and more. The integration of 3D modeling and dynamic datasets could further refine predictions, offering more personalized and context-aware assessments.

ARO-The Scientific Journal of Koya University
Vol. XIII, No. 1 (2025), Article ID: ARO.11850. 12 pages
Doi: 10.14500/aro.11850
Received: 07 October 2024; Accepted: 01 January 2025
Regular review paper; Published: 01 February 2025
†Corresponding author's e-mail: ali.hikmat@auas.edu.krd

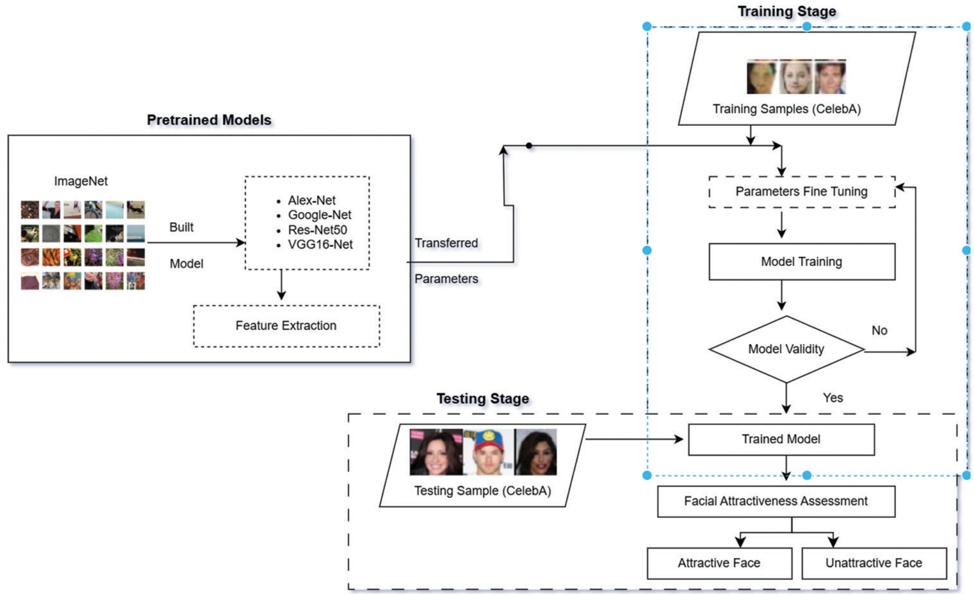

Fig. 1. The classification model of facial image attractiveness based on transfer learning path (Saeed, Abdulazeez and Ibrahim, 2022).

In the context of MTL, various paths have been proposed to facilitate knowledge sharing between tasks. These paths can be largely categorized into three groups based on the relationship between tasks or features. These approaches take into account both task relationships and trait correlations. While the max-margin method excels in discrimination, the Bayesian method offers greater flexibility in combining rich prior information (Taherkhani, et al., 2021).

Incorporating task relationships during training: This approach considers the relationships between tasks during the training phase. For example, one method treats the MTL challenge as a multi-label classification problem, utilizing prior knowledge of closely related labels. The model uses a max-margin multi-label formulation and integrates correlation-based interactions between labels into the prediction loss function.

Discovering common feature structures: Methods falling into this category target to identify common feature structures shared across all tasks or explore related features that are common among them. Such as a non-parametric Bayesian model is proposed to learn a commonly hidden space for all related tasks. This shared subspace captures the relationships between tasks, where each task's parameters (e.g., weight vectors in linear Support vector Machine SVM) are explained as a linear combination of basis tasks forming the latent shared subspace.

Recent way considering both tasks and feature correlations: The third category includes recent models, such as the max-margin method and Bayesian method.

From FBP to Personalized Facial Beauty assessment, the objective is to develop a model that can estimate facial attractiveness as accurately as humans. Traditional approaches often trust geometric features or global appearance descriptors (such as Color Histograms, Local Binary Patterns, Histograms of Oriented Gradients [HOG], and Gabor Filters) to learn FBP. However, these handcrafted attributes heavily rely on heuristic rules. With the remarkable success of deep learning, FBP can now be optimized using (CNNs) in an end-to-end manner. It's significant to note that most existing FBP methods are designed to learn esthetics at the population level (Lin, et al., 2023).

Currently, Artificial Intelligence (AI) and image improvement, receiving and quantifying beauty remains an interrogation able yet fascinating pursuit. This literature review shows the landscape of beauty multi-functional predictions grounded in face attributes analysis. With the proliferation of deep learning methodologies, researchers have investigated leveraging facial features to predict different sides of beauty, including attractiveness, age, gender, and ethnicity. This review begins by studying the foundational theories of beauty perception, encompassing both evolutionary and cultural perspectives. It proceeds to discuss the methodologies employed in beauty prediction models, encompassing feature extraction, dataset curation, and algebraic frameworks for instance CNNs and generative adversarial networks (GANs) (Cowen, et al., 2021).

Furthermore, this review investigates the implications and applications of beauty multi-task predictions across diverse domains, including cosmetics, healthcare, and social media. It highlights the potential societal impacts, ethical considerations, and biases inherent in beauty prediction algorithms. Moreover, it reviews the challenges faced in developing universal standards of beauty and the implications for inclusivity and diversity. Finally, this review outlines some future directions for research in this field, emphasizing the need for interdisciplinary collaboration, robust evaluation metrics, and transparency in algorithmic decision-making. By synthesizing insights from psychology, computer vision, and sociology, this literature review aims to provide a comprehensive overview of the current state and prospects of beauty multi-task predictions based on face attributes. The current literature review is structured according to the

http://dx.doi.org/10.14500/aro.11850

following. First, the datasets and image rating systems critical for training beauty prediction models are introduced. Second, several related works offering a comparative analysis of key results are highlighted. Finally, a conclusion and future directions for research, emphasizing the necessity for ethical and inclusive advancements in beauty prediction systems are presented.

II. Facial Beauty Dataset and Image Rating

The establishment of a definitive standard for beauty remains challenging due to its subjective nature. Therefore, FBP lacks extensively accepted authoritative data, making the construction of large-scale databases difficult. This section reviews the datasets commonly used in FBP studies, emphasizing their strengths, limitations, and the diversity they provide. An ideal dataset should encompass a range of attributes, including gender and age for the diversity of facial images. However, the available datasets often lack diversity, limiting model generalizability. For rating methods, facial images are typically annotated with beauty scores using human raters. The accuracy of these ratings significantly impacts the training of beauty prediction models. Despite the studies suggesting that facial beauty is a universal concept that can be learned through machine learning, establishing a definitive standard for beauty remains challenging due to the subjective nature of individual judgments. Consequently, FBP lacks widely accepted authoritative data, making the construction of large-scale databases difficult. A diverse range of attractiveness is essential for datasets used in facial image beauty research, which underscores the importance of selecting an appropriate face database for optimal model performance. Facial images can be sourced from the Internet, digital cameras, 3D scanners, existing facial image databases, and computer-generated images. Furthermore, these images must be evaluated using various rating methods. They are then labeled with beauty scores to establish the ground truth for the learning process and validate the model (Gan, et al., 2020).

