Paper 88-1 has been designated as a Distinguished Paper at Display Week 2025. The full-length version of this paper appears in a Special Section of the *Journal of the Society for Information Display (JSID)* devoted to Display Week 2025 Distinguished Papers. This Special Section will be freely accessible until December 31, 2025 via:

https://sid.onlinelibrary.wiley.com/doi/full/10.1002/jsid.2090

Authors that wish to refer to this work are advised to cite the full-length version by referring to its DOI:

https://doi.org/10.1002/jsid.2090

Enhancing Face Recognition Accuracy for Under-Display Cameras via Image Restoration

Kyusu Ahn†,*, Chanwoo Park‡, and Jaejin Lee†,‡
†Dept. of Data Science, Seoul National University, Republic of Korea
‡Dept. of Computer Science and Engineering, Seoul National University, Republic of Korea
*CAE Team, Research Center, Samsung Display Co., Ltd., Yongin, Republic of Korea

Abstract
Under-display Camera (UDC) allows for a full-screen display by placing the camera beneath the screen, eliminating the need for display holes. However, diffraction caused by display pixels results in the degradation of UDC image quality, including reduced transmittance, noise, blur, and flare. This leads to decreased facial recognition accuracy, compromising UDC's practicality. While recent research has focused on image restoration for UDC, studies on UDC facial recognition remain insufficient. Previous UDC face recognition studies use synthetic UDC datasets without flare to train degradation models and apply them to high-quality datasets of human faces. However, they do not address real UDC datasets or Face ID performance, a common smartphone application. This paper presents three UDC facial recognition datasets generated by a generative adversarial network (GAN) trained on synthetic and real-world datasets with flares. We aim to analyze the impact of UDC restoration on Face ID accuracy, contributing to the expansion of UDC face recognition research.

Author Keywords
Under-Display Camera; Under-Panel Camera; Face Recognition; Image Restoration; Deep Neural Networks

1. Introduction
Under-display Camera (UDC) enables a full-screen display by embedding the camera beneath the display panel, addressing the demand for higher screen-to-body ratios. However, this innovation presents significant challenges, including complex degradation issues such as noise, blur, reduced transmittance, and flare, all resulting from the diffraction and attenuation of light caused by the display's pixel array.

Several studies have addressed UDC degradation through two primary approaches: developing UDC datasets [1-4] and proposing UDC image restoration models [4-7]. Despite these advancements, many restoration models have yet to leverage the real-world UDC dataset, UDC-SIT [1]. While Ahn et al. propose SFIM [7] to restore actual degraded images, their work does not extend to face recognition (e.g., pair matching). Although UDC-degraded face images have been studied, these areas remain underexplored. Tan et al. introduce UDC-DMNet [8], a two-stage model for learning UDC degradation using the P-OLED dataset [2]. They synthesize UDC-degraded face images by applying UDC-DMNet to high-quality face images from FFHQ [9] and CelebA [10]. However, they focus on restoration rather than applying them to face recognition. Similarly, Wang et al. [11] use MPGNet [12], trained on the P-OLED dataset, to synthesize UDC-degraded face images from facial expression recognition (FER) datasets such as RAF-DB [13], FERPlus [14], and KDEF [15]. They further propose LRDif [11] to enhance the performance in UDC-FER tasks, predicting emotions like surprise, fear, disgust, happiness, sadness, anger, and neutrality. However, they do not address face pair matching, the most common application in modern smartphones.

This paper introduces three UDC-degraded face recognition datasets, leveraging MPGNet [12] to model UDC-specific degradations from the P-OLED, SYNTH, and UDC-SIT datasets. Using the pre-trained MPGNet, we apply these learned degradations to the high-quality Labeled Faces in the Wild (LFW) dataset [16] to create UDC-degraded face recognition datasets. A notable aspect of our work is the inclusion of flare, a key characteristic of UDCs in both the SYNTH and UDC-SIT datasets. The UDC-SIT dataset, in particular, uniquely mirrors real-world UDC artifacts [1], making it critical for advancing research in UDC face recognition. Unlike prior studies [11], which primarily model the P-OLED dataset's degradations for face emotion recognition, our approach focuses on face recognition in a more realistic UDC environment by incorporating both synthetic and real-world UDC degradations, as seen in the UDC-SIT dataset.

The contributions of this paper are as follows:

- We propose three synthetic UDC face recognition datasets generated using MPGNet. To the best of our knowledge, this is the first work to explore UDC face recognition, focusing on pair matching.
- Through extensive experiments, we demonstrate that restoring UDC-degraded images enhances face recognition accuracy significantly.
- We highlight the importance of advancing UDC face recognition research using a real-world UDC dataset.

