Towards Universal Computational Aberration Correction in Photographic Cameras: A Comprehensive Benchmark Analysis

(Supplementary Material)

Xiaolong Qian[1,*] Qi Jiang[1,*] Yao Gao[1] Lei Sun[2,†] Zhonghua Yi[1]
Kailun Yang[3] Luc Van Gool[2] Kaiwei Wang[1,†]
[1]Zhejiang University [2]INSAIT, Sofia University "St. Kliment Ohridski" [3]Hunan University

In this supplementary material, §1 describes a comparison between our benchmark and existing benchmarks. §2 provides additional implementation details on the datasets and benchmark, and also presents further simulation results compared with *Zemax*, along with the generation process of the pseudo-GT. §3 describes the ODE calculation process and highlights its importance for CAC. §4 provides detailed CAC results. §5 presents ablation studies to validate our key findings. §6 demonstrates the advantages of UniCAC considering aspherical lens designs. Finally, § 7 discusses the limitations of our work and outlines directions for future research.

1. Comparison with Existing Benchmarks

To clearly highlight the scale, diversity, and rigorous physical definitions of our proposed dataset, we provide a detailed comparison with existing cross-lens CAC benchmarks in Table 1. Compared to prior works that rely on small-scale, handcrafted, or unspecified lens samples, UniCAC features a large-scale stratified test set with strictly defined optical design ranges and introduces the ODE for objective aberration quantification.

2. More Implementation Details

2.1. Dataset Details

The key to constructing a dataset is to ensure the diversity of lens aberrations. To generate a credible benchmark, we sampled based on the proposed ODE, ensuring a reasonable distribution of CAC difficulty. In terms of training and validation data, we used 873 and 75 lens images, respectively, to simulate aberration images on the DIV2K and Flickr2K [11] datasets. For the test set, we select additional 120 lenses, ensuring that the imaging quality, uniformity, and chromatic aberration distribution are reasonable. The GT images of the test set are taken with a Sony α6600 camera and a Sony $18-135mm$ lens, with a total of 26 images. We apply the test lens to the 26 images one by one for aberration simulation. Specifically, regarding the image formation model, we simulate the broadband degradation of the optics by densely ray-tracing across the continuous visible spectrum (*e.g.*, 400nm to 700nm). These wavelength-dependent PSFs are subsequently integrated with a standard camera response function to convert the continuous spectral degradations into highly accurate RGB kernels, ensuring maximum physical fidelity to real-world sensors. Furthermore, our simulation explicitly integrates the full ISP pipeline, including Bayer patterning and demosaicking. To bridge the sim-to-real gap, we apply synchronized perturbations (primarily white balance gains) to both LQ and GT images as ISP domain randomization. This strategy forces the model to learn aberration features that are invariant to specific color renderings. According to the research needs, these 120 lenses are reasonably selected for analysis to ensure the comprehensiveness and accuracy of the research results.

2.2. Lens Generation Details

OptiFusion [4] is an automatic design algorithm that generates physically-constrained optical lenses based on specified design requirements. It integrates Genetic Algorithm (GA), Simulated Annealing Algorithm (SAA), and ADAM optimization. We first construct a spherical lens dataset based on OptiFusion. The diversity of design specifications ensures that a variety of aberrations are present in our benchmark. The value ranges for design specifications are: $1\sim6$ for piece number with an interval of 1, $20°\sim40°$ for half FoV with an interval of $10°$, and $2.0\sim4.0$ for F number with an interval of 1.0. Fig. 1(a) presents several representative lenses with different specifications included in our benchmark. Notably, the aperture position is determined

*Equal contribution.
†Corresponding authors.

Work	Key Specs (FoV, F/#, Elem.)	Test Set (# Lenses)	Lens Type	Aberration Quantification
Li *et al.* [8]	Unspecified	Small / Unknown	Unspecified	N/A (Random)
Gong *et al.* [5]	Unspecified	∼5	Unspecified	N/A (Random)
OmniLens [7]	Discrete (6 Specific Sets)	6 (Handcrafted)	Spherical + Aspherical	Spot RMS Radius
UniCAC (Ours)	**40∼80°, F/2∼4, 1∼6 Elem.**	**120 (Unseen)**	**Spherical + Aspherical**	**ODE (Proposed)**

Table 1. Comparison with existing CAC benchmarks. UniCAC features strictly defined optical design ranges and a large-scale stratified test set, whereas prior works rely on unspecified or discrete lens samples.

