Image Restoration for Under-Display Cameras: A Review of Current Technologies

Jewon Yoo*, Kyusu Ahn*, Jebum Cho*, Jong Beom Hong*, and Seungin Baek*
***CAE Team, Research Center, Samsung Display Co., Ltd., Yongin, Republic of Korea**

Abstract
Under-Display Cameras (UDCs) are a promising technology that integrates cameras beneath the display, enabling a larger screen-to-body ratio. However, positioning the camera beneath the display presents distinct challenges, as the light must pass through multiple display layers, leading to image and video degradation. These degradations include reduced transmittance, color distortion, blurring, noise, and flare, significantly impacting the quality of captured images or videos. These degradations significantly impact applications like photography, video recording, and facial recognition, which require high image accuracy and clarity. This paper comprehensively surveys recent technologies to mitigate various UDC degradations. We categorize the research into two key perspectives: datasets and deep-learning models. We explore studies focusing on creating synthetic and real-world UDC datasets that capture UDC degradation characteristics, facilitating research on UDC restoration methods. Additionally, we examine advancements in deep-learning techniques designed to restore UDC-distorted image and video quality, highlighting future opportunities for improving UDC performance.

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

1. Introduction
Under-Display Camera (UDC) is a novel technology to maximize the screen-to-body ratio, addressing consumer demand for immersive, bezel-free displays. Unlike traditional designs using notches or punch holes, UDC integrates the camera sensor beneath the display panel, as shown in Figure 1.

In this setup, the display fully covers the camera, functioning as part of the screen during regular use. When the camera is activated, the display section above the sensor switches off, allowing light to pass through to the camera. The pixel density in the UDC area is reduced to support light transmission compared to other regions. However, even with reduced density, the display pixels cause light diffraction, resulting in image and video quality degradation.

There has been active research to address UDC degradation, which can be categorized into two paradigms. The former focuses on creating UDC datasets to simulate actual UDC distortion [1-2, 6, 8, 11, 14, 19-21, 23]. The latter emphasizes the development of deep-learning models designed to restore UDC-degraded images, focusing on enhancing visual quality and maintaining critical features [2, 23, 25-32].

This paper aims to provide a comprehensive review of UDC research, offering a valuable reference for understanding the field's current state and identifying potential directions for future improvement. The rest of the survey is organized as follows. Section 2 provides a comprehensive overview of UDC degradation, detailing its formulation and the various factors that cause image distortion. Section 3 describes various datasets used in UDC research, including image, video, and face recognition datasets. Section 4 explores the deep-learning models to tackle UDC degradations. Section 5 summarizes the paper and outlines future directions.

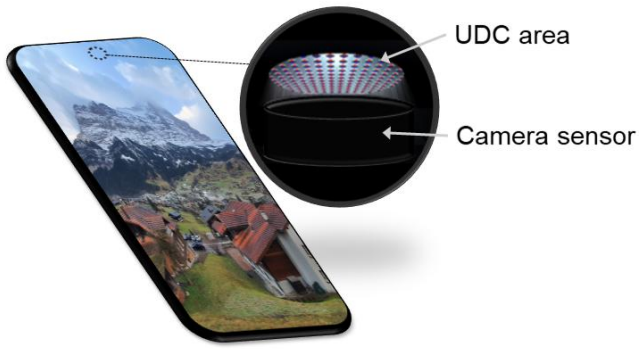

Figure 1. Illustration of UDC technology. The display region above the camera sensor features a lower pixel density than the normal area. Since light passes through the display layers, it undergoes diffraction, which results in image and video quality degradation.

2. UDC Degradation and Formulation
Many studies represent UDC degradation with the following formulation:

$$y = (\gamma * x) \otimes k + n, \tag{1}$$

where y represents the UDC-degraded image, and x is the clean image. The degradation is modeled by γ, k, and n. The light attenuation, denoted as γ, is caused by the layers stacked on the display through absorption, scattering, and reflection. Diffraction arises from the display's pixel grid structure, which affects the light passing through and introduces the point spread function (PSF), represented by k. The PSF is convolved with the image, attenuated by γ, to simulate diffraction. Noise, expressed as n, is added to the resulting image. Here, $\otimes$ denotes the convolution operation. The representative UDC degradation types are outlined below.