A. Facial Images Databases

An ideal database should encompass a diverse array of facial images, representing all beauty levels across different genders, ages, races, poses, and expressions. In addition, it is essential for the dataset to include both ends of the facial attractiveness spectrum, featuring both highly attractive and very unattractive faces, to provide a comprehensive representation of the concept of beauty (Mu, 2013). However, due to privacy concerns, the majority of datasets are not publicly accessible. In addition, landmarks are critical for evaluating the geometry cue, which is lacking in some of the benchmarks (Aarabi, et al., 2001). The majority of existing datasets are relatively small (Eisenthal, Dror and Ruppin, 2006; Gunes and Piccardi 2006). Moreover, they are neutral (postures, expression, attractiveness), insufficiently diverse, and most of them are limited for female facial beauty (Altwaijry and Belongie 2013; Kalyta, et al., 2022).

Gender is an important factor in facial beauty analysis, with many facial databases primarily focusing on female subjects. These datasets often yield better results in beauty classification than those centered on male subjects. This discrepancy may be due to the relative simplicity of analyzing and computing beauty in female faces compared to male faces. Consequently, researchers frequently use images of females or a mix of both genders in their studies, whereas datasets comprising exclusively male images are less common (Ahmadimehr and Moridani 2020). Due to the absence of datasets that meet the specific conditions required for (FBP), such as considering diverse attractiveness levels, particularly extremes of attractiveness and unattractiveness, researchers have resorted to using general face recognition datasets cited in references (Schmid, Marx and Samal. 2008), as well as facial expression datasets from (Xu, et al., 2018) for their FBP studies especially in the initial stages. The limitations posed by variations in facial posture and expression are critical factors that can constrain both dataset availability and model performance. Consequently, existing FBP databases predominantly consist of frontal face images with neutral expressions.

In spite of the rarity of facial image datasets especially designed for FBP and related tasks, several well-known face datasets are mostly used as benchmarks for this purpose as shown in Table I. Gray, et al. (2010) carried out the first attempt to address the issue of facial feature landmark localization and imposing stringent limits on training samples by building a large number of FBP database of 2056 females known as Hot or Not HotorNot. It is a very difficult dataset since the images are under unconstrained conditions on the background, expression, position, lighting, race and age with no use of landmarks, as well as the problem of predicting the locations of landmark features sufficiently. The same notion was also used for constructing a benchmark called the Large-Scale Asian Facial Beauty Database LSAFBD (Zhai, et al., 2016), which comprised 20,000 labeled images of gender and 80,000 unlabeled ones. However, both datasets are based on apparent features only and are not accessible publicly.

Most prediction performances of facial esthetic are fulfilled on nearly small facial datasets through employing conventional machine learning or shallow network learning methods (Zhai, et al., 2020). As a result, small datasets often lead to overfitting during the training of the model. This makes the prediction model less effective. In addition to this, datasets that are built under specific computational constraints limit the performance and flexibility of the model. This complicates the comparisons between models that are developed with various methodologies. For example, the widely-used FBP databases the South China University of Technology SCUT-FBP5500 (Zhai, et al., 2020) and the South China University of Technology SCUT-FBP (Xie, et al., 2015) have produced promising results in many researches. However, SCUT-FBP is made up of only 500 Asian female faces, restricting the performance of models trained on this data. Conversely, the SCUT-FBP5500 dataset, though relatively large, has limitations in terms of lighting, blurriness, and positioning, which can

TABLE I
FBP Dataset Benchmarks

Database Name	References	Year	Size	Gender	Age	Landmark	Ethnicity	Pose	Expression	Beauty score
Celeb A	Saeed and Abdulazeez, 2022	2022	200K	F/M	Different	5	Different	Different	Different	2
LFWA Labeled Faces the Wild Attributes	Fan, et al., 2019	2019	13,233	F/M	Different	N/A	Different	Different	Different	N/A
RAF-DS	Zheng, et al., 2024	2024	29672	F/M	Different	68	Diverse	Almost frontal	Different	N/A
CASIA-WebFace	Sekhar, et al., 2024	2024	500000	F/M	Different	68	Diverse	Diverse	Different	N/A
M2B	Nguyen, et al., 2013	2021	1240	F	19-40	N/A	Western/Eastern	Different	Different	10
HotorNot	Xu, et al., 2018	2018	2056	F	18-40	N/A	Diverse	Almost frontal	Different	10
Beauty 799	Chen, Xu and Zhang, 2014	2014	799	F	N/A	98	Diverse	Frontal	Almost Neutral	3
SCUT-FBP	Xie, et al., 2015	2015	500	F	N/A	84	Asian	Frontal	Neutral	5
LSAFBD	Gan, et al., 2023	2023	20000	F/M	Different	5	Asian	Almost frontal	Different	5
SCUT-FBP5500	Lebedeva, Guo and Ying, 2022	2022	5500	F/M	15–60	86	Asian/Caucasian	Frontal	Neutral	5
MEBeauty	Lebedeva, Guo and Ying, 2022	2022	2550	F/M	Different	N/A	Diverse	Different	Different	10
CFD	Gan, et al., 2023	2023	597	F/M	17–65.	86	Diverse	Almost frontal	Different	7
SCface	Grgic, Delac and Grgic, 2011	2011	4160	F/M	Different	N/A	Diverse	Different	Neutral	N/A

impact the effectiveness of attractiveness prediction models (Lebedeva, Guo and Ying, 2021). Moreover, researchers have shown that the SCUT-FBP5500 dataset has an imbalanced distribution of beauty scores, potentially influencing correlation analyses. Another public dataset, the CelebFaces Attributes dataset Celebrity Attributes (CelebA) (Liu, et al., 2015), contains over 200,000 images with 40 attribute annotations, including attractiveness. Despite its extensive range of poses and backgrounds, CelebA is mostly utilized in computer vision tasks such as face recognition, detection, component localization, editing, and synthesis. While it has been frequently employed for facial beauty classification, its binary beauty scores (attractive and unattractive) simplify computation but may not fairly evaluate beauty. The range of beauty scores affects considerably the fairness of beauty assessment. This makes the establishment of robust ground truth fundamental for both learning and model validation processes.

Real-world Affective Faces Dataset (RAF-DS) consists of over 30,000 facial images annotated with different expressions and facial attributes. It was labeled through crowdsourcing, variation, and ensuring high reliability. While primarily designed for expression recognition tasks in FBP (Zheng, et al., 2024). The Surveillance Cameras Face database (SCface) (Grgic, Delac and Grgic, 2011) includes 4,160 facial images captured by surveillance cameras in real-world conditions. The database contains images of 130 individuals taken from diverse angles and with low quality to simulate surveillance scenarios. SCface serves as a crucial resource for face recognition tasks in challenging environments, for example, scenarios with varying distances and inconsistent lighting conditions.