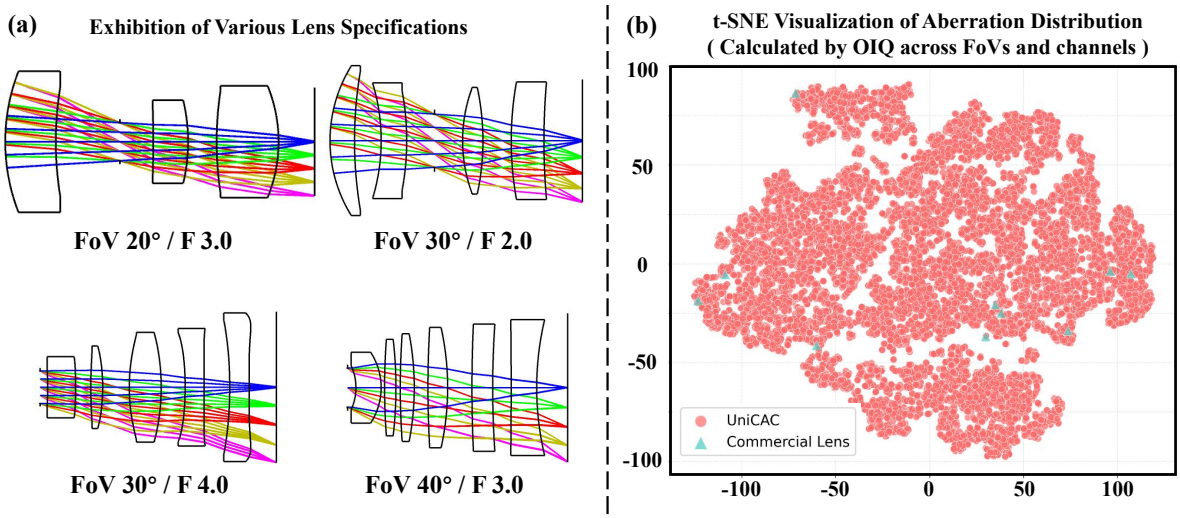

Figure 1. (a) Examples of automatically designed lenses with diverse specifications. (b) t-SNE visualization of aberration distributions in our benchmark. Best viewed when zoomed in.

based on the piece number, which can be located before the first piece, after the last piece, or in the middle of the lens. With all sets of specifications fed into OptiFusion, the initial spherical lens dataset is constructed.

Moreover, we extend OptiFusion's spherical surface parameters to encompass aspherical surfaces. Given a Cartesian coordinate system (x, y, z), the z-axis coincides with the optical axis, while (x, y) forms the transverse plane. Let $r = \sqrt{x^2 + y^2}$ and $\rho = r^2$, then the height h of the traditional spherical surface is defined as:

$$h(\rho) = \frac{c\rho}{1 + \sqrt{1 - c^2\rho}}, \tag{1}$$

where c is the curvature. Further, the height of the standard aspherical surface is defined as:

$$h(\rho) = \frac{c\rho}{1 + \sqrt{1 - \alpha\rho}} + \sum_{i=2}^{n} a_{2i}\rho^i, \tag{2}$$

where $\alpha = (1 + k)c^2$ with k being the conic coefficient, and a_{2i}'s are higher-order coefficients. By setting the conic coefficient and aspherical coefficients as variables, we continue to use OptiFusion to further optimize all the lenses in the initial spherical lens dataset, so as to establish an additional aspherical lens dataset. Combining the initial spherical lens dataset and the aspherical lens dataset, we can achieve a large lens library as our lens data foundation. It is worth noting that the automatically designed lenses conform to physical constraints.

2.3. Lens Aberration Distribution Analysis

To validate the effectiveness and reliability of our benchmark, we perform a t-SNE analysis on the aberration distribution of the benchmark lenses. Following the approach in [5], we select several well-known commercial lenses from ZeBase as references. As shown in Fig. 1(b), the aberration distribution in our benchmark is broad and diverse, effectively covering the range of common aberrations found in real-world optical systems.