- **Noise**: Due to low light conditions, both the desired signal and unintended noise are amplified.
- **Blur**: Diffraction causes the image to lose sharpness and detail, resulting in blurring that makes fine features more difficult to distinguish.
- **Transmittance decrease**: The display layers, which absorb and scatter light, reduce the brightness.
- **Flare**: In the UDC setting, the light interacts with the display panel above the camera sensor, leading to unwanted reflections, diffraction, or scattering.

These degradation factors collectively contribute to the challenges in improving UDC image quality and pose significant obstacles for tasks such as face recognition, object detection, and general image enhancement.

ISSN 0097-996X/25/5602-1214-$1.00 © 2025

3. UDC Datasets

This section demonstrates the existing UDC datasets, such as the image, face recognition, and video datasets. Since the actual depiction of UDC degradation is critical, we delve into aspects such as noise, blur, reduced transmittance, and flare. Specifically, accurately modeling spatially variant flares, light-source dependent flares and temporally varying flares is essential for advancing research on UDC restoration.

3.1. Still Image Datasets

Table 1 presents four publicly available UDC image datasets. These datasets provide paired images, including ground-truth (GT) and UDC-degraded counterparts, designed for training.

Zhou et al. [1] introduce the first UDC image dataset, T-OLED and P-OLED, alongside an image acquisition system known as the Monitor Camera Imaging System (MCIS). They display images on a 4K Liquid Crystal Display (LCD) monitor and capture them with a static camera. To capture GT images, the camera is covered with a thin glass matching the thickness of an Organic Light-Emitting Diode (OLED) panel, ensuring accurate pixel alignment without additional adjustments. UDC-degraded images are generated by placing display panels, such as transparent-OLED (T-OLED) or pentile-OLED (P-OLED), in front of the camera. While the T-OLED and P-OLED datasets are pioneering work to stimulate UDC research, they lack the actual UDC degradations, such as the distinct UDC flare characteristics.

Feng et al. [2] improve the UDC dataset. They use High Dynamic Range (HDR) images from the HDRI Haven dataset [3] as the GT. They convolve the measured PSF with the GT images to simulate UDC degradation. The PSF is measured by positioning a white point light source one meter away from the OLED display of the ZTE Axon 20 [4]. The proposed SYNTH dataset depicts UDC flares; however, it does not capture the distorted flares addressed by Yoo et al. [5] and Kim et al. [30].

Feng et al. [6] propose a pseudo-real dataset by capturing GT images using an iPhone 13 Pro [7] camera and UDC-degraded images with ZTE Axon 20 UDC [4]. Due to the varying specifications of the two cameras, they encounter challenges with geometric misalignment and domain discrepancies. To address the geometric misalignment between the paired images, they propose AlignFormer; however, this approach introduces occlusion regions. The alignment accuracy, measured by the Percentage of Correct Keypoints (PCK), is 58.75% when $\alpha = 0.01$. PCK evaluates the ratio of correctly aligned keypoint pairs to the total number of keypoint pairs, where a pair is considered correctly aligned if $d < \alpha \times \max(H, W)$. Here, d represents the positional difference between matched keypoints, α is the threshold, and H and W denote the height and width of the image, respectively. The pseudo-real dataset accurately depicts the spatially variant flares.

Ahn et al. [8] propose a real-world dataset called UDC-SIT, which accurately depicts UDC degradation. They use an image-capturing system to capture paired images. They cut the UDC area of the Samsung Galaxy Z-Fold 3 [9] and attach it to a lid onto the Samsung Galaxy Note 10's standard camera [10]. Paired images are captured by alternately opening and closing the lid. Since the system uses a single camera, geometric misalignment is reduced compared to the pseudo-real dataset. However, unavoidable misalignment occurs due to the lid's movement. To mitigate this, Discrete Fourier Transform (DFT) is employed to align the two images, ensuring no artifacts (e.g., occlusions). The alignment accuracy measured by PCK is 97.26% when $\alpha = 0.01$. The UDC-SIT dataset effectively captures spatially variant flares.

