Beyond Pipeline Mimicry: Expert-Level Aesthetic ISP via Reward-Guided Flow

Tong Qiao[1], **Kepeng Xu**[1], **Gang He**[1,*], **Zhenyang Liu**[1] and **Wenxin Yu**[2]
[1]Xidian University
[2]Southwest University of Science and Technology

Abstract

Recent deep learning-based ISP methods are primarily constrained by mimicking fixed camera pipelines, consequently struggling to achieve expert-level aesthetic quality. To this end, we propose AesISP, the first expert-level aesthetic ISP framework formulated as a reward flow model. Addressing the ill-posed nature of ISP, AesISP establishes a Privileged Prior Distillation paradigm. By leveraging expert-image-guided latent proxies, we decompose the intractable task into a tractable, proxy-driven learning process. Meanwhile, we propose MeanFlow++, which reformulates the flow matching objective via target-prediction parameterization and a Progressive Temporal Curriculum. By evolving from learning instantaneous to long-range average velocities, this mechanism rectifies transport trajectories and uses deterministic mapping to impose explicit constraints, anchoring the generation trajectory to the expert manifold. Finally, we introduce a Multi-dimensional Reward-driven Reinforcement Learning approach.Leveraging Group Relative Policy Optimization (GRPO) to balance trade-offs across dimensions such as color and lighting, it steers the model to converge precisely on expert-level aesthetic standards. Experiments demonstrate that AesISP outperforms state-of-the-art methods in both quantitative metrics and aesthetic quality.

1 Introduction

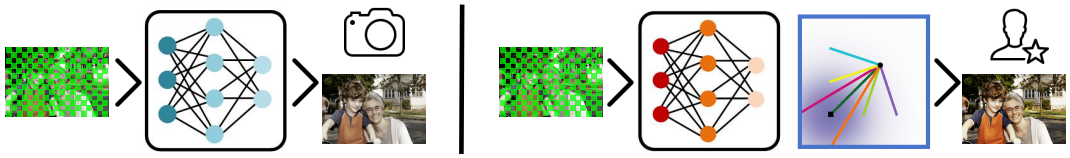
In the field of image signal processing (ISP), a significant paradigm shift is occurring, transitioning from traditional hardware-driven logic to data-driven deep learning (DL) models. Most mainstream DL-ISP approaches adopt a supervised learning framework, utilizing paired data to fit deterministic mappings prescribed by standard engines or commercial cameras (e.g., RawPy or Canon). These methods accurately model the transformation from sensor data to sRGB space. Recent end-to-end real-world camera pipelines further show that jointly learning RAW-to-sRGB processing is effective under practical imaging conditions [Xu *et al.*, 2024]. However, they are inherently high-fidelity proxies for a fixed imaging pipeline, which limits their flexibility and artistic adaptability.

This deterministic nature, while precise, fundamentally conflicts with the art of photography. High-quality imaging requires adaptive, stylistic adjustments based on specific visual semantics and mood, rather than rigid approximations. By enforcing a one-size-fits-all logic across different scenes, existing methods degrade into static converters, prioritizing rule compliance over aesthetic expression. To overcome this limitation, we propose a shift toward Aesthetic ISP (AesISP), aiming to reconstruct the processing paradigm. We focus on moving away from fidelity-driven simulation to aesthetics-driven generation, aligning more closely with human expert preferences to produce images that balance signal integrity with exceptional visual appeal.

Building such an aesthetic ISP system presents unique challenges, requiring a reevaluation of existing learning paradigms. Directly regressing Raw data to expert-level aesthetic images through traditional models is highly challenging due to the ill-posed nature of the mapping. To address this, we introduce the concept of Privileged Prior Distillation. Rather than directly tackling this high-dimensional generation task, we decompose it into two more tractable stages. In the first stage, we propose an expert-validated condensing model (Aesthetic Condensing Network (ACN)) and a Raw-to-RGB transformation model (Raw-to-SRGB Mapping Model (R2SM)). The ACN condenses the complex transformation from Raw to aesthetic sRGB into a compact, low-dimensional aesthetic ISP latent variable, z_{aes}. Subsequently, the R2SM uses the aesthetic prior z_{aes} to guide the Raw-to-RGB transformation. Since z_{aes} is a low-dimensional abstraction of the complex color grading logic, it significantly improves the tractability of modeling compared to the full pixel space. This strategy not only circumvents the inherent ambiguity of direct mapping but also forces the model to learn the direction of aesthetic adjustments, avoiding mechanical memorization of pixel values.

In the second stage, we develop an Expert Prior Generator (EPG) to model the distribution of expert-level aesthetic latents, z_{aes}, from Raw inputs alone. The generated prior then guides R2SM toward high-quality sRGB rendering.

*Corresponding authors.



a). Previous methods: Learning fixed traditional ISP. b). Our method: Learning expert-level aesthetic ISP.

Figure 1: **From pipeline emulation to aesthetic generation.** (a) Existing deep ISPs [Zhang *et al.*, 2021] typically use supervised regression to mimic fixed hardware pipelines (e.g., Canon), often failing to capture semantic-adaptive aesthetics. (b) AesISP formulates the task as reward-guided flow generation, leveraging an expert prior and multi-dimensional rewards to optimize for both fidelity and expert-level aesthetics.

We observe that vanilla MeanFlow faces a key bottleneck in aesthetic ISP: its linear flow assumption is ill-suited to the high-curvature mapping from the linear Raw domain to the nonlinear aesthetic manifold. We therefore introduce two core improvements:

- **Progressive Temporal Curriculum**: We use a progressive temporal curriculum to bridge the geometric gap between local tangents and global secants on curved manifolds. Training begins with local denoising on the expert prior manifold and gradually expands the sampled time span, enabling stable learning of long-range transformations.
- **Target-Predictive Parameterization**: We introduce target-predictive parameterization to suppress high-frequency noise in velocity regression. Since the expert aesthetic prior, z_{aes}, lies on a compact low-dimensional manifold, target prediction is more stable than direct regression in high-dimensional Raw space. Projecting the predicted velocity back to signal space imposes explicit manifold constraints, keeping predictions anchored to the expert aesthetic prior.