2.4. Aberration Simulation Evaluation Details

More comparison with *Zemax*: In Fig. 2, we present additional comparison results for multi-element optical systems to further validate the fidelity of our simulation system. The minor discrepancies mainly stem from three aspects: ray tracing accuracy, the precision of entrance pupil position-

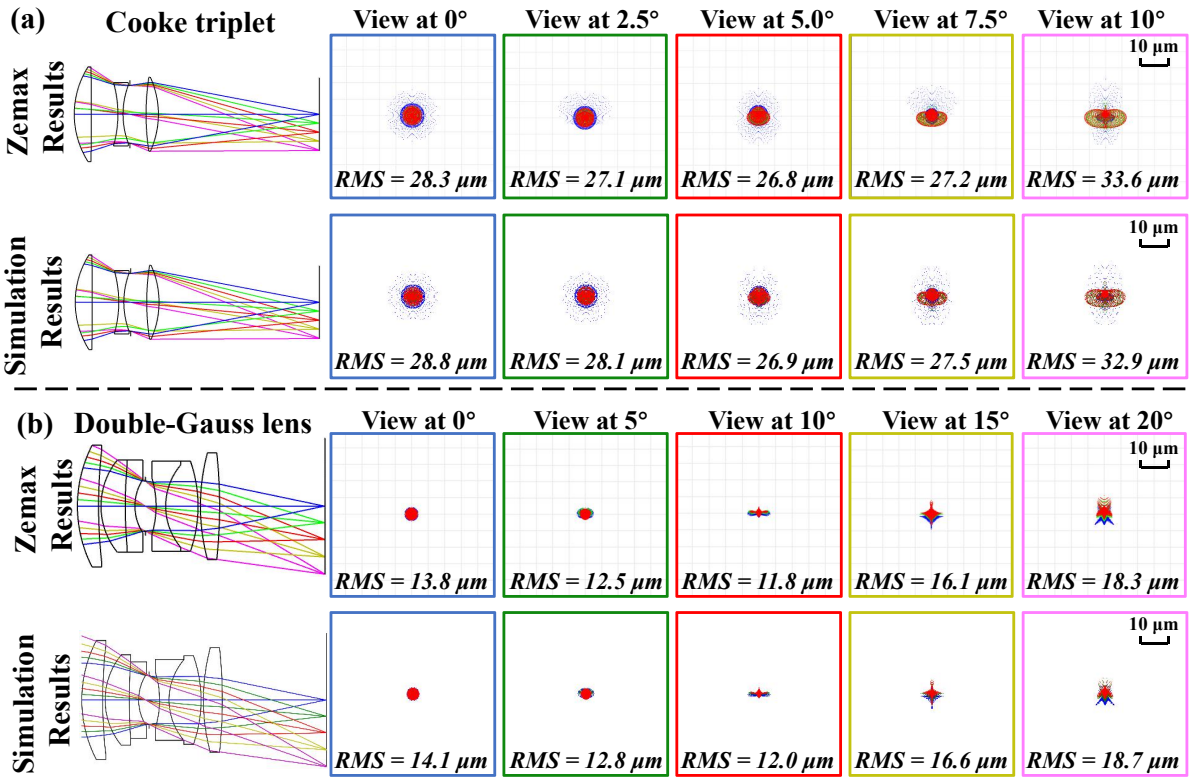

Figure 2. More comparison with *Zemax*. For additional validation, we test the simulation system using Cooke triplet lens and Double-Gauss lens.

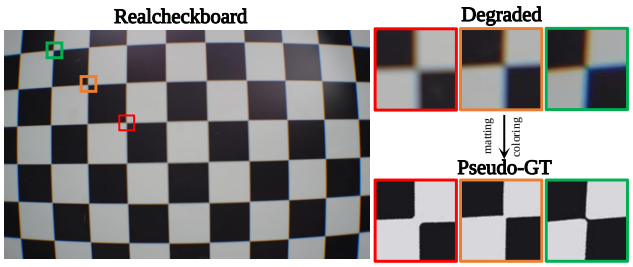

Figure 3. Comparison between the real aberration image and the extracted pseudo-GT image.

ing, and the method of ray sampling at the entrance pupil.

Pseudo-GT: Obtaining pixel-aligned Ground Truth (GT) in the validation of aberration simulation and real-world photography is challenging. According to [2], we generate pseudo-GT by performing matting and coloring on aberration images, shown in Fig. 3. However, the edges in the pseudo-GT differ from those in the real GT due to severe aberrations causing errors in the edge detection operators. Given this situation, we should focus more on the visual effects of aberrations caused by the PSF as demonstrated in Fig. 4 in the main text.

3. More Discussion on the Optical Degradation Evaluator

3.1. Calculation Details

As introduced in the main paper, we compute the Optical Degradation Evaluator (ODE) by first calculating sub-OIQ values across five fields of view (FoVs) and three color channels for a degraded checkerboard. The overall OIQ is then obtained by averaging these values.