Table 1. Representative UDC image datasets.

Datasets	Scene	Flare presence	Publication
T/P-OLED [1]	Synthetic		CVPR '21
SYNTH [2]	Synthetic	✔	CVPR '21
Pseudo-real [6]	Real	✔	CVPR '23
UDC-SIT [8]	Real	✔	NeurIPS '23

3.2. Face Recognition Datasets

Face recognition involves various tasks, including facial emotion recognition, face verification, and pair matching. Facial emotion recognition involves analyzing facial expressions to identify an individual's emotional state, such as happiness, sadness, anger, or surprise. Identity verification examines whether a specific face image corresponds to a predefined identity within a known set. In contrast, pair matching evaluates two face images to determine whether they represent the same person.

In the UDC setting, image degradation significantly impacts face recognition accuracy. Several studies have focused on UDC facial recognition datasets to address this challenge, as outlined in Table 2.

Tan et al. [11] use the GT facial images from the FFHQ [12] and CelebA [13] datasets. They generate UDC-degraded images by applying UDC-DMNet [9], trained on the P-OLED dataset [1], to the clean images. Similarly, Wang et al. [14] use MPGNet [15], also trained on the P-OLED dataset [1], to synthesize UDC-degraded facial images using GT images such as RAF-DB [16], FERPlus [17], and KDEF [18] datasets. Their work focuses on predicting emotions. These datasets, however, fail to capture actual UDC degradation characteristics such as flares.

Ahn et al. [19] use MPGNet [15], trained on the UDC-SIT dataset [8], to synthesize UDC-degraded facial images. Their approach aims to replicate actual UDC degradation using a real-world dataset; however, they also utilize a Generative Adversarial Network (GAN) for image synthesis. Ahn et al. [20] propose the UDC-VIT dataset, which includes 22 volunteers captured from various angles. They analyze the impact of UDC degradation and facial recognition accuracy.

3.3. Video Datasets

Table 3 describes the representative UDC video datasets. Chen et al. [21] propose the PexelsUDC-T and PexelsUDC-P datasets. To

Table 2. Representative UDC facial recognition datasets.

Datasets	Scene	Flare presence	Publication
Tan et al. [11]	Synthetic		TCSVT '23
Wang et al. [14]	Synthetic		arXiv '24
LFW-UDC [19]	Synthetic	✔	SID '25
UDC-VIT [20]	Real	✔	arXiv '25

Table 3. Representative UDC video datasets.

Datasets	Scene	Flare presence	Publication
PexelsUDC-T/P [21]	Synthetic		arXiv ’23
VidUDC33K [23]	Synthetic	✔	AAAI ’24
UDC-VIT [20]	Real	✔	arXiv ’25

obtain clean videos, they use videos from the Pexels database [22]. They propose a two-stage generation method to learn the UDC degradation from the T-OLED and P-OLED datasets [1]. The generator employs a customized U-Net as its core architecture. A light attenuation module adjusts the pixel value accordingly to incorporate the formulation in Equation (1). Furthermore, heteroscedastic noise is added to enhance the realism of the synthesized UDC degradations within the proposed framework.

Liu et al. [23] introduce the VidUDC33K dataset. Following Feng et al.’s approach [2], they convolve the measured PSF to the GT videos. Since the temporally variant flares are the unique characteristics of UDC videos, they simulate the dynamic changes of the PSF during the motion by computing the inter-frame Homography matrix $H_{t-1\to t}$, which is formulated as:

$$k_t = \mathrm{T}(\mathrm{k}_{\mathrm{t-1}}, H_{t-1\to t}) = \left|\mathcal{F}\left(\mathcal{H}_{t-1\to t}^{-1}\left(F^{-1}\left(sqrt(k_{t-1})\right)\right)\right)\right|^2, \quad (2)$$

$$\mathcal{H}_{t-1\to t}^{-1} = \mathcal{M}(I_{t-1}^{GT}, I_t^{GT}),$$

where k_t is the diffraction kernel (i.e., PSF), T is the transform functions using $\mathcal{H}_{t-1\to t}^{-1}$ to perspective warp the PSF of the previous frame k_{t-1}. $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ are the Fourier transform and inverse Fourier Transform, respectively. $\mathcal{M}(\cdot)$ indicates the matched part used to compute the Homography matrix between consecutive frames. Although they simulate the dynamic variations of the PSF, the changes in PSF and flare shapes with different light sources are not realistic, and temporally variant flares are often absent.