Chinese Academy of Sciences Institute of Automation WebFace (CASIA-WebFace) dataset provided by the Chinese Academy of Sciences Institute of Automation, consists of around 494,000 facial pictures belonging to 10,575 different individuals. It is primarily created for use in face recognition research and encompasses images obtained from the web spanning various conditions such as pose, expression, and lighting. This diversity makes it an excellent choice for training and evaluating face recognition algorithms, although its primary focus is not on beauty evaluation (Sekhar, et al., 2024). In Gan, et al. (2023), the Chicago Face Database is a thorough collection of images showcasing faces with a variety of demographic characteristics, such as age, ethnicity, and gender. It offers precise evaluations on aspects, such as attractiveness, trustworthiness, and other social traits, which enhances its utility for research in social perception and facial features analysis. This dataset is commonly employed in psychological with behavioral studies, emphasizing social sciences instead of computational beauty forecasting.

Recently, a Multi-Ethnic (ME Beauty) dataset consists of 2,550 in-the-wild facial images of both males and females, named ME Beauty (Lebedeva, Guo and Ying, 2022), has been introduced and is expected to become a prominent resource for further study in facial beauty assessment. However, the personalized nature of FBP may be influenced by the relatively considerable number of raters. Existing facial datasets primarily concentrate on static features. To address dynamic FBP, Weng, et al. (2021) developed the first Video-based Facial Attractiveness Prediction (VFAP) dataset, which includes 1,430 short video clips of facial performances from TikTok. Variations in beauty scores may take place because of using likes, comments, and forward as indicators of attractiveness. In addition to this, the gender distribution in TikTok's facial performance videos is uneven which makes the beauty rankings of male faces less easily explained in comparison to those of female faces.

The challenges outlined in this review necessitate innovative computational approaches, which are discussed in the next section. Building on the dataset limitations mentioned in the previous section. Researchers have developed sophisticated computational methods to enhance FBP systems. For Deep Learning, techniques such as CNNs enable the end-to-end optimization of beauty prediction, leveraging features such as facial symmetry and texture. Regarding MTL, integrating auxiliary tasks, such as age or gender prediction, MTL frameworks enhance the robustness of beauty prediction models. These methods demonstrate the potential to advance beauty prediction systems, setting the stage for a deeper exploration of related works.

III. Related Works

The related work is divided into three parts mainly: CNN, MTL, and Machine learning. This section synthesizes findings from the available research, highlighting contributions to FBP while identifying gaps for future studies. For the comparison of computational approaches, the Studies using CNN architectures, such as VGG16, EfficientNetV2B0, and ResNet50, balance computational cost against predictive performance. Lightweight models prioritize efficiency but may compromise precision. Challenges and future directions indicate key limitations include dataset biases, overfitting, and subjective beauty standards. Addressing these requires improved dataset diversity, advanced multimodal methods, and fairness-aware models.

A. CNN

Saeed and Abdulazeez (2021). Advanced deep learning techniques, particularly deep CNNs (DCNN), have been employed to assess facial attractiveness – a task shaped by subjectivity and cultural influences. Challenges in this domain include the lack of comprehensive datasets encompassing diverse attributes such as age, gender, culture, ethnicity, and facial expressions, as well as the absence of standardized evaluation metrics for FBP systems. Despite their relevance in areas, such as beauty product recommendations and cosmetic procedures, FBP studies face hurdles, such as limited datasets, resource constraints, and inconsistent assessment criteria. Current approaches address data scarcity through semi-supervised learning, data augmentation, and GANs. Developing diverse datasets and incorporating 3D facial models, along with varied rating methods, can enhance predictive accuracy and foster further advancements, paving the way for future research in FBP technologies.

Favorskaya and Pakhirka (2023) Introduced an innovative multi-task CNN for age-group classification enhanced by auxiliary tasks such as the identification of gender and facial expression analysis. The model features three interconnected CNN branches sharing initial feature extraction layers with relatively straightforward architectures. Different from the traditional approaches, the age-group classification task is framed as a ranking problem rather than a regression, improving prediction accuracy. Tested on five Internet Movie Database (IMDB), University of Tennessee, Knoxville Face Dataset UTKFace, and Morphological MORPH II—the model outperformed Densely Connected Convolutional Networks (DenseNet) and queeze-and-Excitation Networks (SENet), achieving a mean accuracy of 96.8% for age-group ranges (18–25), (26–40), and (41–65). This work highlights the benefits of leveraging auxiliary attributes and demonstrates the potential of MTL in facial analysis. The study overlooks substantial challenges such as dataset imbalance, real-world applicability, and the inclusion of diverse facial attributes, all of which can impact the robustness and fairness of the model.

Gao, et al. (2018) addressed facial attractiveness prediction (FAP) using a deep MTL strategy. The proposed model considers both texture and shape (geometric landmarks) by incorporating two tasks: predicting attractiveness scores and localizing fiducial landmarks. A lightweight CNN is designed to learn these tasks effectively, even with limited training data. Evaluated on the SCUT-FBP dataset, the method achieves a correlation score of 0.92, demonstrating its efficacy. Furthermore, the model outperforms several cutting-edge approaches and reliably compares facial images pre- and post-make-up or beautification. This research highlights the advantages of a MTL framework for precise and efficient facial beauty analysis.

B. MTL

Another study (Savchenko, 2021) introduced an innovative training pipeline for compact CNNs, achieving cutting-edge performance in recognizing facial expressions and attributes across various datasets. Utilizing MTL and pre-training on the extensive Visual Geometry Group Face (VGGFace2) dataset, the method enhances resilience to face extraction and alignment challenges. The models, trained on tightly cropped faces, deliver high accuracy, impressive speed, and a small footprint, making them ideal for mobile devices. Although the models perform exceptionally well, reliance on conventional classifiers (e.g., SVMs, random forests) slightly limits their accuracy below state-of-the-art levels. This study highlights the potential of lightweight architectures for efficient decision-making in practical scenarios. The robustness to face extraction and alignment is stated, but the potential limitations or challenges of relying on cropped faces without margins are not addressed. Moreover, the discussion on challenges specific to emotion recognition is limited.