2. Background
UDC Datasets. Extensive research has been conducted on UDC datasets. The T-OLED/P-OLED datasets [2] capture paired images displayed on a monitor, but flares are almost absent due to the monitor's limited dynamic range. The SYNTH dataset [4] generates flares by convolving the measured point spread function (PSF) of the ZTE Axon 20 with clean images, but it lacks noise and spatially variant flares. Feng et al. [3] introduce a pseudo-real dataset by capturing paired images of similar scenes with two cameras (e.g., ZTE Axon 20 UDC and iPhone 13 Pro). However, this approach leads to geometric misalignment, and alignment accuracy remains a challenge despite their attempt to correct it using AlignFormer [3]. In contrast, Ahn et al. [1] propose a real-world UDC-SIT dataset with a dedicated image-capturing system. Based on discrete Fourier transform (DFT), their alignment method enables high accuracy by effectively correcting geometric misalignments. The pseudo-real and UDC-SIT datasets capture realistic UDC degradation, including spatially variant flares.

UDC Facial Datasets. Wang et al. [11] and Tan et al. [8] propose UDC datasets containing human faces generated by a GAN-based model trained on the T-OLED/P-OLED datasets [2]. However, these datasets fail to capture actual UDC degradations, particularly flares. Tan et al. focus solely on image restoration without addressing face-related tasks, while Wang et al. propose LRdif to

ISSN 0097-996X/25/5602-1218-$1.00 © 2025

evaluate UDC-FER performance using the flare-free P-OLED dataset, neglecting face pair matching tasks.

UDC Restoration. There are many UDC restoration models to leverage UDC degradations. DISCNet [4] incorporates the domain knowledge (i.e., PSF) of the UDC image formation model as prior information to guide the diffraction removal. UDC-UNet [5] incorporates a condition branch for spatially variant modulation and a kernel branch that utilizes prior PSF knowledge. ECFNet [6] takes input images at multiple scales, gradually producing results that transition from coarse to fine detail. Although not explicitly designed for UDC restoration, we will also use SRGAN [17], which enhances low-resolution images, to restore UDC-degraded face images in this research.

3. UDC Face Recognition Dataset

This section outlines the definition of face recognition and details the method used to generate the UDC face recognition dataset.

3.1. Face Recognition

Face recognition entails comparing two faces to establish a match. It is typically used to confirm an individual's identity against an identification document. A well-known example is the Face ID feature in smartphones, which is widely utilized for unlocking devices and enabling secure transactions. Recently, financial institutions have started to adopt this technology, underscoring the importance of accuracy in face recognition for identity authentication across various sectors.

3.2. Method

We employ MPGNet [12] to generate UDC-degraded face recognition datasets, following a two-step process illustrated in Figure 1. In the first step, MPGNet is trained on three UDC datasets: P-OLED [2], SYNTH [3], and UDC-SIT [1]. This training results in three distinct pre-trained MPGNet models, each capturing the unique degradation characteristics of the respective datasets. In the second step, these pre-trained models degrade high-quality images from the LFW dataset [16]. The outputs are three UDC-degraded face recognition datasets: LFW-P-OLED, LFW-SYNTH, and LFW-UDC-SIT, each corresponding to the degradation learned from the P-OLED, SYNTH, and UDC-SIT datasets, respectively.

As shown in Figure 1, the LFW-P-OLED dataset shows an excessive transmittance decrease. It is essential to highlight that only LFW-SYNTH and LFW-UDC-SIT datasets exhibit flare artifacts around light sources, with LFW-UDC-SIT uniquely capturing additional degradation such as noise, blur, and reduced transmittance. These features of the UDC-SIT dataset reflect real-world UDC degradation, as detailed by Ahn et al. [1]. Given that these unique UDC artifacts can impact face recognition accuracy, conducting research on face recognition under actual UDC degradation is crucial for advancing reliable identity authentication in UDC environments.

4. Experiments

This section emphasizes the importance of UDC image restoration in enhancing face recognition tasks. First, we train deep neural networks to restore UDC-degraded face images. Then, we apply these pre-trained models to various UDC-degraded face datasets. Finally, we evaluate the impact of restoration on face recognition accuracy by comparing results before and after restoration.