Ahn et al. [20] propose the UDC-VIT dataset using the video-capturing system. They use a beam-splitter to acquire the paired videos. They cut the UDC area of the Samsung Galaxy Z-Fold 5 [24] and attach it to one side of a beam-splitter. The UDC-VIT dataset accurately depicts the actual UDC degradations, including spatially and temporally variant flares.

4. Restoration Models for UDC Restoration

This section provides an overview of deep-learning models designed to address UDC degradations. Section 4.1 focuses on image restoration models; Section 4.2 discusses face recognition models and Section 4.3 covers video restoration models.

4.1. UDC Image Restoration Models

UDC image restoration models can be categorized into two main paradigms. The first leverages prior knowledge of the diffraction blur kernel (i.e., PSF) as input to guide the correction of diffraction artifacts. The second approach harnesses the representational capabilities of Deep Neural Networks (DNNs) to directly learn UDC degradation patterns.

The first approaches include DISCNet [2], UDC-UNet [25], SV-Net [30], and a method proposed by Kwon et al. [26]. DISCNet incorporates the domain knowledge of UDC image formation (i.e., PSF). UDC-UNet integrates prior information through kernel branches and adapts spatial variations using condition branches. SV-Net uses the Brown-Conrady distortion model [32] with coordinate embedding to facilitate the CNN's learning of space-variant distortions by mapping local image patches to their corresponding spatial coordinates. Kwon et al. propose a controllable image restoration algorithm that addresses spatially variant blur and noise in UDC images using pixel-wise UDC-specific kernel representation and a noise estimator.

The second approaches include BNUDC [27], FSI [28], SFIM [29], and U-Net using the guided filter [31]. BNUDC decouples UDC image degradation into two components, high-frequency diffraction, and low-frequency spatial attenuation, enabling them to be learned separately. FSI combines frequency learning to remove spectral bias, spatial learning to enhance local details, and a dual transfer unit for cross-domain interaction with a color transform module for color enhancement. SFIM employs a multi-level architecture combining spatial and frequency domain blocks through an attention-based integration module. U-Net-based model uses the guided filter to reduce computational load, making the network five times lighter while maintaining effective reconstruction through simple local linear operations.

4.2. UDC Face Recognition Models

Wang et al. [14] introduce LRDif, a method designed to improve performance in facial expression recognition (FER) tasks specific to UDC images. It aims to predict a range of emotions, including surprise, fear, disgust, happiness, sadness, anger, and neutrality. No methods have been proposed for pair-matching tasks using UDC-degraded datasets.

4.3. UDC Video Restoration Models

Research on UDC video restoration remains largely unexplored. DDRNet [23] employs a recurrent network that integrates bi-directional propagation and multi-scale feature processing. Yoo et al. [5] propose a U-Net-based network incorporating a guided filter, reducing computation by nearly 90% by downsampling FHD input frames to SD resolution.

5. Discussion and Conclusion

This paper reviews existing UDC datasets and restoration methods, highlighting the progress and challenges in this field. Capturing real-world UDC images requires significant time and effort, highlighting the need for research to create efficient datasets that effectively exhibit actual degradation. For example, generative models (e.g., diffusion models) can be trained to learn UDC degradation. Wave optic simulations with more precise modeling of display and camera systems can be another choice [5]. While restoration methods for UDC images have proven effective in many cases, they still struggle to recover regions severely degraded by intense flares. Furthermore, research on UDC video restoration is limited, with challenges like frame-to-frame inconsistencies leading to flickering artifacts. Improving temporal coherence in video restoration models is crucial for future work. Furthermore, there is a need for research on pair-matching tasks for face recognition using UDC-degraded datasets.

6. References

1. 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.
2. 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.