Together, these advancements stabilize Raw-to-RGB conversion while preserving subtle aesthetic details.

Given the subjective and multi-dimensional nature of aesthetic quality, pixel-level losses and scalar rewards cannot capture nuanced perceptual trade-offs in arrangement, color harmony, and overall appeal. RAW-centric multimodal studies further suggest that sensor-level information can support high-level vision-language reasoning [Xu *et al.*, 2025], motivating a closer link between low-level imaging and semantic aesthetic assessment. We therefore introduce a multi-dimensional reward-driven reinforcement learning approach that uses Visual Language Models (VLMs) to optimize aesthetic criteria such as arrangement, color balance, and emotional tone, guiding the model toward expert-level standards.

As the first expert-level aesthetic ISP framework built on a reward-flow model, we present *AesISP*. The main contributions of our work are as follows.

- **AesISP Framework**: The first expert-level aesthetic ISP based on a reward flow model, surpassing traditional fixed-pipeline imitation paradigms to achieve rendering that combines both high fidelity and artistic expression.
- **Privileged Prior Distillation**: A novel approach that decomposes the ill-posed Raw-to-expert mapping into a reference-guided surrogate learning process, effectively constraining the solution space.
- **MeanFlow++ Mechanism**: A mechanism that combines target prediction with progressive curriculum learning to reconstruct flow objectives, anchoring the expert manifold robustly through explicit constraints.
- **Multi-Dimensional Aesthetic Reward Learning**: Utilizing Visual Language Models (VLM) to optimize multiple aesthetic criteria, such as arrangement and color harmony, aligning the model with human aesthetic preferences.

2 Related Work

Neural Image Signal Processing Deep learning has revolutionized the design philosophy of ISP pipelines. Recent enhancement and fusion studies show that perceptual reconstruction often requires illumination-aware modeling and adaptive objectives beyond pixel-wise fidelity [Bai *et al.*, 2025a; Bai *et al.*, 2025c]. This trend is also reflected in broader computational imaging tasks, where degradation estimation, modality interaction, and signal-specific dynamics are exploited for light-field restoration, event-based deblurring, and optical-flow modeling [Xiao and Xiong, 2025; Xiao *et al.*, 2026; Zhao *et al.*, 2021; Zhao *et al.*, 2026]. In image and video coding, recent benchmarks and learned prediction/filtering modules further indicate that high-quality reconstruction depends on task-specific priors and signal-adaptive processing [Li *et al.*, 2025; Li *et al.*, 2024a; Li *et al.*, 2024b]. Classical regression models such as CSANet [Hsyu *et al.*, 2021], AWNet [Dai *et al.*, 2020], and LiteISP [Zhang *et al.*, 2021] essentially serve as neural approximators for traditional hardware ISPs. Even generative methods like ISPDiffuser [Ren *et al.*, 2025] and D&R [Mo *et al.*, 2025], despite improving texture details, remain constrained by the "pipeline fitting" paradigm, where the ultimate goal is to replicate the outputs of existing camera algorithms. In contrast, AesISP shifts ISP from "fidelity-oriented simulation" to "aesthetic-oriented generation," directly mapping RAW inputs to expert-level aesthetic representations beyond standard pipeline constraints.

Generative Models While Diffusion Probabilistic Models (DPMs) [Ho *et al.*, 2020] and Continuous Normalizing Flows (CNFs) [Chen *et al.*, 2018] excel in generation tasks, their iterative integration processes often incur prohibitive inference latency. To address this, Flow Matching (FM) [Lipman *et al.*, 2022] introduces a simulation-free paradigm, achieving deterministic transport by regressing fixed velocity fields.

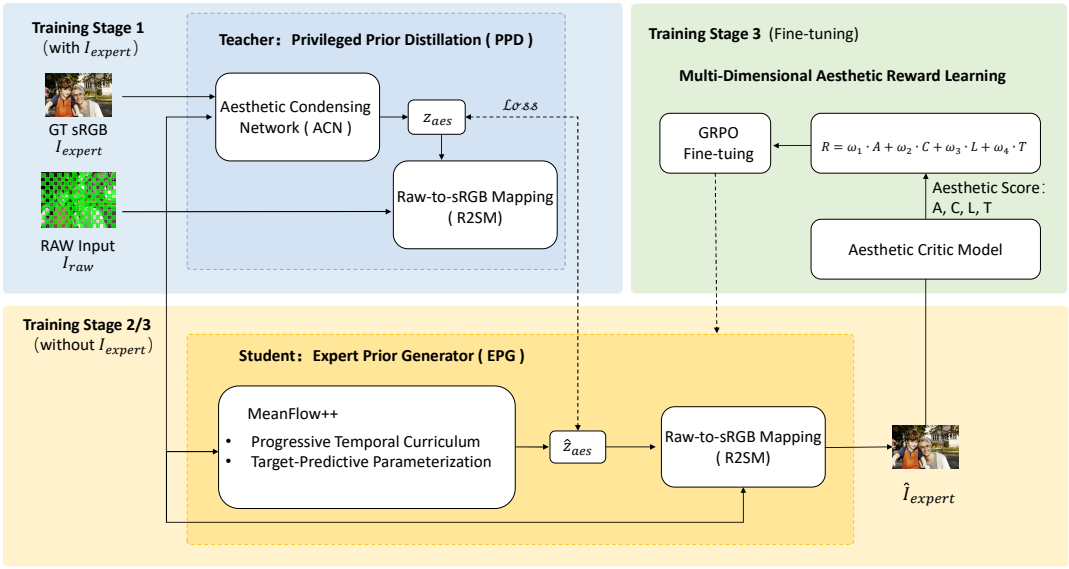


Figure 2: **The framework of AesISP. (a) Privileged Prior Distillation (PPD):** An Aesthetic Condensing Network (ACN) extracts expert intent from I_{expert} into a compact latent z_{aes} to condition the Raw-to-sRGB mapping. **(b) Generative Prior Modeling:** The Expert Prior Generator (EPG) predicts $\hat{z}_{aes}$ directly from I_{raw} using MeanFlow++, which incorporates target-predictive parameterization for robust generation. **(c) Multi-Dimensional Aesthetic Reward Learning:** The model is fine-tuned via GRPO using multi-dimensional aesthetic scores derived from VLMs to align with expert-level standards.

