OpticalFBP: A Deep Optical Imaging Model for Privacy-Protective Facial Beauty Prediction*

Jiho Bae, Minjae Lee, and Suwon Lee†

Gyeongsang National University, South Korea
{dream_cacao_jh, wjdchs0129, leesuwon}@gnu.ac.kr

Abstract

Facial beauty prediction (FBP) often compromises personal privacy as it requires identifiable facial images. We propose **OpticalFBP**, a deep optical imaging framework that predicts facial beauty without capturing recognizable faces. A learnable optical lens is optimized to erase identity features while preserving beauty-related cues. The system is trained in two stages: (*i*) lens optimization with landmark guidance and (*ii*) beauty regression on privacy-free images. **OpticalFBP** enables accurate and privacy-preserving beauty prediction at the imaging stage itself.

Keywords: Facial Beauty Prediction, Privacy-free, Deep Optical Imaging

1 Introduction

Facial beauty prediction (FBP) [1] has gained increasing attention in computer vision and human–computer interaction due to its potential applications in cosmetic recommendation, aesthetic evaluation, and social media analysis. However, conventional FBP systems rely heavily on facial images captured by regular cameras, which inherently expose personal identity and raise significant privacy concerns. Once a facial image is recorded, it can be easily misused or leaked, resulting in irreversible identity disclosure. Inspired by the concept of optical privacy-preserving imaging introduced in OpticalDR [2], we aim to move privacy protection from the digital stage to the physical level. Instead of acquiring facial image and then processing it with anonymization algorithms, our proposed system directly captures privacy-free optical images through a learnable lens.

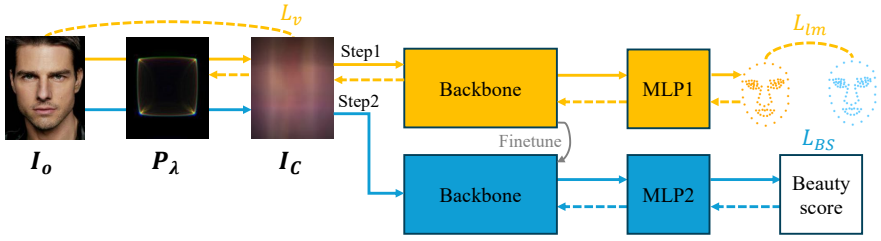

Figure 1: OpticalFBP Overview

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-17, December 16-18, 2025, Sapporo, Japan.

†Corresponding author

2 Method

As illustrated in Fig. 1, the proposed **OpticalFBP** framework aims to achieve privacy-protective facial beauty prediction by integrating a learnable optical lens with a deep regression network. Through trained optical lenses, we remove personally identifiable information during the image capture process and utilize the deidentified images in the facial attractiveness prediction model. To achieve this, we introduce a two-step learning process.

Step 1: Learning a Privacy-Preserving Optical Lens: The goal of this stage is to learn a lens that erases identifiable facial details while retaining geometric and structural cues relevant to beauty perception. As proposed by OpticalDR [2], we parameterize the optical lens as a differentiable imaging module that simulates light propagation and image formation on the sensor. Specifically, the Zernike polynomial coefficients are treated as learnable parameters to train the lens surface to prevent human identification. The optical lens is jointly optimized with a backbone under twoz complementary objectives: (i) visual dissimilarity from the original image, and (ii) preservation of facial structure via landmark consistency. The overall optimization target can be formulated as:

$$\mathcal{L}_{\text{ol}} = -\|I_c - I_o\|_2^2 + \frac{\beta}{N}\sum_{i=1}^{N}\|\hat{\mathbf{p}}_i - \mathbf{p}_i\|_2^2, \tag{1}$$

where I_c denotes the image captured through the learnable optical system, and I_o is the original RGB image. β is a weighting factor balancing privacy and structural fidelity. $\hat{\mathbf{p}}_i$ and $\mathbf{p}_i$ represent the predicted and ground-truth coordinates of the i-th landmark, respectively, and N is the total number of landmarks. The first term increases the distance between the two images, training the lens to capture images that prevent human identification. The second term preserves structural consistency by aligning the predicted facial landmarks $\hat{\mathbf{p}}_i$ with the ground-truth landmarks $\mathbf{p}_i$. Through this process, we remove human recognition information while preserving the facial structural information necessary for FBP.

Step 2: Facial Beauty Regression: In the second stage, the optical parameters are fixed, and the generated privacy-free images are used to train a facial beauty prediction network. A convolutional encoder(backbone) extracts visual representations from I_c, followed by a regression head to estimate the predicted beauty score $\hat{y}$. The model is trained using the mean squared error (MSE) between the predicted and ground-truth scores:

$$L_{\text{bs}} = \|\hat{y} - y\|_2^2, \tag{2}$$

where y represents the ground-truth beauty score. This simple yet effective objective allows the model to learn aesthetic representations directly from privacy-preserving optical inputs without requiring any access to identifiable data. Through the above two-step optimization, **OpticalFBP** enables end-to-end privacy-preserving beauty prediction at the imaging stage itself. The proposed framework ensures that personally identifiable facial features are never captured or stored, while enabling aesthetic evaluation.

References

[1] Lingyu Liang, Luojun Lin, Lianwen Jin, Duorui Xie, and Mengru Li. Scut-fbp5500: A diverse benchmark dataset for multi-paradigm facial beauty prediction. In *2018 24th International conference on pattern recognition (ICPR)*, pages 1598–1603. IEEE, 2018.

[2] Yuchen Pan, Junjun Jiang, Kui Jiang, Zhihao Wu, Keyuan Yu, and Xianming Liu. Opticaldr: A deep optical imaging model for privacy-protective depression recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1303–1312, 2024.