For U_s, we merge channel data per FoV and compute the coefficient of variation (CV) across FoVs. For U_c, we combine FoV data for each channel and calculate the CV across channels. Additionally, we present the specific computation details of OIQE [6], which is a key component of ODE in Fig. 4. After calculating the MTF using the edge method, OIQE is obtained by averaging MTF50 and MTF Area. The CV is calculated as follows:

$$CV = \frac{Std(\{OIQ_i | i = 1, ..., N\})}{Avg(\{OIQ_i | i = 1, ..., N\})}. \quad (3)$$

3.2. Importance of ODE

Traditional metrics such as MTF or Strehl ratio focus solely on optical performance, whereas **ODE** combines both optical quality (through MTF-based OIQE) and image fidelity

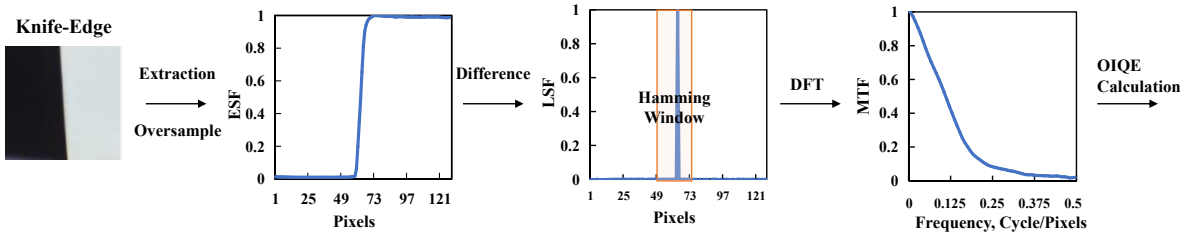

Figure 4. Pipeline for OIQE calculation.

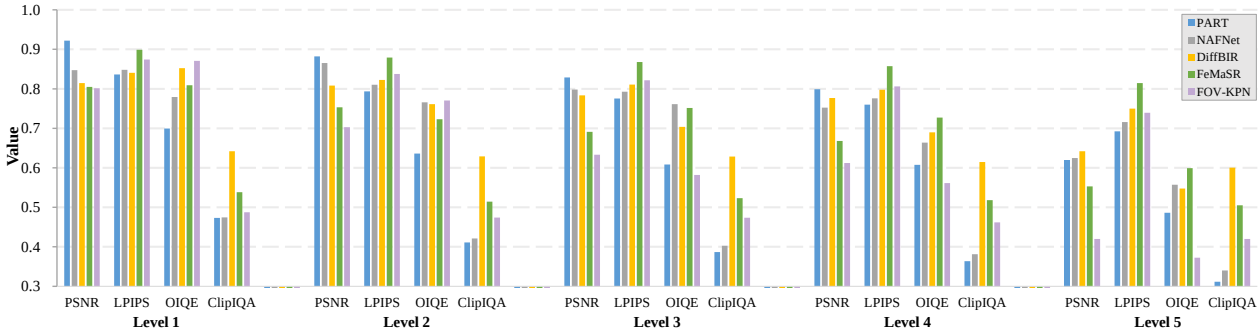

Figure 5. **Overall performance comparison of different methods across five aberration severity levels.** To enhance visual clarity, we normalize PSNR to a range of $20dB$ to $30dB$ and transform LPIPS values by subtracting them from 1.

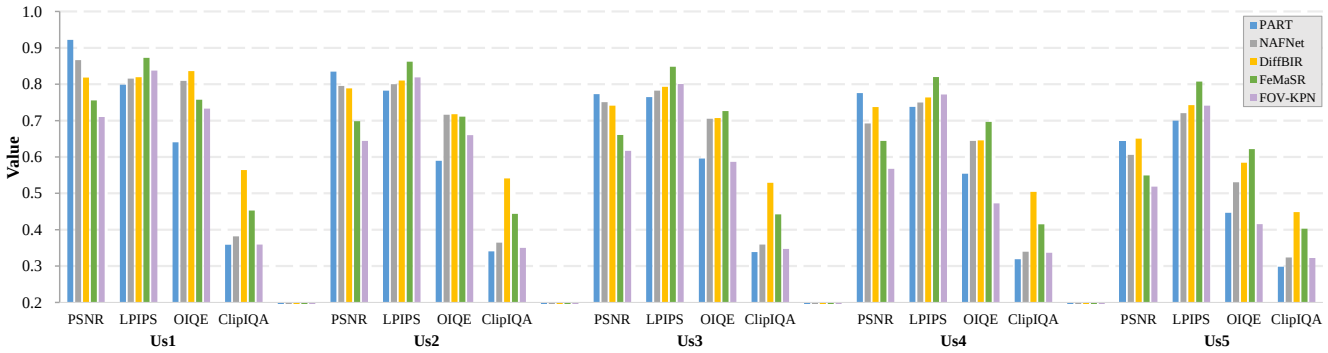

Figure 6. Overall performance comparison of different methods under 5 image spatial uniformity levels.