3. Poly Haven [Online]. Available: https://polyhaven.com/.
4. ZTE Corporation. ZTE Axon 20 [Online]. Available: https://global.ztedevices.com/products/zte-axon-20-5g/.
5. Yoo, J. et al. "7-5: Synthetic Dataset for Improving UDC Video Restoration Network Performance". *In: SID Symposium Digest of Technical Papers. Wiley Online Library.* 2022, pp. 58–60.
6. 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.
7. Apple Inc. iPhone 13 Pro [Online]. Available: https://www.apple.com/by/iphone-13-pro/specs/.
8. 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.
9. Samsung Electronics Co., Ltd. Samsung Galaxy Z Fold 3 [Online]. Available: https://www.samsung.com/us/smartphones/galaxy-z-fold3-5g/.
10. Samsung Electronics Co., Ltd. Samsung Galaxy Note 10 [Online]. Available: https://www.samsung.com/my/smartphones/galaxy-note10/specs/.
11. 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.
12. 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.
13. 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.
14. Wang, Z., Zhang, K. and Sankaranarayana, R., "LRDif: Diffusion Models for Under-Display Camera Emotion Recognition," *arXiv*, 2024.
15. 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.
16. 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.
17. 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.
18. 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.
19. Ahn, K., Park, C. and Lee, J., "Enhancing Face Recognition Accuracy for Under-Display Cameras via Image Restoration," *In: SID Symposium Digest of Technical Papers. Wiley Online Library.* 2025.
20. Ahn, K., Kim, J. S., Lee, S., Lee, H. G., Ko, B., and Lee, J., "UDC-VIT: A Real-World UDC Video Dataset for Under-Display Camera," *arXiv preprint arXiv:2501.18545*. 2025.
21. Chen, Xuanxi, et al. "Deep Video Restoration for Under-Display Camera." arXiv preprint arXiv:2309.04752 (2023).
22. Pexels [Online]. Available: https://www.pexels.com.
23. Liu C, Wang X, Fan Y, Li S, Qian X. "Decoupling degradations with recurrent network for video restoration in under-display camera," *In Proceedings of the AAAI Conference on Artificial Intelligence*. 2024 Mar 24 (Vol. 38, No. 4, pp. 3558-3566).
24. Samsung Electronics Co., Ltd. Samsung Galaxy Z Fold 5 [Online]. Available: https://www.samsung.com/ae/smartphones/galaxy-z-fold5/specs/.
25. Liu X, Hu J, Chen X, Dong C. "Udc-unet: Under-display camera image restoration via u-shape dynamic network," *In European Conference on Computer Vision*. 2022 Oct 23 (pp. 113-129).
26. Kwon, K. et al. "Controllable image restoration for under-display camera in smartphones," *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 2021 (pp. 2073-2082).
27. Koh J, Lee J, and Yoon S. "Bnudc: A two-branched deep neural network for restoring images from under-display cameras," *In Proceedings of the ieee/cvf conference on computer vision and pattern recognition.* 2022 (pp. 1950-1959).
28. Liu C, Wang X, Li S, Wang Y, Qian X. "Fsi: Frequency and spatial interactive learning for image restoration in under-display cameras," *In Proceedings of the IEEE/CVF International Conference on Computer Vision.* 2023 (pp. 12537-12546).
29. Ahn, K., Kim, J., Park, C., Kim, J. S., and Lee, J. "Integrating Spatial and Frequency Information for Under-Display Camera Image Restoration," *arXiv preprint arXiv:2501.18517*. 2025.
30. Kim, D. et al., "54-1: Development of UDC Image Restoration Technology Using Space Variant CNN," I*n SID Symposium Digest of Technical Papers*. 2024. (Vol. 55, No. 1, pp. 735-737).
31. Brown D., "Decentering Distortion of Lenses," Photogramm. *Engn Remote Sensing*. 1966;444462.
32. Kim, D. et al., "18-4: Fast Image Restoration and Glare Removal for Under-Display Camera by Guided Filter," I*n SID Symposium Digest of Technical Papers*. 2021. (Vol. 52, No. 1, pp. 222-223).