4.1. UDC Image Restoration

We train ECFNet, UDC-UNet, DISCNet, and SRGAN on the LFW-P-OLED, LFW-SYNTH, and LFW-UDC-SIT datasets and assess their restoration performance on their respective test sets.

4.2. UDC Face Recognition

Face recognition encompasses various tasks, including but not limited to the following [16]:

- **Face Verification:** Given a face image, identify the individual from known identities.
- **Pair Matching:** Given two face images, determine if they belong to the same person.

In our study, we focus on the pair-matching task and use four face recognition models from the DeepFace library [18]: VGG-Face [19], Facenet [20], Dlib [21], and ArcFace [22].

4.3. Quantitative Results

We assess face recognition accuracy against PSNR, SSIM, and LPIPS. The average recognition accuracy is calculated across 1,000 image pairs, comparing the face recognition performance between

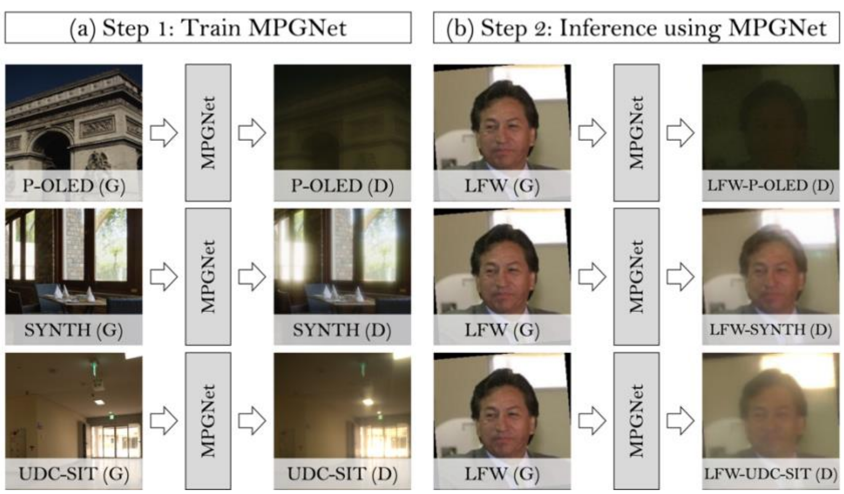

Figure 1. Schematic illustration of the process for generating UDC-degraded face recognition datasets. The process involves two main steps. (a) Training MPGNet [12] on UDC datasets, such as P-OLED [2], SYNTH [3], and UDC-SIT [1], where (G) and (D) denote ground-truth and degraded images, respectively. (b) Using the pre-trained MPGNet from step 1 to generate UDC-degraded datasets by applying it to the LFW dataset [16].

human 1 (GT, Input, or Restored) and human 2 (ground truth (GT)).

The results in Figure 2 illustrate the significant impact of UDC degradation on face recognition accuracy. For instance, as shown in Figure 2 (a), (d), and (g), Input images with PSNRs of 9.21, 21.16, and 17.00 achieve recognition accuracies of 50.8%, 72.8%, and 48.7%, respectively. In contrast, the GT images yield an accuracy of 88.3%, while Restored images from ECFNet with PSNRs of 28.37, 37.32, and 32.07 achieve recognition accuracies of 72.7%, 84.8%, and 78.9%.

The P-OLED and SYNTH datasets differ from the UDC-SIT dataset regarding UDC degradation realism. For example, the P-OLED dataset exhibits excessive transmittance reduction, while the SYNTH dataset lacks sufficient noise and transmittance reduction. As a result, the Input's PSNR of the P-OLED dataset is remarkably lower than that of the UDC-SIT dataset, while the Input's PSNR of the SYNTH dataset is relatively higher than that of the UDC-SIT dataset. Performance on the SYNTH dataset across different models is more consistent than the UDC-SIT dataset, leading to similar face recognition accuracies. This highlights the need for research using actual UDC-degraded face datasets to gauge face recognition performance accurately under actual UDC conditions.

5. Limitations

Prior research uses GANs trained on the P-OLED dataset to create face datasets, but this approach lacks UDC-specific degradations. To address this, we train MPGNet on the SYNTH and UDC-SIT datasets, which better represent UDC artifacts. However, since these datasets are not from real-world scenes, they cannot fully replicate actual degradation. Furthermore, the LFW dataset has limited images containing the light source, restricting our ability to assess flare impact on face recognition accuracy.