(via PSNR and SSIM), providing a more comprehensive quantification of CAC difficulty. **ODE** demonstrates *a strong linear relationship with restoration quality*, offering valuable guidance for computational imaging researchers in selecting appropriate lenses.

Beyond replacing the **RMS radius**, **ODE** also *bridges the gap between optical design and CAC*, offering unique insights to guide optical engineers. In practical scenarios, optical design is often constrained by factors like cost, volume, weight, or manufacturing feasibility, making it difficult to design optical systems with optimal performance. **ODE**, *a framework for quantifying CAC difficulty*, empowers designers to make targeted trade-offs.

4. More CAC Results

Quantitative results under 5 different aberration severity levels. We systematically demonstrate the quantitative metrics of 10 methods across various levels of aberration severity in Tab. 5. All methods show a decline in overall performance as the severity of aberration increases, indicating a strong correlation between CAC results and ODE. The results indicate that the PART [6], which leverages PSF information, achieves the highest PSNR from level L1 to L4. However, in the most severe aberration level (L5), the diffusion-based DiffBIR [10] outperforms others, achieving the best results. When focusing on perceptual metrics, methods like FeMaSR [1] and DiffBIR [10], which utilize GAN or diffusion training paradigms, exhibit significant ad-

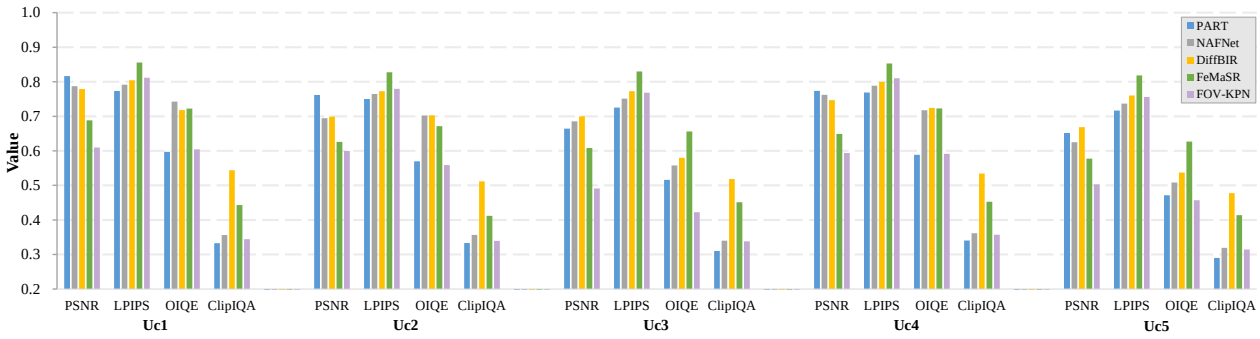

Figure 7. Overall performance comparison of different methods under 5 image chromatic aberration levels.

vantages over other approaches. Additionally, we select five representative networks to provide a more detailed comparison of the specific metric differences across different severity levels, as illustrated in Fig 5.

Quantitative results under 5 different spatial uniformity levels. The detailed metrics are presented in Tab. 6 under different levels of spatial uniformity. Spatial uniformity also affects CAC results, as network performance declines with increasing spatial non-uniformity. To better illustrate metric variation across spatial uniformity levels, Fig. 6 presents results from five representative methods.

Quantitative results under 5 different chromatic aberration levels. Given that chromatic aberration has a minimal impact on CAC results, we have showcased the performance of five representative methods across different chromatic aberration levels. As illustrated in Fig. 7, the model's CAC performance exhibits a low correlation with levels of chromatic aberration, indicating that it is not a primary factor influencing CAC results.

5. Validation of Our Observations.

To verify whether our key observations can provide actionable guidance for CAC model design, we conduct a series of controlled experiments by making targeted modifications inspired by the insights. This not only helps validate the observations themselves but also demonstrates their practical value in improving CAC performance. In Tab. 2, we evaluate the effectiveness of the training strategy and the incorporation of clear image priors. The GAN-based training method improves perceptual quality, while the clear image prior enhances both image fidelity and perceptual realism.

