NEGATIVE-FILM RESTORATION VIA THE HETEROGENEITY OF CHANNEL-INDUCED GEOMETRIC FEATURES

Anonymous authors
Paper under double-blind review

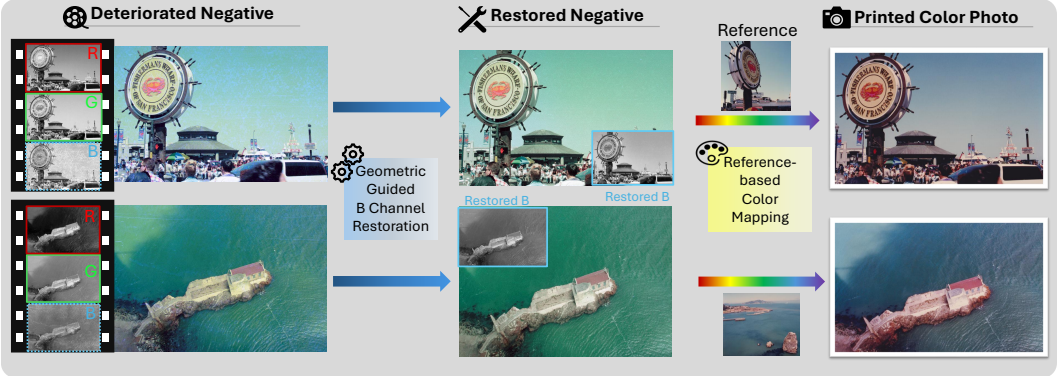

Figure 1: Overview of our geometry-guided negative-film restoration framework. Given a degraded negative, we use cross-channel discrepancies in frozen DA3 (Lin et al., 2025) features to guide blue channel restoration. We then establish DA3-based geometric correspondences between the restored negative and a reference printed photo to transfer region-aware appearance cues and produce a coherent print-like image.

ABSTRACT

Color film negatives are vulnerable to channel-heterogeneous deterioration from long-term chemical aging, with the blue (B) channel often more severely corrupted than other channels. Existing restoration methods typically process degraded RGB images holistically and therefore do not fully exploit the complementary cross-channel information. In this work, we propose a geometry-guided approach based on a key observation: different color channels depict the same underlying scene geometry and should yield consistent features from a pretrained 3D foundation model in reliable regions, whereas deterioration introduces cross-channel discrepancies. These discrepancies guide spatially adaptive restoration of the B channel while preserving the R and G channels. We further use the geometric features to guide reference-conditioned print color mapping by associating geometrically and semantically related regions across the restored negative and a reference printed photo. Extensive experiments demonstrate that our approach achieves superior negative-film restoration and more coherent color mapping than existing methods. Our code is available at the GitHub repository.

1 INTRODUCTION

Film negatives preserve a large portion of the pre-digital visual record, including historical events, cultural heritage, and personal archives in physical form. Unlike digital photographs or printed photos, color negatives store red (R), green (G), and blue (B) information in physically distinct color-sensitive layers of the film. Given this layer-wise storage mechanism, long-term chemical aging can produce *channel-heterogeneous deterioration* (Liu et al., 2025). In particular, the scanned B channel is more susceptible to degradation and often exhibits more severe corruption than the R and G channels, as illustrated in the left of Fig. 1. Such degradation can range from high-frequency

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

mottled artifacts to large-area gradual fading, while the other channels remain comparatively intact and preserve reliable scene structure.

This channel-heterogeneous deterioration distinguishes negative-film restoration from both conventional old-photo restoration methods (Wan et al., 2020) and general-purpose image restoration methods (Potlapalli et al., 2023; Wang et al., 2026a), which jointly restore the RGB image without explicitly exploiting cross-channel degradation asymmetry. When applied to deteriorated negatives, these methods fail to fully exploit the reliable structural information preserved in the relatively intact R and G channels to guide the restoration of the corrupted B channel, limiting their ability to cope with the diverse and spatially varying deterioration within the B channel.

To better exploit such cross-channel structural information, our core insight is that, despite their distinct intensity distributions, the three color channels of a negative depict the same underlying scene geometry and should therefore yield consistent geometric features in reliable regions. Channel-specific deterioration breaks this consistency, i.e. when the B channel is severely corrupted, its geometric features deviate from other relatively channels. This cross-channel discrepancy thus provides a natural cue for localizing structurally unreliable regions. Based on this idea, we treat the R, G, and B channels as channel-wise images and jointly extract their intermediate geometric features using Depth Anything 3 (DA3) (Lin et al., 2025). We then compare the B-channel features with those from the relatively intact R and G channels, using the resulting cross-channel geometric inconsistency as spatial guidance to amplify restoration updates in more inconsistent regions.