6. Conclusion

This paper addresses UDC-degraded face recognition, focusing on pair matching by generating datasets that better reflect actual UDC artifacts. Prior work on FER tasks with synthetic datasets overlooks key degradations like spatially variant flares. We leverage the SYNTH and UDC-SIT datasets with MPGNet to create realistic datasets, significantly improving the representation of UDC effects despite not being real-world captures. Our evaluation shows how UDC-specific artifacts affect recognition accuracy, particularly in pair matching, the most common UDC application. This highlights

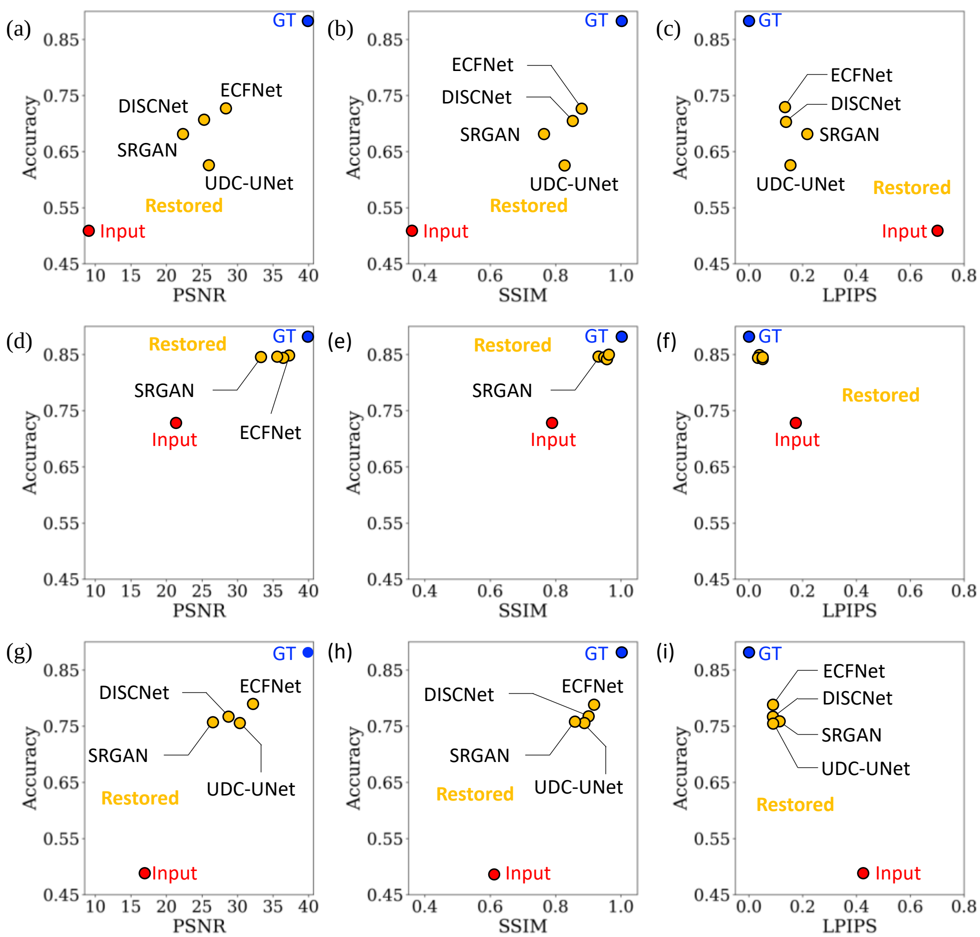

Figure 2. Face recognition accuracy. Face recognition accuracy of Input refers to the comparison between Input (human 1) and GT (human 2), that of GT refers to the comparison between GT (human 1) and GT (human 2), and that of a deep neural network (e.g., ECFNet) refers to the comparison between Restored by ECFNet (human 1) and GT (human 2). Images restored by deep neural networks with better performance in (a) PSNR, (b) SSIM, and (c) LPIPS tend to achieve higher recognition accuracy. The restoration metrics (e.g., PSNR, SSIM, and LPIPS) are evaluated between human 1 (Input, GT, or Restored) and human 1 (GT). The PSNR value between the two GTs is plotted as 40.00 for easy observation. (a)-(c) LFW-P-OLED dataset. (d)-(f) LFW-SYNTH dataset. (g)-(i) LFW-UDC-SIT dataset.

the need for further research into UDC face recognition for reliable identity authentication in real-world settings.