Fan, et al. (2019) presented an innovative MTL framework for predicting face attributes, focusing on a smile and gender recognition tasks. The architecture addresses the challenges of conventional MTL models by integrating attention modules into task-specific layers, enabling the model to learn disentangled representations for predicting face attributes through softmax layers. Experiments on the LFWA and FotW both datasets demonstrate the model's superior performance over traditional MTL architectures and state-of-the-art methods. Future directions include expanding the model to additional face feature tasks, applying attention mechanisms to other layers, and exploring dynamic weighting methods to enhance predictive capabilities. The research gap lies in the limited scope of tasks (smile and gender), static task relationships, and restricted application of attention mechanisms to task-specific layers. In addition, the scalability, dynamic weighting, and generalizability of the model across diverse datasets are unexplored. The research issue focuses on enhancing MTL frameworks to dynamically model task dependencies, generalize across varied attributes, and improve computational efficiency.

The 2M BeautyNet, a multi-input multi-task network designed for FBP is conducted by Gan, et al. (2020). The network prioritizes beauty prediction as its core task while leveraging gender recognition as a secondary task through multi-task TL to boost performance. By transferring pre-trained network parameters across datasets, the model enhances accuracy. It employs a multi-task loss automatic learning strategy to balance contributions from each task, preventing dominance by any single task. After training, a random forest classifier replaces the softmax classifier for improved results. Experimental findings on the SCUT FBP5000 with LSFBD databases show the model achieves an FBP accuracy of 68.23%, surpassing alternative approaches. Future enhancements include exploring additional tasks for MTL, integrating local information, and accounting for other factors influencing facial beauty. This methodology shows promise for applications in esthetic surgery, cosmetic recommendations, and facial beautification, but the research shows encouraging outcomes in forecasting facial attractiveness and fails to consider the influence of various demographic characteristics, such as gender, age, or the inclusion of 3D facial traits, which could improve the adaptability and practical use in real-world scenarios of the model.

Another study by Lin, et al. (2021) described an innovative multi-task network that is created to simultaneously identify faces and forecast facial characteristics (such as gender and age). Instead of treating face detection and attribute prediction as two separate tasks, the study has merged them into a unified model. It has also introduced an optimization technique based on noise estimation to dynamically adjust the weights of multi-task losses, enhancing task equilibrium. Experiments conducted on the CelebA dataset demonstrate that this approach achieves an excellent performance in terms of both accuracy and speed.

C. Machine Learning

Lebedeva, Guo and Ying (2023) focus on the personalized evaluation of facial attractiveness, aiming to accommodate the unique beauty preferences of individuals. Unlike conventional models that predict universal beauty norms, the proposed method uses meta-learning to capture shared beauty preferences during meta-training and adapts to new individuals with minimal data in the meta-testing phase. Tests on a diverse dataset of facial attractiveness – varying in age, ethnicity, gender, and expression – demonstrate the technique's ability to discern individual preferences and outperform current models in personalized scenarios. Future research can explore alternative machine learning techniques to enhance accuracy with limited annotated samples, investigate 3D FAP, utilize active learning to optimize data, and integrate the approach into recommender systems to address cold-start issues. The main point of the research is evaluating personalized facial attractiveness. However, it is experiencing difficulties in managing scarce labeled data, incorporating diverse cultural preferences, and ensuring efficient scalability for practical applications such as recommendation systems. Bridging these deficiencies can improve forecast accuracy, address cold-start problems, and allow for effective deployment. Machine-learning methods are applied to explore the impact of facial geometric attributes on attractiveness and emotional interpretation, focusing on changes after rhinoplasty. Analysis of data from 42 patients using multivariate regression revealed that increased nasolabial and nasofrontal angles correlated with greater attractiveness. Neural networks proved most effective in recognizing facial expressions, emphasizing the importance of mouth, eyebrows, and eye shape in emotion identification. These results suggest that highlighting these geometric features in rhinoplasty can enhance esthetic outcomes. Furthermore, findings from emotion classification support an improved version of the Facial Action Coding System, demonstrating its reliability in linking facial structure and emotions. The research highlights the role of data-driven approaches in shaping plastic surgery practices (Štěpánek, Kasal and Mestak, 2018). In addition, Zhang, Xiao and Lu (2018) inducted models to predict and enhance facial attractiveness. The Geometric + Principal Component Analysis Network (Geo + PCANet) model uses geometric features from facial landmarks to estimate beauty indices closely aligning with human assessments. Skin esthetics are improved with multi-level median filtering, while facial geometry is refined using the moving least squares technique. In addition, an average facial beautification model is proposed to enhance overall appearance. The findings validate the efficacy of these approaches, though challenges remain, particularly in the computational efficiency of skin enhancement. Future work will focus on refining these methods for greater practical applicability. The objective of this research is to create effective, scalable models that incorporate sophisticated neural structures for enhanced and computationally efficient facial beautification. Nevertheless, the issue of the study is rooted in depending on the geometric attributes and neglecting investigation into deep learning approaches to improve beauty prediction and enhancement (Lin, et al., 2023) focuses on personalized FBP (PFBP), aiming to predict individual esthetic preferences from a few personalized images. Different from the traditional models that generalize facial attractiveness, PFBP treats each user as a distinct meta-task within the Few-Shot Learning (FSL) framework. The proposed MetaFBP framework includes a universal feature extractor to identify common esthetic traits and a high-order predictor that adapts swiftly to personal preferences, overcoming issues of slow adaptation and overfitting in traditional methods. Experiments on newly created PFBP benchmarks show the effectiveness of the framework, confirming its ability to make accurate, user-specific beauty predictions. The study also discusses the limitations of conventional FSL methods and introduces a learning-to-learn mechanism designed for faster and more effective adaptation in PFBP tasks.

The review of some research on FBP, which investigates dataset, attributes, algorisms, challenges, future directions, strengths, and metrics is presented in Table II. The comparative analysis of the computational approaches for predicting beauty judgments using facial attributes reveals several strengths, such as the high accuracy and scalability of methods, including multi-task TL, which benefits from shared learning frameworks. Furthermore, models, including ShadowFace3D showcase robustness and generalization through diverse datasets and novel integrations of geometry and texture. However, limitations persist, including data insufficiencies and ethnic and gender biases, which hinder model fairness and inclusivity. Moreover, advanced CNNs may encounter problems such as overfitting and significant computational requirements. Research gaps contain the lack of various datasets, particularly for features, such as hairstyles, smiles, and 3D geometries, and a reliance on one-dimensional inputs without multi-modal integration. Future directions should prioritize enhancing data diversity, integrating multimodal approaches, exploring advanced attention mechanisms for feature enhancement, and developing subjectivity-aware models to better reflect human perceptions of beauty. Furthermore, investigating cross-cultural esthetics through region-specific data collection will provide deeper insights into the variability of beauty standards. By comparing these approaches, this section underlines the necessity of interdisciplinary solutions to enhance FBP systems.