The above structural restoration alone cannot recover the color appearance that we see in the printed photograph, because the restored image remains the color characteristics of the negative. Traditionally, the transformation from a film negative to a printed photograph involved a nonlinear photochemical color-mapping process. We further explore whether geometric priors can facilitate color mapping for negatives with a reference printed photo. We observe that regions with similar geometric and semantic content across the restored negative and the reference printed photo can provide meaningful cues for color mapping. For instance, grass regions in the reference printed photo can guide the color mapping of similar regions in the negative. We found that the intermediate geometric features of DA3 can also capture such associations, and transfer color distributions from the reference printed photo to related regions of the restored negative, producing more plausible print-like colors. Therefore, our structural restoration and color-mapping modules share a similar design.

Our contributions are summarized as follows:

- We exploit cross-channel inconsistencies in geometry-aware features of film negatives to enable spatially-varying restoration that amplifies residual updates in more inconsistent regions.
- We further exploit region-level similarities in geometric features across a restored negative and a reference printed photo to identify geometrically and semantically related regions, enabling more coherent reference-conditioned color mapping.
- We effectively leverage the geometric priors encoded in the pretrained 3D foundation model for both negative-film restoration and subsequent color mapping, outperforming existing combinations of restoration and color-mapping methods on this task.

2 Related Work

Legacy Photo Restoration. Legacy photo restoration initially focused on detecting and repairing localized defects such as scratches, cracks, and missing regions through interpolation or inpainting (Bruni & Vitulano, 2004; Chang et al., 2005). As real old photographs often contain multiple interacting degradations, subsequent learning-based methods moved toward unified restoration frameworks. Wan et al. (Wan et al., 2020; 2022) addressed the domain gap between synthetically degraded and real old photographs through latent-space translation, while later approaches modeled mixed degradations using learned degradation priors, unified pipelines, or damage-specific representations (Zhao et al., 2021; Xu et al., 2023; Cai et al., 2023; 2024). While prior methods primarily target generic aging artifacts, we address film-specific deterioration with a characteristic cross-channel dependency. This structure provides an intrinsic cue for restoration that existing old-photo pipelines do not explicitly exploit, while our framework further integrates reference-guided printed-color reconstruction within the same restoration pipeline.

Image Restoration. Image restoration has evolved from task-specific models (Zhang et al., 2017; Cai et al., 2016; Li et al., 2017; Ren et al., 2019; Cho et al., 2021) to general-purpose architectures, including multi-stage networks (Zamir et al., 2021), Transformers (Liang et al., 2021; Zamir et al., 2022), activation-free networks (Chen et al., 2022), state-space models (Guo et al., 2024), and diffusion bridges (Wang et al., 2026b). All-in-one restoration further handles multiple degradations through learned degradation representations or prompts (Li et al., 2022; Potlapalli et al., 2023; Wang et al., 2023), vision–language or instruction priors (Luo et al., 2024; Conde et al., 2024), diffusion-based representations (Zheng et al., 2024), frequency-adaptive modulation (Cui et al., 2025), multi-granularity estimation (Lin et al., 2026), and retrieval-based knowledge (Wang et al., 2026a). However, these methods are not explicitly designed for spatially varying channel-heterogeneous degradation, especially gradual deterioration that is difficult to localize from the corrupted channel alone. We instead leverage structurally preserved channels through a geometry-aware prior to identify cross-channel inconsistency and guide restoration toward degraded regions.

Reference-based Color Mapping. Reference-based color mapping transfers chromatic and tonal characteristics from a reference image to a target image while preserving its underlying content and structure. Classical approaches primarily transfer global or local color statistics and distributions (Reinhard et al., 2001; Pitie et al., 2005; Wang et al., 2010). Later learning-based methods improve structural preservation and reference-style modeling through deep feature transformations, photorealistic constraints, or learned color mappings (Luan et al., 2017; Yoo et al., 2019; Ho & Zhou, 2021; Ke et al., 2023; Ko et al., 2026). Recent approaches further introduce explicit interactions between the source and reference to achieve content-aware color mapping. ColorFM (He et al., 2026) employs semantic-prior-guided hierarchical color coupling and cross-attention to improve semantic consistency between source and reference, while SA-LUT (Gong et al., 2025) uses content–style cross-attention to generate a spatial context map for adaptive color transformation with a 4D LUT. Unlike prior methods that introduce semantic or appearance features for color transfer, we reuse the same geometry-aware representation that drives restoration to establish cross-image reference cues, providing a unified geometric prior across restoration and color reconstruction.