As illustrated in Tab. 3, incorporating FoV and PSF information effectively improves the model's CAC performance in terms of both image fidelity and perceptual quality.

6. On the Effectiveness of UniCAC

To demonstrate the effectiveness of our dataset in considering both spherical and aspherical lenses, we compare it with another dataset based on automatic design that only considers spherical lenses [7]. All experiments are conducted under the same experimental environment and settings. As proven in Tab. 4, our UniCAC outperforms AODLib because it incorporates aspherical lenses into the overall pipeline.

Method	GAN	Codebook	PSNR↑	LPIPS↓	ClipIQA↑
FeMaSR	✗	✗	26.83	0.147	0.491
	✓	✗	26.86	0.145	0.512
	✓	✓	**26.94**	**0.136**	**0.520**

Table 2. Ablation results of training scheme and clean image prior

Method	FoV	PSNR↑	LPIPS↓	ClipIQA↑
DFUnet	✗	25.88	0.253	0.377
	✓	**26.33**	**0.252**	**0.393**
Method	PSF	PSNR↑	LPIPS↓	ClipIQA↑
PART	✗	27.22	0.264	0.370
	✓	**28.10**	**0.228**	**0.389**

Table 3. Ablation results of optical prior

7. Limitations and Future Work

While our comprehensive benchmark yields valuable insights, several areas remain open for further improvement. First, although the extended AOD algorithm [4] has automatically designed a sufficient number of spherical and aspherical lenses, our benchmark primarily focuses on photographic cameras. We have not yet considered advanced optical components such as wavefront-based metasurfaces and Diffractive Optical Elements (DOEs), nor does it include other imaging devices like microscopes, telescopes, and endoscopes, which also suffer from aberrations but with an emphasis on different types. Future research can further incorporate these complex optical systems into the lens library to enhance the overall dataset's aberration diversity. Furthermore, the aberration images used in this study are

CAC Dataset	Model	PSNR↑	SSIM↑	LPIPS↓	FID↓	OIQE↑	ClipIQA↑
AODLib [7]	FOV-KPN[ICCV21] [2]	23.90	0.756	0.244	74.51	0.477	0.430
	SwinIR[ICCVW21] [9]	25.05	0.793	**0.226**	65.18	0.418	**0.418**
	FeMaSR[MM22] [1]	24.15	0.748	0.235	67.73	0.568	0.476
UniCAC (Ours)	FOV-KPN[ICCV21] [2]	**26.34**	**0.824**	**0.184**	**50.27**	**0.631**	**0.463**
	SwinIR[ICCVW21] [9]	**27.22**	**0.848**	0.264	**54.41**	**0.513**	0.370
	FeMaSR[MM22] [1]	**26.94**	**0.841**	**0.136**	**34.59**	**0.722**	**0.520**

Table 4. Quantitative evaluation of the UniCAC.

based on optical simulation [3]. To isolate intrinsic lens aberrations, we currently fix the object distance at infinity, reserving depth-dependent PSFs for future work. Although the simulation accuracy has been validated, future work can incorporate depth-coupled degradations and include real-world images captured using various types of lenses to enhance the comprehensiveness and effectiveness of the dataset.

In the future, we will extend existing AOD algorithms to automatically design a wider variety of optical systems, enhancing the diversity of our dataset. Additionally, it will be essential to develop a more general and powerful CAC architecture based on this enriched dataset.

References

[1] Chaofeng Chen, Xinyu Shi, Yipeng Qin, Xiaoming Li, Xiaoguang Han, Tao Yang, and Shihui Guo. Real-world blind super-resolution via feature matching with implicit high-resolution priors. In *MM*, 2022. 4, 6

[2] Shiqi Chen, Huajun Feng, Keming Gao, Zhihai Xu, and Yueting Chen. Extreme-quality computational imaging via degradation framework. In *ICCV*, 2021. 3, 6

[3] Shiqi Chen, Huajun Feng, Dexin Pan, Zhihai Xu, Qi Li, and Yueting Chen. Optical aberrations correction in postprocessing using imaging simulation. *ACM Transactions on Graphics*, 2021. 6

[4] Yao Gao, Qi Jiang, Shaohua Gao, Lei Sun, Kailun Yang, and Kaiwei Wang. Exploring quasi-global solutions to compound lens based computational imaging systems. *IEEE Transactions on Computational Imaging*, 2025. 1, 5