The review provides a critical assessment of methodologies used for predicting facial beauty by examining the trade-offs in CNN architectures and dataset biases. CNN studies, such as those utilizing VGG16 or ResNet50, often balance complexity and computational cost against performance. Conversely, lightweight models prioritize efficiency but may compromise on precision. Dataset biases are a significant concern, as many lack diversity in ethnicity, age, and gender, leading to generalized models across various populations. The review underscores challenges such as overfitting from small datasets, imbalanced beauty score distributions, and subjective beauty judgments influenced by cultural factors. To identify effective approaches, the review explores how TL, data augmentation, and MTL can address dataset limitations and enhance model robustness. It is highly necessary to stress the importance of incorporating varied datasets, rectifying imbalances, and utilizing fairness metrics to comprehensively assess models. This critical synthesis aims to guide future research toward more inclusive and efficient beauty prediction systems.

Traditional machine learning methods, such as SVR and KNN, leverage handcrafted features like geometric landmarks and global descriptors. While computationally efficient, these approaches struggle with capturing complex relationships in large-scale datasets. Techniques such as local binary patterns, Gabor filters, and HOG focus on specific visual patterns but require extensive preprocessing and are sensitive to variations in lighting and pose. MTL improves model performance by incorporating auxiliary tasks, such as gender recognition, to enable robust feature sharing; however, it requires careful balancing of tasks and dataset preparation. GANs contribute by simulating esthetic transformations, including facial symmetry enhancements, though they are computationally demanding and complex to train. PFBP utilizes meta-learning frameworks like FSL to adapt models to individual preferences, promoting inclusivity but presenting scalability challenges. Advances in dataset diversity, with datasets like MEBeauty and RAF-DS, address demographic variations and biases, aligning outputs with inclusive beauty standards. These developments demonstrate a shift from rule-based approaches to adaptive, data-driven systems, enhancing generalization and fairness in beauty prediction tasks. This review synthesizes an understanding from computer vision, psychology, and ML to provide a comprehensive overview of FBP. It emphasizes the significance of various datasets, ethical considerations in developing inclusive beauty prediction models, and fairness in algorithmic decisions.

IV. Conclusion and Future Directions

This review has provided a comprehensive synthesis of computational methodologies for predicting beauty judgments from facial attributes, addressing the interplay between individual, social, and cultural factors. By leveraging deep learning frameworks, feature extraction techniques, and ensemble models, significant progress has been made in automating beauty predictions. The work underscores the importance of diverse and robust datasets, detailing their limitations and the challenges they introduce, such as data variation, subjective judgments, and algorithmic biases. One of the major contributions of this review is its integration of interdisciplinary perspectives – spanning psychology, and sociology, with computer vision – to present a holistic understanding of beauty prediction. It emphasizes how advancements in transfer and MTL have propelled the field, enabling nuanced modeling of complex facial attributes. Moreover, the synthesis highlights the transformative potential of these models across applications in healthcare, cosmetics, social media, and more. Despite these advances, there remains a critical need for addressing the ethical and societal implications of beauty algorithms, particularly concerning inclusivity and fairness across cultural contexts. The integration of 3D modeling and dynamic datasets could further refine predictions, offering more personalized and context-aware assessments. Future research should prioritize developing datasets that encompass greater demographic diversity and employ active learning techniques to mitigate biases. In addition, collaborative efforts between fields – incorporating ethics, computational sciences, and the arts – could redefine the standards and applications of beauty prediction models. Innovations in model interpretability, explainability, and user-centric design are essential for fostering trust and expanding the real-world utility of these algorithms. This review serves as a foundation for advancing beauty prediction, urging the community to adopt transparent, inclusive, and multidisciplinary approaches in the journey ahead.

TABLE II
The Summary of Facial Beauty Computational Models that Utilized Features-based Approaches

References	Categories and description	
Gan, et al., 2020	Dataset	LSFBD and SCUT-FBP5500
	Attributes	Gender and Skin
	Algorithm	Random Forest Classifier.
	Limitations	How to establish a versatile, effective multi-input multi-task network that merges local information and other elements that affect facial beauty.
	Future Direction	How to establish a versatile, effective multi-input multi-task network that merges local information and other elements that affect facial beauty.
	Strengths	Use of multi-task transfer learning in 2M BeautyNet, improving facial beauty prediction by leveraging knowledge from related tasks for better performance.
	Metrics	Accuracy up to 68.23%
Panić, Marjanović, and Bezdan, 2024	Dataset	diverse
	Attributes	Age
	Algorithm	CNN
	Limitations	Dataset Composition, and Bias Across Ethnic Groups.
	Future Direction	Use pre-trained CNN models, particularly VGG19, which are further fine-tuned aiming at the prediction of age.
	Strengths	Focus on Bias means addresses the critical issue of ethnic bias in facial age prediction models, supporting fairness and inclusivity.
	Metrics	Mean Absolute Error (MAE): 7.946 tested on African faces. 7.677 tested on Asian faces.
Cao, et al., 2020	Dataset	SCUT-FBP5500
	Attributes	All Facial Attributes
	Algorithm	CNN
	Limitations	Designing active Network. Attention Mechanism. Significance Distribution among features. To address this, the paper presents a joint spatial-wise and channel-wise attention (SCA) block. This approach helps exploit inner correlations among features and leads to a better representation of facial features.
	Future Direction	The improvement of Network Structures and Multimodal methods: integrating the prediction of facial beauty with other modalities, such as voice or body language
	Strengths	Utilization of Deep Learning and Scalability
	Metrics	• (MAE): 0.2595; Root-Mean-Square Error (RMSE) = 0.3397 Root-Mean-Square Error (RMSE)) = 0.8570
Chen, et al., 2021	Dataset	CelebHair
	Attributes	Hairstyle
	Algorithm	CNNs and Spatial Transformer Network (STN).
	Limitations	The initial challenge that is highlighted in the abstract is the absence of appropriate hairstyle-related datasets that are necessary for improving the recommendation application hairstyle.
	Future Direction	The hairstyle try-on experience needs to be refined for users through employing Interface GAN.
	Strengths	Creation of a new large-Scale database
	Metrics	Accuracy=85.45%
Moridani, et al., 2023	Dataset	Most Beautiful Women Faces (MBWFs).
	Attributes	All Facial Attributes
	Algorithm	K-Nearest Neighbors (KNN). Support Vector Regression (SVR).
	Limitations	Beauty Subjectivity and Feature Extraction.
	Future Direction	Future researches need to be done to improve attractive models through exploring more features. This results study is useful for the industry, modeling, and marketing of beauty, where ranking attractiveness is significant.
	Strengths	Human-like Evaluation that's mean The method mimics human judgment, achieving alignment together with human perceptions of facial beauty prediction.
	Metrics	Coefficient of determination (R2) =0.9902. RMSE=0.0056. and Mean Absolute Percentage Error (MAPE) =0.0856.
Gao, et al., 2018	Dataset	SCUT-FBP database
	Attributes	All Facial Attributes
	Algorithm	CNN
	Limitations	Inadequate Label Information, Overfitting, and intricacy of facial features.
	Future Direction	It is suggested that the proposed deep multi-task learning that is based on a prediction model joint with landmark localization is effective for Facial Attractiveness Prediction.
	Strengths	Getting better accuracy and attribute generalization.
	Metrics	Correlation=0.92

 http://dx.doi.org/10.14500/aro.11850

TABLE II
(Continued)