Building on this, MeanFlow [Geng *et al.*, 2025] further realizes ultra-fast one-step generation by modeling average velocity. Although these velocity-centric frameworks have succeeded in general synthesis tasks, applying them to aesthetic ISP reveals a fundamental theoretical bottleneck. As pointed out by Li et al. [Li and He, 2025], velocity is inherently an off-manifold quantity belonging to the high-dimensional ambient space; directly regressing it introduces significant task-irrelevant high-frequency noise interference. More critically, the high-curvature nature of the Raw-to-Aesthetic mapping induces severe geometric misalignment: standard training objectives cannot reconcile the significant discrepancy between global average velocity (secant) and local instantaneous velocity (tangent), leading to ambiguous optimization directions. To surmount these limitations, we propose MeanFlow++. This framework redirects the optimization objective toward manifold-anchored signal prediction, circumventing high-entropy noise in the velocity field. Furthermore, the introduction of a progressive time curriculum eliminates geometric misalignment caused by high curvature, ensuring the model robustly captures the complex mapping from Raw to aesthetic space.

3 Methodology

3.1 Overview

AesISP is designed to reconstruct the intricate mapping between raw sensor data ($\mathbf{I}_{raw}$) and expert-level aesthetic images ($\mathbf{I}_{expert}$). As illustrated in Figure 2, to tackle the inherent ill-posed nature of this translation, our framework operates through three tightly coupled stages, effectively decomposing the high-dimensional generation task into tractable proxy learning processes.

First, we establish a Privileged Prior Distillation paradigm. During training, utilizing $\mathbf{I}_{expert}$ as privileged information, we employ an Aesthetic Condensing Network (ACN) to distill complex color grading logic into a compact, low-dimensional aesthetic ISP latent variable (z_{aes}). This latent variable subsequently serves as a prior to guide the Raw-to-SRGB Mapping Model (R2SM), thereby circumventing the ambiguity associated with direct mapping.

Second, to accurately model the distribution of z_{aes} solely from $\mathbf{I}_{raw}$ during inference, we introduce the Expert Prior Generator (EPG), underpinned by our novel MeanFlow++ mechanism. Addressing the complex mapping from the linear Raw domain to the non-linear aesthetic manifold, MeanFlow++ integrates a Progressive Temporal Curriculum with Target-Predictive Parameterization. By evolving from learning instantaneous local denoising to modeling long-range deterministic transport paths, it constructs a flow trajectory with explicit structural constraints, robustly anchoring the prediction to the expert manifold.

Finally, to ensure alignment with human aesthetic preferences, we incorporate a Multi-Dimensional Aesthetic Reward Learning mechanism. This approach leverages Visual Language Models (VLMs) to capture high-level semantics such as arrangement and color harmony, guiding the model fine-tuning to produce sRGB images that balance high signal fidelity with exceptional artistic expressiveness.

3.2 Privileged Prior Distillation Paradigm

Mapping raw sensor data ($\mathbf{I}_{raw}$) to expert-level aesthetic images ($\mathbf{I}_{expert}$) constitutes an inherently ill-posed inverse problem. Directly learning this complex, one-to-many mapping is

arduous and prone to suboptimal convergence. To address this, we propose the Privileged Prior Distillation (PPD) strategy, designed to regularize the solution space by leveraging privileged information available during training.

Specifically, we introduce the Aesthetic Condensing Network (ACN), denoted as $\mathcal{E}_{\text{ACN}}$. Unlike traditional encoders that focus solely on content features, $\mathcal{E}_{\text{ACN}}$ takes both $\mathbf{I}_{raw}$ and $\mathbf{I}_{expert}$ as joint inputs to distill the underlying aesthetic adjustment logic. The network compresses this complex nonlinear transformation into a compact, low-dimensional aesthetic ISP latent variable z_{aes}:

$$z_{\text{aes}} = \mathcal{E}_{\text{ACN}}(\mathbf{I}_{raw}, \mathbf{I}_{expert}) \tag{1}$$

where z_{aes} serves as a low-dimensional proxy for high-dimensional aesthetic operations, effectively decoupling the aesthetic grading instructions from the image content.

Subsequently, we employ the Raw-to-SRGB Mapping Model (R2SM), denoted as $\mathcal{M}_{\text{R2SM}}$, which acts as a conditional execution engine. $\mathcal{M}_{\text{R2SM}}$ utilizes the aesthetic intent encoded in z_{aes} to guide the physical rendering process from $\mathbf{I}_{raw}$ to the sRGB space. We jointly optimize ACN and R2SM by minimizing the pixel-wise reconstruction error between the generated output and the expert ground truth:

$$\mathcal{L}_{\text{distill}} = \|\mathcal{M}_{\text{R2SM}}(\mathbf{I}_{raw}, z_{\text{aes}}) - \mathbf{I}_{expert}\|_1 \tag{2}$$

Through this distillation paradigm, we decompose the ambiguous direct regression task into two tractable subproblems: aesthetic intent extraction via ACN and controlled image reconstruction via R2SM. This strategy not only significantly reduces modeling complexity but also forces z_{aes} to form a compact and semantically continuous manifold, laying a robust foundation for the subsequent flow-based prediction stage.

3.3 MeanFlow++ Mechanism

Following the Privileged Prior Distillation stage, we obtain a training set of paired data $\{(\mathbf{I}_{raw}, z_{\text{aes}})\}$, where $z_{\text{aes}} = \mathcal{E}_{\text{ACN}}(\mathbf{I}_{raw}, \mathbf{I}_{expert})$ represents the low-dimensional aesthetic latent prior distilled by the frozen ACN. The objective of this second stage is to model the conditional distribution $p(z_{\text{aes}} \mid \mathbf{I}_{raw})$: given solely $\mathbf{I}_{raw}$, the model must generate a latent code z that lies precisely on the expert aesthetic manifold. To achieve efficient and high-fidelity generation, we construct the MeanFlow++ framework upon the average-velocity modeling principle of MeanFlow [Geng *et al.*, 2025]. We introduce two complementary advancements—a progressive temporal curriculum and target-predictive parameterization—to address the challenges of fitting this complex mapping from the linear Raw domain to the non-linear aesthetic manifold.