[5] Jin Gong, Runzhao Yang, Weihang Zhang, Jinli Suo, and Qionghai Dai. A physics-informed low-rank deep neural network for blind and universal lens aberration correction. In *CVPR*, 2024. 2

[6] Qi Jiang, Shaohua Gao, Yao Gao, Kailun Yang, Zhonghua Yi, Hao Shi, Lei Sun, and Kaiwei Wang. Minimalist and high-quality panoramic imaging with PSF-aware transformers. *IEEE Transactions on Image Processing*, 2024. 3, 4

[7] Qi Jiang, Yao Gao, Shaohua Gao, Zhonghua Yi, Xiaolong Qian, Hao Shi, Kailun Yang, Lei Sun, and Kaiwei Wang. OmniLens: Towards universal lens aberration correction via lenslib-to-specific domain adaptation. *arXiv preprint arXiv:2409.05809*, 2024. 2, 5, 6

[8] Xiu Li, Jinli Suo, Weihang Zhang, Xin Yuan, and Qionghai Dai. Universal and flexible optical aberration correction using deep-prior based deconvolution. In *ICCV*, 2021. 2

[9] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. SwinIR: Image restoration using swin transformer. In *ICCVW*, 2021. 6

[10] Xinqi Lin, Jingwen He, Ziyan Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Yu Qiao, Wanli Ouyang, and Chao Dong. DiffBIR: Toward blind image restoration with generative diffusion prior. In *ECCV*, 2024. 4

[11] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, Lei Zhang, Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, et al. NTIRE 2017 challenge on single image super-resolution: Methods and results. In *CVPRW*, 2017. 1

Metric	Aberration Severity	Model									
		PART	DFUnet	FOV-KPN	MIMOUnet	NAFNet	Restomer	Uformer	FeMaSR	PromptIR	DiffBIR
PSNR↑	L1	29.22	27.57	28.02	28.24	28.47	27.77	28.87	28.05	27.77	28.14
	L2	28.82	27.07	27.03	28.22	28.65	27.77	28.66	27.53	27.63	28.08
	L3	28.29	26.45	26.33	27.52	27.98	27.33	28.10	26.91	27.35	27.84
	L4	27.99	25.81	26.12	27.39	27.53	27.04	27.98	26.68	26.94	27.77
	L5	26.19	24.75	24.20	25.43	26.25	25.29	26.12	25.53	25.12	26.42
SSIM↑	L1	0.907	0.895	0.881	0.912	0.915	0.906	0.911	0.885	0.903	0.830
	L2	0.881	0.860	0.842	0.885	0.889	0.882	0.886	0.855	0.878	0.821
	L3	0.868	0.843	0.823	0.871	0.877	0.871	0.875	0.841	0.866	0.815
	L4	0.858	0.831	0.810	0.860	0.867	0.858	0.865	0.830	0.853	0.809
	L5	0.818	0.792	0.765	0.819	0.830	0.817	0.826	0.796	0.811	0.784
LPIPS↓	L1	0.164	0.171	0.126	0.159	0.152	0.176	0.162	0.101	0.180	0.160
	L2	0.206	0.226	0.162	0.204	0.190	0.209	0.199	0.121	0.216	0.177
	L3	0.224	0.252	0.178	0.226	0.207	0.226	0.216	0.132	0.234	0.189
	L4	0.240	0.269	0.194	0.243	0.224	0.246	0.232	0.143	0.255	0.202
	L5	0.307	0.341	0.261	0.314	0.284	0.309	0.293	0.186	0.322	0.250
FID↓	L1	26.245	31.457	29.238	27.651	23.464	48.498	36.715	23.092	50.689	32.741
	L2	33.112	40.898	39.963	34.338	30.221	52.635	43.093	28.540	53.165	34.505
	L3	38.016	46.582	43.220	38.394	34.174	54.949	48.039	31.982	56.094	38.269
	L4	48.594	56.814	56.357	48.930	43.376	66.733	57.585	37.809	67.738	42.851
	L5	72.351	85.647	82.588	75.949	64.716	90.950	78.637	51.548	96.022	56.908
OIQE↑	L1	0.700	0.766	0.871	0.859	0.779	0.806	0.826	0.809	0.825	0.853
	L2	0.636	0.642	0.770	0.820	0.766	0.774	0.755	0.723	0.782	0.761
	L3	0.608	0.620	0.582	0.776	0.761	0.731	0.746	0.751	0.755	0.704
	L4	0.607	0.596	0.561	0.715	0.664	0.684	0.696	0.727	0.702	0.690
	L5	0.486	0.433	0.372	0.540	0.557	0.527	0.546	0.599	0.532	0.548
ClipIQA↑	L1	0.473	0.469	0.487	0.477	0.475	0.392	0.438	0.538	0.387	0.642
	L2	0.411	0.412	0.474	0.406	0.421	0.341	0.388	0.496	0.331	0.629
	L3	0.387	0.391	0.474	0.381	0.403	0.321	0.369	0.489	0.312	0.629
	L4	0.364	0.371	0.462	0.352	0.381	0.297	0.346	0.480	0.283	0.615
	L5	0.311	0.325	0.420	0.302	0.340	0.255	0.301	0.416	0.244	0.600