References	Categories and description	
Yuan, et al., 2024	Dataset	University of Tennessee, Knoxville Face Dataset (UTKFace dataset).
	Attributes	Age, Gender, and Race
	Algorithm	CNN
	Limitations	The Imbalance of data search for means treatment imbalanced database where certain features may be underrepresented.
	Future Direction	Both Enhanced data augmentation as well as improved model architectures implies designing more efficient model architectures to deal with the intricacy and improve performance.
	Strengths	Outweigh in employing uncertainty-based weighting in MTL enhancing the precise and equitable estimation of different facial attributes.
	Metrics	Age=64.74, Gender=90.91, and Race=79.98
Jamoliddin and Yoo, 2022	Dataset	UTKFace dataset
	Attributes	Age as well as Gender
	Algorithm	CNN
	Limitations	Limited Dataset: Small-scale CNNs often lack access to extensive, diverse datasets. This leads to overfitting and poor generalization to new data. To Ensure sufficient variability in the training data is indispensable for the enhancement of model robustness.
	Future Direction	Improving Accuracy seeks to Enhance the performance of the model with advanced architectures or additional data.
	Strengths	Strength lies in its efficient employ of a small-scale CNN construction to achieve accurate gender and age classification, balancing performance and computational cost.
	Metrics	F1 Score=0.90
Vahdati and Suen, 2021	Dataset	SCUT-FBP5500
	Attributes	All Facial Attributes
	Algorithm	CNN
	Limitations	The complexity of the Model as well as partiality and balance.
	Future Direction	Improving both model architecture and addressing subjectivity.
	Strengths	Innovative use of multi-task and multi-stream CNN To forecast facial attractiveness by thoroughly examining each facial feature and its impact.
	Metrics	Accuracy: 95%; Correlation Coefficient: 0.9; F1 Score: 0.9
Xiao, et al., 2021	Dataset	ShadowFace3D
	Attributes	All facial attributes
	Algorithm	CNN
	Limitations	3D Data complexity. Dataset Availability means There could be a shortage of large, high-quality 3D facial datasets annotated with attractiveness ratings that can limit model training and generalization.
	Future Direction	Investigate how our deep learning network interprets facial attractiveness by employing geometric as well as textural features. To figure out cross-cultural esthetics, more various data will be collected. Beauty3DFaceNet and ShadowFace3D will be used for applications, such as enhancing 3D facial attractiveness.
	Strengths	Fusion of deep geometry and texture features using Beauty3DFaceNet, enabling accurate and comprehensive 3D facial attractiveness prediction.
	Metrics	Pearson correlation coefficient=0.742
Wang, et al., 2017	Dataset	Morphological (MORPH II). CelebA. LFWA.
	Attributes	Age, gender, and race.
	Algorithm	CNN
	Limitations	Heterogeneity of Attributes indicates that facial attributes, such as gender, age, and expressions vary noticeably in terms of characteristics and data needs. This requires various information for the accuracy prediction.
	Future Direction	Examine both illumination and pose normalization methods and automatic attribute category grouping approach as well for efficient attribute prediction.
	Strengths	Strength lies in its application of deep multi-task learning to jointly predict diverse facial attributes, enhancing efficiency besides accuracy through shared attribute learning.
	Metrics	Accuracy: Age: 85.3, Gander: 98, and Race: 96.6
Yin, et al. 2020	Dataset	LFW, WIDER FACE, Queen's University Machine Learning - Surveillance Face Dataset (QUML-SurvFace), and SCface datasets.
	Attributes	All Facial Attributes
	Algorithm	Feature Adaptation Network (FAN)
	Limitations	Disentanglement learning feature and adaptation feature. By first disentangling the face features into identity and non-identity components, it facilitates our adaptation network to require both feature-level and image-level similarity regularizations. This framework is appropriate for both paired and unpaired training,
	Future Direction	The main aim is to learn the features of robust identity for FR. These features are used to produce a normalized face with enhancement facial details and neutral Pose Illumination Expression (PIE).
	Strengths	Normalizes and improves face recognition performance in opposed surveillance scenarios.
	Metrics	LFW Accuracy=95.2%, QMULAccuracy=70.88%, and SCface=90.3

TABLE II
(Continued)

References	Categories and description	
Sagonas, et al., 2016	Dataset	Diverse
	Attributes	N/A
	Algorithm	Constrained Local Mod (CLM) and Coarse-to-Fine Auto-Encoder Networks (CFAN)
	Limitations	Authors compare their approaches with other state-of-the-art. They do so by applying, in many cases, completely different databases for training compared to the ones that the other methods were originally trained on.
	Future Direction	Robustness to Lighting and Quality Variations mean enhancing the robustness of detection algorithms to varying illumination conditions and image qualities and enhanced Datasets as well.
	Strengths	Providing an overall benchmark database and evaluation framework, advancing the development and comparison of strong face alignment algorithms under real-world conditions
	Metrics	Accuracy=95.8%
Savchenko, 2021	Dataset	UTKFace, Affective Facial Expressions Network Dataset) AffectNet(, Acted Facial Expressions in the Wild (AFEW), and Video-based Group-level Affect and Face Dataset) VGAF datasets(
	Attributes	Age and Gender
	Algorithm	CNN, Support Vector Machines (SVM), and Random forests.
	Limitations	Handling variability in data means handling variations in real-world data's occlusions, stances, illumination, and face expressions. In addition, the efficiency of models indicates that there must be a balance in models in terms of both computing efficiency and intricacy so as to be used on tools that have constrained resources.
	Future Direction	Improved Model Efficiency hints focusing on the creation of more efficient lightweight models that involve less computational power when maintain high accuracy. Enhanced MTL indicates Further explore and refine MTL approaches to enhance the simultaneous recognition of facial expressions and attributes.
	Strengths	Its use of lightweight neural networks with MTL to effectively and accurately recognize facial expressions and attributes simultaneously.
	Metrics	Accuracy: 94.0% on the AffectNet dataset Accuracy: 88.7% on the CelebA dataset.
Liang, et al., 2017	Dataset	SCUT-FBP
	Attributes	Eyes, lips, and overall symmetry
	Algorithm	Support vector regression (SVR)
	Limitations	Constraints of Asian Female face on the SCUT-FBP dataset
	Future Direction	Construct a large-scale benchmark database in a later study.
	Strengths	Use of region-aware scattering convolution networks to capture detailed and localized facial features, enhancing the accuracy of FBP.
	Metrics	PC=0.83
Mao, et al., 2020	Dataset	CelebA. And LFWA datasets
	Attributes	All Facial Attributes
	Algorithm	Deep Multi-task Multi-label CNN, (DMM-CNN)
	Limitations	The study investigates difficulties such as handling data imbalance.
	Future Direction	Improving model accuracy seeks to Enhancing the CNN architecture or examining new strategies so as to make the accuracy of facial attribute classification in better shape.
	Strengths	Its use of a deep multi-task, multi-label CNN architecture, enabling accurate classification of multiple facial attributes simultaneously by leveraging shared attribute learning.
	Metrics	Mean accuracy CelebA=91.70% and Mean accuracy LFWA=86.56%
Rohani, Farsi and Mohamadzadeh, 2023	Dataset	IMDB-WIKI (Internet Movie Database) and GENKI-4K datasets
	Attributes	Smile, age, and gender
	Algorithm	CNN
	Limitations	About features have fewer samples, causing bias with suboptimal performance.
	Future Direction	Develop architectures to share knowledge while preserving task accuracy.
	Strengths	Employing deep multi-task convolutional neural networks to efficiently classify multiple facial features, leveraging shared representations for improved performance and efficiency.
	Metrics	Smile accuracy=96.63 Gender accuracy=93.20 Age accuracy=68.92