Progressive Temporal Curriculum

To enable efficient one-step generation (1-NFE), MeanFlow++ must ultimately master long-range transport, mapping noise $\boldsymbol{\epsilon}$ directly to the aesthetic prior z_{aes} across the full time interval (i.e., $r = 0, t = 1$). While MeanFlow theoretically supports learning average velocities over arbitrary intervals [Geng *et al.*, 2025], exposing the model to long-range supervision immediately creates a significant optimization bottleneck. At the early stages, the model has not yet established a stable correspondence between the photographic semantics in $\mathbf{I}_{raw}$ and the geometry of the z_{aes} manifold.

To mitigate this, we introduce a Progressive Temporal Curriculum. This strategy resolves geometric misalignment by organizing learning from local tangent fitting to global trajectory alignment. The curriculum proceeds dynamically by controlling the time gap $\Delta = t - r$:

1. Instantaneous Stage ($\Delta \approx 0$): The objective reduces to standard flow matching [Lipman *et al.*, 2022; Geng *et al.*, 2025], where the model learns instantaneous velocities to capture local manifold tangents.
2. Short-Range Stage (Small Δ): The model expands its receptive field to learn average velocities over short intervals, consolidating stability near the manifold.
3. Long-Range Stage (Large Δ): Finally, the model learns global transport across large gaps, acquiring the capability for direct Raw-to-Aesthetic mapping.

In practice, we sample $\Delta \sim \mathcal{U}(0, \Delta_{\max})$ and gradually anneal $\Delta_{\max}$ from 0 to 1 during training.

Target-Predictive Parameterization

Standard flow-based approaches typically regress a conditional velocity field $\mathbf{u}$. However, the velocity field resides in a high-entropy, off-manifold space, making direct regression prone to variance and error accumulation. In our context, the target z_{aes} lies on a structured, semantically meaningful *expert aesthetic manifold*. Regressing velocity often leads to "drift", where the integral of the predicted flow deviates from this precise manifold.

To address this, we reformulate the learning paradigm as Target-Prediction. Instead of optimizing the velocity field directly, we leverage the MeanFlow identity [Geng *et al.*, 2025] to project the predicted velocity back into the signal space, enforcing an explicit constraint on the terminal state. Specifically, given noise $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and an interpolated state $\mathbf{z}_t = (1-t)z_{\text{aes}} + t\boldsymbol{\epsilon}$, we derive a closed-form estimate of the aesthetic prior, denoted as $\hat{\mathbf{z}}_0$:

$$\hat{\mathbf{z}}_0(\mathbf{z}_t, r, t) \triangleq \boldsymbol{\epsilon} - \mathbf{u}_\theta(\mathbf{z}_t, r, t \mid \mathbf{I}_{raw}) - (t-r)\frac{\partial \mathbf{u}_\theta}{\partial t} \tag{3}$$

where $\mathbf{u}_\theta$ is the network prediction and $\frac{\partial \mathbf{u}_\theta}{\partial t}$ is computed via automatic differentiation. This formulation essentially inverts the flow equation to predict the origin z_{aes}. Consequently, we supervise the model directly in the target space:

$$\mathcal{L}_{\text{MF++}} = \mathbb{E}\left[\mathcal{L}_{\text{rec}}(\hat{\mathbf{z}}_0, z_{\text{aes}})\right] \tag{4}$$

where $\mathcal{L}_{\text{rec}}$ denotes a robust reconstruction loss function. This objective mathematically enforces Manifold Anchoring: by minimizing the discrepancy between the predicted origin $\hat{\mathbf{z}}_0$ and the ground truth z_{aes}, we constrain the flow trajectory to terminate precisely on the expert manifold, rather than merely matching intermediate velocities. This significantly stabilizes the generation process and ensures that the synthesized latents faithfully preserve the intricate color grading and tonal logic distilled from the experts.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	NIMA ↑	MUSIQ ↑	MANIQA ↑	Q-Align ↑	MUSIQ-AVA ↑	CNNIQA ↑
AWNet[Dai *et al.*, 2020]	24.330	0.8814	0.1086	4.7591	63.83	0.3621	3.8242	4.7922	0.5712
CSANet[Hsyu *et al.*, 2021]	23.074	0.8478	0.1329	4.7539	64.58	0.3685	3.8069	4.7574	0.5944
LiteISPNet[Zhang *et al.*, 2021]	24.693	0.8810	0.0895	4.8248	64.90	0.3869	3.9179	4.8361	0.5846
MW-ISPNet[Ignatov *et al.*, 2020]	23.487	0.8824	0.1014	4.8141	64.37	0.3656	3.8355	4.8056	0.5791
MobileIE[Yan *et al.*, 2025]	23.932	0.8448	0.1506	4.7194	60.74	0.3423	3.6233	4.7404	0.5468
MetaISP[Souza and Heidrich, 2024]	24.461	0.8754	0.1336	4.7591	63.15	0.3621	3.6854	4.6929	0.5239
FourierISP[He *et al.*, 2024]	24.758	0.8830	0.1126	4.7933	64.60	0.3873	3.8400	4.8236	0.5769
ISPDiffuser[Ren *et al.*, 2025]	23.007	0.8453	0.1620	4.8067	61.26	0.3112	3.7535	4.5742	0.4687
AesISP (Ours)	**26.173**	**0.9096**	**0.0675**	**4.8542**	**65.97**	**0.4039**	**3.9956**	**4.8995**	0.5942

Table 1: **Comprehensive quantitative comparison**. We compare our AesISP against state-of-the-art baselines on the aesthetic ISP task. Our method achieves strong performance across fidelity and aesthetic metrics. Specifically, the substantial gains in aesthetic metrics (e.g., Q-Align, NIMA) validate our model's capability to master expert-level aesthetic semantics beyond simple pipeline mimicry.