Table 5. Quantitative results under 5 different aberration severity levels.

Metric	Spatial Uniformity	Model									
		PART	DFUnet	FOV-KPN	MIMOUnet	NAFNet	Restomer	Uformer	FeMaSR	PromptIR	DiffBIR
PSNR↑	Us1	29.22	26.87	27.10	28.30	28.66	27.98	28.89	27.56	27.78	28.18
	Us2	27.89	26.10	26.44	27.56	27.95	27.22	28.17	26.98	27.07	27.89
	Us3	27.73	25.97	26.17	27.07	27.50	26.76	27.74	26.61	26.69	27.41
	Us4	27.37	25.61	25.67	26.71	26.92	26.62	27.42	26.44	26.64	27.37
	Us5	26.43	24.75	25.19	25.65	26.06	25.42	26.11	25.49	25.35	26.51
SSIM↑	Us1	0.885	0.865	0.846	0.890	0.894	0.887	0.891	0.859	0.882	0.830
	Us2	0.826	0.849	0.834	0.877	0.883	0.874	0.879	0.851	0.870	0.826
	Us3	0.865	0.843	0.824	0.868	0.873	0.865	0.871	0.843	0.861	0.818
	Us4	0.811	0.827	0.811	0.851	0.858	0.850	0.855	0.831	0.846	0.811
	Us5	0.829	0.806	0.790	0.827	0.836	0.828	0.831	0.814	0.822	0.798
LPIPS↓	Us1	0.202	0.220	0.163	0.199	0.184	0.206	0.195	0.128	0.215	0.181
	Us2	0.190	0.240	0.181	0.219	0.200	0.225	0.212	0.138	0.232	0.190
	Us3	0.235	0.262	0.199	0.238	0.218	0.243	0.228	0.152	0.250	0.207
	Us4	0.236	0.285	0.228	0.270	0.251	0.273	0.259	0.180	0.280	0.236
	Us5	0.300	0.320	0.259	0.311	0.279	0.303	0.293	0.193	0.314	0.258
FID↓	Us1	36.281	44.166	41.581	35.766	32.920	54.721	45.589	30.527	55.965	36.512
	Us2	38.638	48.958	47.971	42.362	37.217	60.577	51.895	33.350	62.155	38.638
	Us3	48.667	58.617	57.240	51.937	43.631	66.133	57.538	38.260	69.144	43.284
	Us4	54.127	69.734	70.895	66.125	60.343	82.859	73.699	50.195	83.771	54.127
	Us5	65.855	77.619	69.387	73.627	61.036	84.739	76.640	48.737	87.640	56.686
OIQE↑	Us1	0.640	0.665	0.733	0.823	0.809	0.775	0.743	0.757	0.745	0.836
	Us2	0.718	0.555	0.660	0.766	0.716	0.696	0.674	0.711	0.722	0.718
	Us3	0.596	0.599	0.587	0.701	0.705	0.654	0.671	0.726	0.695	0.707
	Us4	0.646	0.528	0.472	0.644	0.644	0.583	0.623	0.696	0.625	0.646
	Us5	0.447	0.429	0.415	0.540	0.530	0.540	0.484	0.621	0.560	0.584
ClipIQA↑	Us1	0.359	0.367	0.359	0.364	0.382	0.325	0.350	0.452	0.315	0.564
	Us2	0.541	0.347	0.350	0.340	0.364	0.304	0.330	0.444	0.295	0.541
	Us3	0.338	0.343	0.347	0.335	0.359	0.302	0.325	0.442	0.294	0.529
	Us4	0.504	0.332	0.337	0.315	0.339	0.290	0.311	0.415	0.281	0.504
	Us5	0.298	0.321	0.322	0.302	0.324	0.275	0.295	0.403	0.268	0.448

Table 6. Quantitative results under 5 different spatial uniformity levels.