References

Aarabi, P., Hughes, D., Mohajer, K., and Emami, M., 2001. The Automatic Measurement of Facial Beauty. In: *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*. IEEE.

Ahmadimehr, S., and Moridani, M.K., 2020. Evaluating facial attractiveness through proportions analysis based on geometric features. *Journal of Image Processing & Pattern Recognition Progress*, 7(2), pp.20-26.

Altwaijry, H., and Belongie, S., 2013. Relative Ranking of Facial Attractiveness. In: *2013 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE.

Bougourzi, F., Dornaika, F., Barrena, N., Distante, C., and Taleb-Ahmed, A., 2023. CNN based facial aesthetics analysis through dynamic robust losses and ensemble regression. *Applied Intelligence,* 53(9), pp.10825-10842.

Cao, K., Choi, K.N., Jung, H., and Duan, L., 2020. Deep learning for facial beauty prediction. *Information*, 11(8), p.391.

Chen, F., Xu, Y., and Zhang, D., 2014. A new hypothesis on facial beauty perception. *ACM Transactions on Applied Perception*, 11(2), pp.1-20.

Chen, Y., Zhang, Y., Huang, Z., Luo, Z., and Chen, J., 2021. CelebHair: A New Large-scale Dataset for Hairstyle Recommendation Based on CelebA. In: *International Conference on Knowledge Science, Engineering and Management.* Springer.

Cowen, A.S., Keltner, D., Schroff, F., Jou, B., Adam, H., and Prasad, G., 2021. Sixteen facial expressions occur in similar contexts worldwide. *Nature*, 589(7841), pp.251-257.

Eisenthal, Y., Dror, G., and Ruppin, E., 2006. Facial attractiveness: Beauty and the machine. *Neural Computation*, 18(1), p.119-142.

Fan, D., Kim, H., Kim, J., Liu, Y., and Huang, Q., 2019. Multi-task learning using task dependencies for face attributes prediction. *Applied Sciences*, 9(12), p.2535.

Favorskaya, M.N., and Pakhirka, A.I., 2023. Age-Group Estimation of Facial Images Using Multi-task Ranking CNN. In: I*nternational KES Conference on Intelligent Decision Technologies*. Springer.

Gan, J., Jiang, K., Tan, H., and He, G., 2020. Facial beauty prediction based on lighted deep convolution neural network with feature extraction strengthened. *Chinese Journal of Electronics*, 29(2), pp.312-321.

Gan, J., Luo, H., Xiong, J., Xie, X., Li, H., and Liu, J., 2023. Facial beauty prediction combined with multi-task learning of adaptive sharing policy and attentional feature fusion. *Electronics*, 13(1), p.179.

Gan, J., Xiang, L., Zhai, Y., Mai, C., He, G., Zeng, J., Bai, Z., Labati, R., Piuri, V., and Scotti, F., 2020. 2M BeautyNet: Facial beauty prediction based on multi-task transfer learning. *IEEE Access*, 8, pp.20245-20256.

Gan, J., Xie, X., Zhai, Y., He, G., Mai, C., and Luo, H., 2023. Facial beauty prediction fusing transfer learning and broad learning system. *Soft Computing*, 27(18), pp.13391-13404.

Gao, L., Li, W., Huang, Z., Huang, D., and Wang, Y., 2018. Automatic Facial Attractiveness Prediction by Deep Multi-task Learning. In: *2018 24th International Conference on Pattern Recognition (ICPR*). IEEE.

Gray, D., Yu, K., Xu, W., and Gong, Y., 2010. Predicting Facial Beauty Without Landmarks. In: *Computer Vision-ECCV 2010: 11th European Conference on Computer Vision. Proceedings, Part VI 11*. Springer, Heraklion, Crete, Greece.

Grgic, M., Delac, K., and Grgic, S., 2011. SCface-surveillance cameras face database. *Multimedia Tools and Applications*, 51, pp.863-879.

Gunes, H., and Piccardi, M., 2006. Assessing facial beauty through proportion analysis by image processing and supervised learning. *International Journal of Human-computer Studies*, 64(12), pp.1184-1199.

Jamoliddin, U., and Yoo, J.H., 2022. Age and gender classification with small scale cnn. *The Journal of the Korea Institute of Electronic Communication Sciences*, 17(1), pp.99-104.

Kalyta, O., Krakb, I., Barmaka, O., Wojcikd, W., and Radiuk, P., 2022. Method of Facial Geometric Feature Representation for Information Security Systems. In: *3rd International Workshop on Intelligent Information Technologies & Systems of Information Security.* Khmelnytskyi, Ukraine.

Lebedeva, I., Guo, Y., and Ying, F., 2021. Transfer learning adaptive facial attractiveness assessment. *Journal of Physics: Conference Series*, 1922, p.012004.

Lebedeva, I., Guo, Y., and Ying, F., 2022. MEBeauty: A multi-ethnic facial beauty dataset in-the-wild. *Neural Computing and Applications*, 34, pp.14169-14183.

Lebedeva, I., Guo, Y., and Ying, F., 2023. Personalized facial beauty assessment: A meta-learning approach. *The Visual Computer*, 39(3), pp.1095-1107.

Liang, L., Xie, D., Jin, L., Xu, J., Li, M., and Lin, L., 2017. Region-aware Scattering Convolution Networks for Facial Beauty Prediction. In: *2017 IEEE International Conference on Image Processing (ICIP*). IEEE.