3.4 Multi-Dimensional Aesthetic Reward Learning

Given the inherently subjective and multi-dimensional nature of aesthetic quality, pixel-level reconstruction losses alone are inadequate for capturing nuanced perceptual differences. Inspired by task-adaptive low-level objectives [Bai *et al.*, 2025c; Bai *et al.*, 2025b], we model aesthetics with multi-dimensional rewards. Traditional scalar rewards fall short in modeling the intricate trade-offs between arrangement, color harmony, and overall visual appeal. To address this, we introduce a novel multi-dimensional reward-driven reinforcement learning approach, termed Multi-Dimensional Aesthetic Reward Learning.

Specifically, we employ Gemini-3 [Google DeepMind, 2025] as the aesthetic critic to evaluate generated images across four dimensions: Arrangement (A), Color Harmony (C), Lighting Tone (L), and Atmosphere (T). These scores are aggregated into a unified reward R_{fused}:

$$R_{fused} = w_1 \cdot A + w_2 \cdot C + w_3 \cdot L + w_4 \cdot T, \quad \text{s.t.} \sum_{i=1}^{4} w_i = 1 \tag{5}$$

For efficient fine-tuning, we utilize Group Relative Policy Optimization (GRPO) [Shao *et al.*, 2024]. For each raw input, we evaluate G candidates and compute the group-relative advantage $\hat{A}_i$:

$$\hat{A}_i = \frac{R_i - \mu_{\text{group}}}{\sigma_{\text{group}} + \epsilon} \tag{6}$$

Here, R_i is the fused reward, and μ_{group} and σ_{group} denote group statistics. This signed advantage promotes above-average candidates while balancing fidelity and aesthetics.

3.5 Training Strategy

We employ a Three-stage Progressive Protocol to train AesFlow on the FiveK dataset. This protocol decouples optimization objectives, ensuring robust convergence across modules.

Stage I: Expert Manifold Construction. This stage establishes a reliable expert latent manifold. We jointly train the Aesthetic Condensing Network (ACN) and the R2SM decoder. Using paired supervision, ACN extracts the latent z_{aes}, from which R2SM reconstructs the expert image. We optimize the reconstruction objective $\mathcal{L}_{rec} = \|\mathcal{M}_{\text{R2SM}}(\mathbf{I}_{raw}, z_{\text{aes}}) - \mathbf{I}_{expert}\|_1$. This constraint forces z_{aes} to encapsulate essential color grading cues, serving as the ground truth anchor for subsequent prior fitting.

Stage II: Generative Prior Fitting via EPG. Fixing the decoder, we train the Expert Prior Generator (EPG) to fit $p(z_{\text{aes}} \mid \mathbf{I}_{raw})$. EPG adopts a Coarse-to-Fine strategy: it first predicts a coarse deterministic estimate, and MeanFlow++ learns to generate the residual $\mathbf{z}_{res}$. To stabilize optimization, we adopt the Adaptive Weighted Objective from MeanFlow [Geng *et al.*, 2025]:

$$\mathcal{L}_{\text{MF++}} = \mathbb{E}\left[\text{sg}(w) \cdot \|\delta\|_2^2\right], \quad w = (\|\delta\|_2^2 + c)^{\gamma - 1} \tag{7}$$

where δ denotes the prediction error, w is the adaptive weight, and $\text{sg}(\cdot)$ is the stop-gradient operator. This objective dynamically balances sample contributions via γ, ensuring a stable transition from local denoising to long-range transport under our progressive curriculum.

Stage III: Aesthetic Preference Fine-tuning. The final stage performs end-to-end fine-tuning using the GRPO mechanism. Leveraging the stochastic nature of MeanFlow++, we synthesize a diverse pool of candidate priors, which are rendered by R2SM. The model is then updated via VLM-driven multi-dimensional rewards, aligning the output with human aesthetic preferences while maintaining pixel-level fidelity.

4 Experiments

4.1 Experimental Settings

Dataset and Metrics We conduct experiments on the MIT-Adobe FiveK dataset [Bychkovsky *et al.*, 2011]. We allocate 4381 image pairs for training and use 200 pairs for testing. To comprehensively assess performance, we employ a holistic set of metrics across three dimensions: (1) Fidelity: PSNR and SSIM measure signal reconstruction accuracy. (2) Perceptual Quality: LPIPS [Zhang *et al.*, 2018] evaluates human-perceived image degradation. (3) Aesthetic & Semantic Quality: We utilize a suite of reference-free metrics, including NIMA [Talebi and Milanfar, 2018], MUSIQ and MUSIQ-AVA [Ke *et al.*, 2021], MANIQA [Yang *et al.*, 2022], CNNIQA [Kang *et al.*, 2014], and the VLM-based metric Q-Align [Wu *et al.*, 2023], to strictly evaluate aesthetic alignment and visual statistics.

Implementation Details Our framework is implemented in PyTorch on an NVIDIA RTX 5090 GPU. We use the AdamW optimizer with an initial learning rate of 2×10^{-4}. Training follows our three-stage progressive protocol. A batch size of 24 is used for the first two stages. In the final stage (GRPO Fine-tuning), we adjust the batch size to 6 with a group size of