7. Acknowledgment

This work was supported in part by the National Research Foundation of Korea (NRF) grant (No. RS-2023-00222663, Center for Optimizing Hyperscale AI Models and Platforms), by the Institute for Information and Communications Technology Promotion (IITP) grant (No. 2018-0-00581, CUDA Programming Environment for FPGA Clusters), by the BK21 Plus programs for Innovative Data Science Talent Education Program (Dept. of Data Science, SNU, No. 5199990914569) and BK21 FOUR Intelligence Computing (Dept. of Computer Science and Engineering, SNU, No. 4199990214639) through the NRF, all funded by the Ministry of Science and ICT (MSIT) of Korea. This work was also supported in part by the Samsung Display Co., Ltd. ICT at Seoul National University provided research facilities for this study.

8. References

[1] Ahn, K., Ko, B., Lee, H., Park, C. and Lee, J., "UDC-SIT: a real-world dataset for under-display cameras," *Advances in Neural Information Processing Systems* 36, Article No. 2962, pp. 67721-67740, 2024.

[2] Zhou, Y., Ren, D., Emerton, N., Lim, S. and Large, T., "Image restoration for under-display camera," *In Proceedings of the ieee/cvf conference on computer vision and pattern recognition*, pp. 9179-9188, 2021.

[3] Feng, R., Li, C., Chen, H., Li, S., Gu, J. and Loy, C.C., "Generating aligned pseudo- supervision from non-aligned data for image restoration in under-display camera," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5013-5022, 2023.

[4] Feng, R., Li, C., Chen, H., Li, S., Loy, C.C. and Gu, J., "Removing diffraction image artifacts in under-display camera via dynamic skip connection network," *In Proceedings of the ieee/cvf conference on computer vision and pattern recognition*, pp. 662-671, 2021.

[5] Liu, X., Hu, J., Chen, X. and Dong, C., "Udc-unet: Under-display camera image restoration via u-shape dynamic network," *In European Conference on Computer Vision*, pp. 113-129, 2022.

[6] Zhu, Y., Wang, X., Fu, X. and Hu, X., "Enhanced coarse-to-fine network for image restoration from under-display cameras," *In European Conference on Computer Vision*, pp. 130-146, 2022.

[7] Ahn, K., Kim, J., Park, C., Kim, J. and Lee, J., "Integrating Spatial and Frequency Information for Under-Display Camera Image Restoration," *arXiv preprint arXiv:2501.18517*, 2025.

[8] Tan, J., Chen, X., Wang, T., Zhang, K., Luo, W. and Cao, X., "Blind face restoration for under-display camera via dictionary guided transformer," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

[9] Karras, T., Laine, S. and Aila, T., "A style-based generator architecture for generative adversarial networks," *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401-4410, 2019.

[10] Liu, Z., Luo, P., Wang, X. and Tang, X., "Deep learning face attributes in the wild," *In Proceedings of the IEEE international conference on computer vision*, pp. 3730-3738, 2015.

[11] Wang, Z., Zhang, K. and Sankaranarayana, R., "LRDif: Diffusion Models for Under-Display Camera Emotion Recognition," *arXiv*, 2024.

[12] Zhou, Y., Song, Y. and Du, X., "Modular degradation simulation and restoration for under-display camera," *In Proceedings of the Asian Conference on Computer Vision*, pp. 265-282, 2022.

[13] Li, S., Deng, W. and Du, J., "Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2852-2861, 2017.

[14] Barsoum, E., Zhang, C., Ferrer, C.C. and Zhang, Z., "Training deep networks for facial expression recognition with crowd-sourced label distribution," *In Proceedings of the 18th ACM international conference on multimodal interaction*, pp. 279-283, 2016.

[15] Calvo, M.G. and Lundqvist, D., "Facial expressions of emotion (KDEF): Identification under different display-duration conditions," *Behavior research methods*, 40(1), pp.109-115, 2008.

[16] Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," *In Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*, 2008.

[17] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W., "Photo-realistic single image super-resolution using a generative adversarial network," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681-4690, 2017.

[18] J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: http://www.atm.com (downloaded 2012, Sep. 10)

[19] Serengil, S., "DeepFace: A Popular Open Source Facial Recognition Library," [Online]. Available: https://viso.ai/computer-vision/deepface/

[20] Schroff, F., Kalenichenko, D. and Philbin, J., "Facenet: A unified embedding for face recognition and clustering," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815-823, 2015.

[21] King, D.E., "Dlib-ml: A machine learning toolkit," *The Journal of Machine Learning Research*, 10, pp.1755-1758, 2009.

[22] Deng, J., Guo, J., Xue, N. and Zafeiriou, S., "Arcface: Additive angular margin loss for deep face recognition," *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4690-4699, 2019.