Lin, L., Shen, Z., Yin, J.L., Liu, Q., Yu, Y., and Chen, W., 2023. MetaFBP: Learning to Learn High-Order Predictor for Personalized Facial Beauty Prediction. In: *Proceedings of the 31st ACM International Conference on Multimedia.*

Lin, Y., Zhibin, G., Zhang, S., Li, L., and Huang, L., 2021. Multi-Task Network and Optimization for Face Detection and Attribute Analysis. In: *2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA*). IEEE.

Liu, Z., Luo, P., Wang, X., and Tang, X., 2015. Deep learning face attributes in the wild. In: *Proceedings of the IEEE International Conference on Computer Vision*. IEEE.

Mao, L., Yan, Y., Xue, J.H., and Wang, H., 2020. Deep multi-task multi-label CNN for effective facial attribute classification. *IEEE Transactions on Affective Computing*, 13(2), pp.818-828.

Moridani, M.K., Jamiee, N., and Saghafi, S., 2023. Human-like evaluation by facial attractiveness intelligent machine. *International Journal of Cognitive Computing in Engineering*, 4, pp.160-169.

Mu, Y., 2013. Computational facial attractiveness prediction by aesthetics-aware features. *Neurocomputing*, 99, pp.59-64.

Nguyen, T.V., Liu, S., Ni, B., Tan, J., Rui, Y., and Yan, S., 2013. Towards decrypting attractiveness via multi-modality cues. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 9(4), p.28.

Panić, N., Marjanović, M., and Bezdan, T., 2024. Ethnic representation matters: Investigating bias in facial age prediction models. *Mathematics*, 12(15), pp.1-30.

Rohani, M., Farsi, H., and Mohamadzadeh, S., 2023. Deep multi-task convolutional neural networks for efficient classification of face attributes. *International Journal of Engineering*, 36(11), pp.2102-2111.

Saeed, J.N., Abdulazeez, A.M., and Ibrahim, D.A., 2022. 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.

Saeed, J.N., and Abdulazeez, A.M., 2021. Facial beauty prediction and analysis based on deep convolutional neural network: A review. *Journal of Soft Computing and Data Mining*, 2(1), pp.1-12.

Saeed, J.N., and Abdulazeez, A.M., 2022. 2D Facial Images Attractiveness Assessment Based on Transfer Learning of Deep Convolutional Neural Networks. In: *2022 4th International Conference on Advanced Science and Engineering (ICOASE*). IEEE.

Sagonas, C., Antonakosa, E., Tzimiropoulosb, G., Zafeirioua, S., and Pantic, M., 2016. 300 faces in-the-wild challenge: Database and results. *Image and Vision Computing*, 47, pp.3-18.

Savchenko, A.V., 2021. Facial Expression and Attributes Recognition Based on Multi-task Learning of Lightweight Neural Networks. In: *2021 IEEE 19th International Symposium on Intelligent Systems and Informatics (SISY*). IEEE, pp.119-124.

Schmid, K., Marx, D., and Samal, A., 2008. Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios. *Pattern Recognition*, 41(8), pp.2710-2717.

Sekhar, J.C., Joel Josephson, P., Chinnasamy, A., Maheswari, V., Sankar, V., and Kalangi, R.R., 2024. Automated face recognition using deep learning technique and center symmetric multivariant local binary pattern. *Neural Computing and Applications*, pp.1-19.

Štěpánek, L., Kasal, P., and Mestak, J., 2018. Evaluation of Facial Attractiveness for Purposes of Plastic Surgery Using Machine-learning Methods and Image

Analysis. In: *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services*. IEEE.

Taherkhani, F., Dabouei, A., Soleymani, S., Dawson, J., and Nasrabadi, N.M., 2021. Tasks structure regularization in multi-task learning for improving facial attribute prediction. arXiv preprint arXiv:2108.04353.

Vahdati, E., and Suen, C.Y., 2021. Facial beauty prediction from facial parts using multi-task and multi-stream convolutional neural networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 35(12), p.2160002.

Wang, F., Han, H., Shan, S., and Chen, X., 2017. Deep Multi-task Learning for Joint Prediction of Heterogeneous Face Attributes. In: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017*). IEEE.

Weng, N., Wang, J., Li, A., and Wang, Y., 2021. Two-stream Temporal Convolutional Network for Dynamic Facial Attractiveness Prediction. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE.

Xiao, Q., Wu, Y., Wang, D., Yang, Y.L., and Jin, X., 2021. Beauty3DFaceNet: Deep geometry and texture fusion for 3D facial attractiveness prediction. *Computers & Graphics*, 98, pp.11-18.

Xie, D., Liang, L., Jin, L., Xu, J., and Li, M., 2015. Scut-fbp: A Benchmark Dataset for Facial Beauty Perception. In: *2015 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE.

Xie, Y., Wang, K., Meng, J., Yue, J., Meng, L., Yi, W., Jung, T.P., Xu, M., and Ming, D., 2023. Cross-dataset transfer learning for motor imagery signal classification via multi-task learning and pre-training. *Journal of Neural Engineering*, 20(5), p.056037.

Xu, L., Xiang, J., and Yuan, X., 2018. CRnet: Classification and Regression Neural Network for Facial Beauty Prediction. *In: Pacific Rim Conference on Multimedia*. Springer, Cham.

Xu, M., Chen, F., Li, L., Shen, C., Lv, P., Zhou, B., and Ji, R., 2018. Bio-inspired deep attribute learning towards facial aesthetic prediction. *IEEE Transactions on Affective Computing*, 12(1), pp.227-238.

Yin, X., Tai, Y., Huang, Y., and Liu, X., 2020. Fan: Feature Adaptation Network for Surveillance Face Recognition and Normalization. In: *Proceedings of the Asian Conference on Computer Vision.*

Yuan, H., He, Y., Du, P., and Song, L., 2024. *Multi-Task Learning Using Uncertainty to Weigh Losses for Heterogeneous Face Attribute Estimation*. [Preprint].

Zhai, Y., Huang, Y., Xu, Y., Gan, J., Cao, H., Deng, W., Labati, R.D., Piuri, V., and Scotti, F., 2020. Asian female facial beauty prediction using deep neural networks via transfer learning and multi-channel feature fusion. *IEEE Access*, 8, pp.56892-56907.

Zhai, Y., Huang, Y., Xu, Y., Zeng, J., Yu, F., and Gan, J., 2016. Benchmark of a Large Scale Database For Facial Beauty Prediction. In: *Proceedings of the 1st International Conference on Intelligent Information Processing.*

Zhang, B., Xiao, X., and Lu, G., 2018. Facial beauty analysis based on features prediction and beautification models. *Pattern Analysis and Applications*, 21, pp.529-542.

Zheng, K., Tian, L., Li, Z., Li, H., and Zhang, J., 2024. Incorporating eyebrow and eye state information for facial expression recognition in mask-obscured scenes. *Electronic Research Archive*, 32(4), pp.2745-2771.