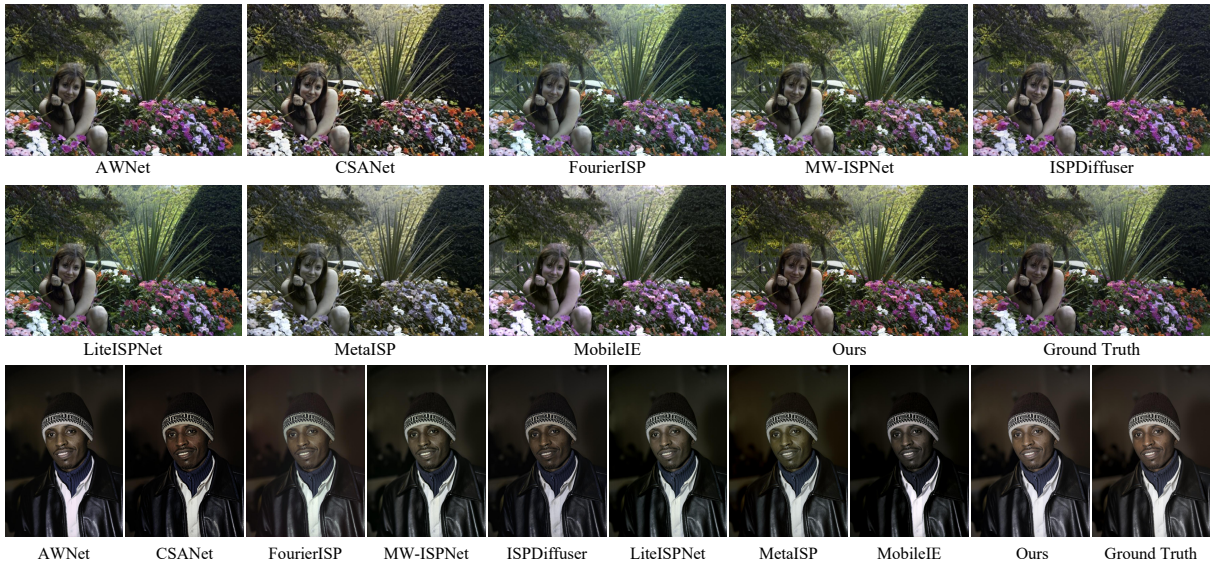


Figure 3: **Visual results.** Existing regression methods (e.g., AWNet) yield desaturated and flat results due to statistical averaging. In contrast, AesISP faithfully restores chromatic vibrancy and tonal hierarchy, surpassing regression smoothness to achieve expert-aligned aesthetic quality closest to the Ground Truth.

$G = 4$ to manage computational costs while enabling effective relative optimization. For inference, thanks to the MeanFlow++ mechanism, we achieve high-fidelity generation with a single function evaluation (NFE=1).

Method	PSNR ↑	SSIM ↑	LPIPS ↓
Baseline (Direct Mapping)	24.69	0.8809	0.0895
Ours (w/ Distillation)	**25.12**	**0.9022**	**0.0739**

Table 2: **Impact of Prior Distillation.** The significant boost in PSNR and SSIM quantitatively confirms the efficacy of PPD: by incorporating privileged priors, the model effectively resolves the ill-posed nature of direct mapping

4.2 Comparison with Existing Methods

We benchmark AesISP against representative regression-based baselines (e.g., LiteISP [Zhang *et al.*, 2021], FourierISP [He *et al.*, 2024]) and the generative ISPDiffuser [Ren *et al.*, 2025].

Quantitative Evaluation. Table 1 details the performance on the FiveK test set. AesISP achieves strong results across fidelity and aesthetic metrics, validating its ability to overcome paradigm limitations:

- Reconciling Fidelity and Aesthetics: AesISP demonstrates strong fidelity and aesthetic quality. This validates our Privileged Prior Distillation: while regression models often converge to a statistical mean of color gradings, resulting in blurry outputs, AesISP leverages the distilled latent z_{aes} to lock onto the unique expert solution, ensuring precise pixel-level restoration.
- Overcoming Mapping Bottlenecks: Although ISPDiffuser also adopts a diffusion-based generative framework, it lags in LPIPS and Q-Align. This confirms that standard diffusion models struggle to master the high-curvature Raw-to-sRGB mapping. In contrast, AesISP maximizes the benefits of MeanFlow++'s Target-Predictive Parameterization and Progressive Temporal Curriculum. These mechanisms explicitly anchor the trajectory to the expert manifold and facilitate long-range transport, ensuring both high generative quality and controllability.

Qualitative Visual Evaluation. As shown in Figure 3, while regression baselines yield desaturated results and diffusion models suffer from texture drift, AesISP generates results that are closest to the Expert-Retouched Reference in terms of color tension and lighting hierarchy, perfectly preserving structural details.

4.3 Ablation Studies and Analysis

Verification of Privileged Prior Distillation

Attempting to model the high-dimensional Raw-to-Expert mapping solely via a monolithic network causes convergence towards the statistical mean of the training distribution, failing to capture distinct expert styles. This regression-to-the-mean phenomenon is visually substantiated by the error heatmap in Fig. 5. Widespread high-error regions (red/green) are observed in the Direct Mapping Baseline. These large-scale deviations are the direct consequence of the model converging to a statistical average when confronted with complex semantics. Unable to lock onto a specific mode, the baseline degrades into a mediocre approximation, resulting in significant divergence from the Ground Truth across broad areas.

To mitigate this, the PPD paradigm decouples the task via z_{aes}, enforcing explicit stylistic guidance. As shown in Table 2, incorporating PPD elevates PSNR to 25.12 dB. This confirms that leveraging privileged priors effectively transcends the limitations of regression-to-the-mean, enabling

Stage	PSNR ↑	SSIM ↑	LPIPS ↓	NIMA ↑	MUSIQ ↑	MANIQA ↑	Q-Align ↑	MUSIQ-AVA ↑	CNNIQA ↑
w/o GRPO (Stage 2)	26.457	0.9113	0.0665	4.8479	65.83	0.4024	3.9816	4.8959	0.5937
w/ GRPO (Stage 3)	26.173	0.9096	0.0675	**4.8542**	**65.97**	**0.4039**	**3.9956**	**4.8995**	**0.5942**

Table 3: **Impact of GRPO on Aesthetic Metrics.** Comparison of Stage 2 (w/o GRPO) and the five-run average of Stage 3 (w/ GRPO). GRPO improves aesthetic scores, especially MUSIQ and Q-Align.

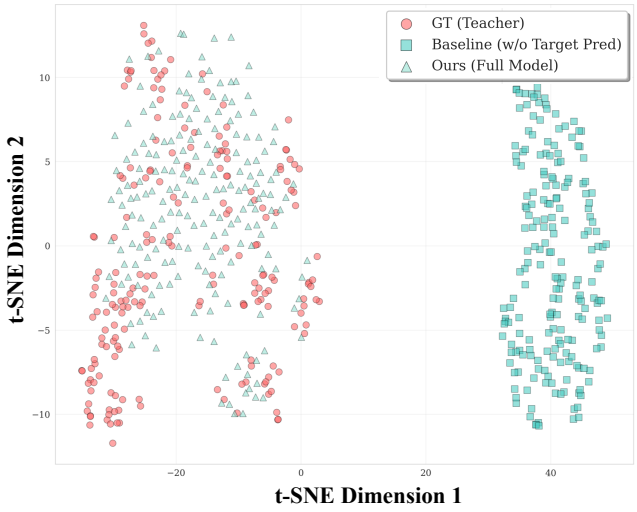


Figure 4: **t-SNE Visualization.** The Baseline exhibits severe distributional shift (isolated cluster), whereas TPP realigns predictions with the Ground Truth, demonstrating precise manifold anchoring.

high-fidelity complex mappings.

Method	PSNR ↑	SSIM ↑	LPIPS ↓
Baseline (Velocity Reg.)	25.12	0.9022	0.0739
+ Target Prediction	25.61	0.9037	0.0708
+ Progressive Curriculum	**26.45**	**0.9112**	**0.0665**

Table 4: **Step-wise Ablation of MeanFlow++.** TPP improves structural fidelity (SSIM) via explicit manifold constraints, while PTC boosts PSNR by correcting geometric misalignment.

Verification of the MeanFlow++ Mechanism

To validate the efficacy of the MeanFlow++ design, we conduct a progressive ablation analysis to investigate how Target-Predictive Parameterization (TPP) and the Progressive Temporal Curriculum (PTC) synergistically resolve manifold anchoring and long-range transport challenges.

Necessity of TPP for Manifold Anchoring. In contrast to indirect velocity field constraints that are prone to high-dimensional noise, TPP imposes explicit constraints by projecting predicted velocities, strictly confining the generation trajectory to the expert manifold. The t-SNE visualization in Fig. 4 confirms this: the unconstrained baseline exhibits a severe distributional shift (isolated cluster), whereas TPP successfully realigns the predicted distribution to overlap with the Ground Truth. This precise anchoring in feature space directly translates into the significant SSIM improvements observed in Table 4.

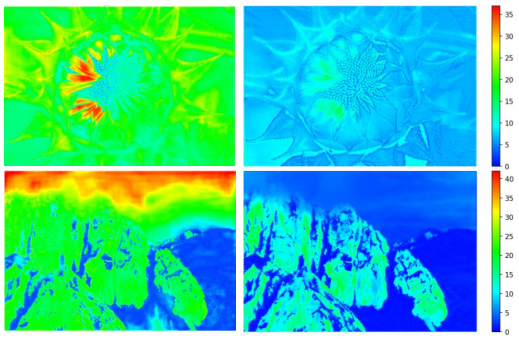


Figure 5: **Visual evidence of regression-to-the-mean.** The widespread high-error regions (red/green) in the baseline visually attest to regression-to-the-mean, a direct consequence of converging to a statistical average rather than a specific style. Conversely, the dominance of low errors (blue) in our method proves precise anchoring to the expert manifold.

Necessity of PTC for Long-Range Transport. Although TPP constrains the endpoint, the non-linear mapping induces geometric misalignment between local tangents and global secants. PTC corrects this by guiding a transition from local denoising to global generation, effectively rectifying the transport path. This strategy triggers a PSNR leap from 25.61 dB to 26.45 dB, confirming its indispensability for resolving long-range transport issues on high-curvature manifolds.

Verification of Multi-Dimensional Aesthetic Reward Learning

Traditional pixel-wise objectives struggle to model intricate trade-offs between arrangement, color harmony, and overall visual appeal. To bridge this gap, we introduce a VLM-driven MDAR designed to capture multi-dimensional aesthetic criteria, including arrangement layout, color balance, and emotional tone. By integrating these multi-faceted dimensions into the learning process, our mechanism effectively guides the model to achieve results that align with expert-level aesthetic standards.

Table 3 shows that GRPO improves perceptual metrics such as Q-Align and MUSIQ, confirming its benefit for expert-level aesthetic alignment despite a slight PSNR drop.

5 Conclusion

We presented AesISP, a reward-guided flow framework for expert-level aesthetic ISP. Combining privileged prior distillation, MeanFlow++, and multi-dimensional reward learning, AesISP improves fidelity and aesthetic quality on MIT-Adobe FiveK. Future work will study sensor generalization and controllable editing.

Acknowledgements

This paper is supported by the construction project of the Fujiang Laboratory Nuclear Medicine Artificial Intelligence Research Center (No. 2023ZYDF074). Tong Qiao and Kepeng Xu contributed equally to this work. Please ask Dr. Gang He for correspondence.

References

[Bai *et al.*, 2025a] Haowen Bai, Jiangshe Zhang, Zixiang Zhao, Lilun Deng, Yukun Cui, and Shuang Xu. Retinexmef: Retinex-based glare effects aware unsupervised multi-exposure image fusion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7251–7261, 2025.

[Bai *et al.*, 2025b] Haowen Bai, Jiangshe Zhang, Zixiang Zhao, Yichen Wu, Lilun Deng, Yukun Cui, Tao Feng, and Shuang Xu. Task-driven image fusion with learnable fusion loss. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 7457–7468, 2025.

[Bai *et al.*, 2025c] Haowen Bai, Zixiang Zhao, Jiangshe Zhang, Yichen Wu, Lilun Deng, Yukun Cui, Baisong Jiang, and Shuang Xu. Refusion: Learning image fusion from reconstruction with learnable loss via meta-learning. *International Journal of Computer Vision*, 133(5):2547–2567, 2025.

[Bychkovsky *et al.*, 2011] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input / output image pairs. In *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[Chen *et al.*, 2018] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[Dai *et al.*, 2020] Linhui Dai, Xiaohong Liu, Chengqi Li, and Jun Chen. Awnet: Attentive wavelet network for image isp. In *European Conference on Computer Vision*, pages 185–201. Springer, 2020.

[Geng *et al.*, 2025] Zhengyang Geng, Mingyang Deng, Xingjian Bai, J Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. *arXiv preprint arXiv:2505.13447*, 2025.

[Google DeepMind, 2025] Google DeepMind. Gemini 3 flash: Frontier intelligence built for speed. https://deepmind.google/models/gemini/flash/, December 2025. Accessed: 2026-01-20.

[He *et al.*, 2024] Xuanhua He, Tao Hu, Guoli Wang, Zejin Wang, Run Wang, Qian Zhang, Keyu Yan, Ziyi Chen, Rui Li, Chengjun Xie, et al. Enhancing raw-to-srgb with decoupled style structure in fourier domain. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 2130–2138, 2024.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[Hsyu *et al.*, 2021] Ming-Chun Hsyu, Chih-Wei Liu, Chao Chen, Chao-Wei Chen, and Wen-Chia Tsai. Csanet: High speed channel spatial attention network for mobile isp. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 2486–2493, 2021.

[Ignatov *et al.*, 2020] Andrey Ignatov, Radu Timofte, Zhilu Zhang, Ming Liu, Haolin Wang, Wangmeng Zuo, Jiawei Zhang, Ruimao Zhang, Zhanglin Peng, Sijie Ren, et al. Aim 2020 challenge on learned image signal processing pipeline. In *European Conference on Computer Vision*, pages 152–170. Springer, 2020.

[Kang *et al.*, 2014] Le Kang, Peng Ye, Yi Li, and David Doermann. Convolutional neural networks for no-reference image quality assessment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1733–1740, 2014.

[Ke *et al.*, 2021] Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5148–5157, 2021.

[Li and He, 2025] Tianhong Li and Kaiming He. Back to basics: Let denoising generative models denoise. *arXiv preprint arXiv:2511.13720*, 2025.

[Li *et al.*, 2024a] Zhuoyuan Li, Jiacheng Li, Yao Li, Li Li, Dong Liu, and Feng Wu. In-loop filtering via trained look-up tables. In *2024 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pages 1–5, 2024.

[Li *et al.*, 2024b] Zhuoyuan Li, Yao Li, Chuanbo Tang, Li Li, Dong Liu, and Feng Wu. Uniformly accelerated motion model for inter prediction. In *2024 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pages 1–5, 2024.

[Li *et al.*, 2025] Zhuoyuan Li, Junqi Liao, Chuanbo Tang, Haotian Zhang, Yuqi Li, Yifan Bian, Xihua Sheng, Xinmin Feng, Yao Li, Changsheng Gao, et al. Ustc-td: A test dataset and benchmark for image and video coding in 2020s. *IEEE Transactions on Multimedia*, 2025.

[Lipman *et al.*, 2022] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

[Mo *et al.*, 2025] Zhipeng Mo, Wenbo Li, and Sihao Ding. Diffuse and refine latent prior with transformers for neural isp. In *2025 IEEE International Conference on Image Processing (ICIP)*, pages 1456–1461. IEEE, 2025.

[Ren *et al.*, 2025] Yang Ren, Hai Jiang, Menglong Yang, Wei Li, and Shuaicheng Liu. Ispdiffuser: Learning raw-to-srgb mappings with texture-aware diffusion models and histogram-guided color consistency. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 6722–6730, 2025.

[Shao *et al.*, 2024] Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Alan Song, Mingchuan Xiao, Yaofei Kelude, Yining Zhang, Kungu Li, Yikang Lv, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

[Souza and Heidrich, 2024] Matheus Souza and Wolfgang Heidrich. Metaisp–exploiting global scene structure for accurate multi-device color rendition. *arXiv preprint arXiv:2401.03220*, 2024.

[Talebi and Milanfar, 2018] Hossein Talebi and Peyman Milanfar. Nima: Neural image assessment. *IEEE transactions on image processing*, 27(8):3998–4011, 2018.

[Wu *et al.*, 2023] Haoning Wu, Zicheng Zhang, Weixia Zhang, Chaofeng Chen, Liang Liao, Chunyi Li, Yixuan Gao, Annan Wang, Erli Zhang, Wenxiu Sun, et al. Q-align: Teaching lmms for visual scoring via discrete text-defined levels. *arXiv preprint arXiv:2312.17090*, 2023.

[Xiao and Xiong, 2025] Zeyu Xiao and Zhiwei Xiong. Incorporating degradation estimation in light field spatial super-resolution. *Computer Vision and Image Understanding*, 252:104295, 2025.

[Xiao *et al.*, 2026] Zeyu Xiao, Zhuoyuan Li, Yang Zhao, Yu Liu, Zhao Zhang, and Wei Jia. Learning dual modality interactions for event-based motion deblurring. *IEEE Transactions on Multimedia*, 2026.

[Xu *et al.*, 2024] Kepeng Xu, Zijia Ma, Li Xu, Gang He, Yunsong Li, Wenxin Yu, Taichu Han, and Cheng Yang. An end-to-end real-world camera imaging pipeline, 2024.

[Xu *et al.*, 2025] Kepeng Xu, Tong Qiao, Zhenyang Liu, Li Xu, and Gang He. End-to-end RAW synergy for elevated vision-language reasoning. In *The First Workshop on Multimodal Knowledge and Language Modeling*, 2025.

[Yan *et al.*, 2025] Hailong Yan, Ao Li, Xiangtao Zhang, Zhe Liu, Zenglin Shi, Ce Zhu, and Le Zhang. Mobileie: An extremely lightweight and effective convnet for real-time image enhancement on mobile devices. *arXiv preprint arXiv:2507.01838*, 2025.

[Yang *et al.*, 2022] Sidi Yang, Tianhe Wu, Shuwei Shi, Shanshan Lao, Yuan Gong, Mingdeng Cao, Jiahao Wang, and Yujiu Yang. Maniqa: Multi-dimension attention network for no-reference image quality assessment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1191–1200, 2022.

[Zhang *et al.*, 2018] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018.

[Zhang *et al.*, 2021] Zhilu Zhang, Haolin Wang, Ming Liu, Ruohao Wang, Jiawei Zhang, and Wangmeng Zuo. Learning raw-to-srgb mappings with inaccurately aligned supervision. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4348–4358, 2021.

[Zhao *et al.*, 2021] Rui Zhao, Ruiqin Xiong, Ziluo Ding, Xiaopeng Fan, Jian Zhang, and Tiejun Huang. Mrdflow: Unsupervised optical flow estimation network with multi-scale recurrent decoder. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(7):4639–4652, 2021.

[Zhao *et al.*, 2026] Rui Zhao, Ruiqin Xiong, Dongkai Wang, Shiyu Xuan, Jian Zhang, Xiaopeng Fan, and Tiejun Huang. Spike camera optical flow estimation based on continuous spike streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(4):4756–4770, 2026.