3 METHOD

We introduce our overall negative restoration pipeline in Fig. 2, which consists of two stages: blue-channel restoration and printed-color mapping. Given a degraded negative, our geometry-guided restoration framework first estimates a degradation-aware uncertainty map from cross-channel discrepancies in geometric features (Sec. 3.1) and uses it to guide restoration of the corrupted blue channel with a U-shaped encoder–decoder built with NAF blocks (Chen et al., 2022) (Sec. 3.2). The restored negative is then transformed into a print-like image through reference-based color mapping guided by geometric correspondence cues (Sec. 3.3). Despite addressing different tasks, both stages follow a common geometry-guided design.

3.1 CROSS-CHANNEL GEOMETRIC INCONSISTENCY ESTIMATION

A key challenge in restoring the corrupted blue channel is determining where degradation occurs and how severely each region is affected. Since the three color channels are spatially aligned observations of the same scene, they should preserve consistent underlying geometry despite their different pixel values. This suggests a simple cue: when the blue channel is severely corrupted, its recovered geometry should deviate from that of the better-preserved R and G channels. A straightforward way to exploit this cue is to compare channel-wise depth predictions. However, predicted depth maps provide only limited sensitivity to such degradation. We therefore exploit intermediate geometry-aware features from the 3D foundation model, where degradation-induced discrepancies remain more evident (Fig. 3), to construct spatial restoration guidance.

Specifically, we adopt DA3 (Lin et al., 2025) as a frozen feature extractor. Given a degraded negative $I^d_{\text{neg}} \in \mathbb{R}^{3\times H\times W}$ with channels R, G, and B_d, where d denotes deterioration, we replicate each channel into a three-channel view and jointly process the three views with DA3:

$$\begin{gathered} V_{\text{R}} = [R, R, R], \qquad V_{\text{G}} = [G, G, G], \qquad V_{\text{B}} = [B_d, B_d, B_d], \\ (D_{\text{R}}, D_{\text{G}}, D_{\text{B}_d}) = \Phi_{L_i}(V_{\text{R}}, V_{\text{G}}, V_{\text{B}}). \end{gathered} \tag{1}$$

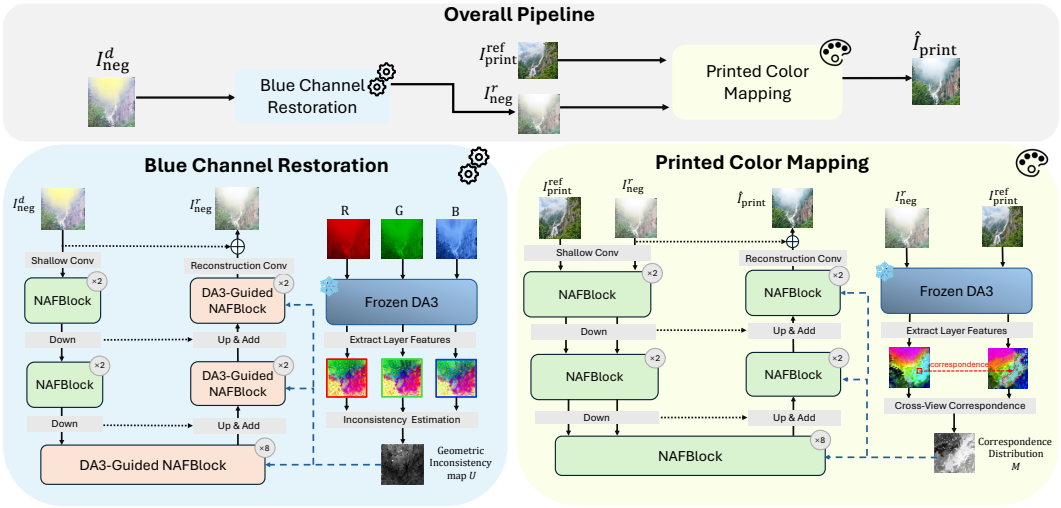

Figure 2: Pipeline overview. We first restore the degraded blue channel using cross-channel geometric inconsistency derived from frozen DA3 features (bottom left), and then reconstruct printed colors through reference-based color mapping guided by cross-view geometric correspondence (bottom right).

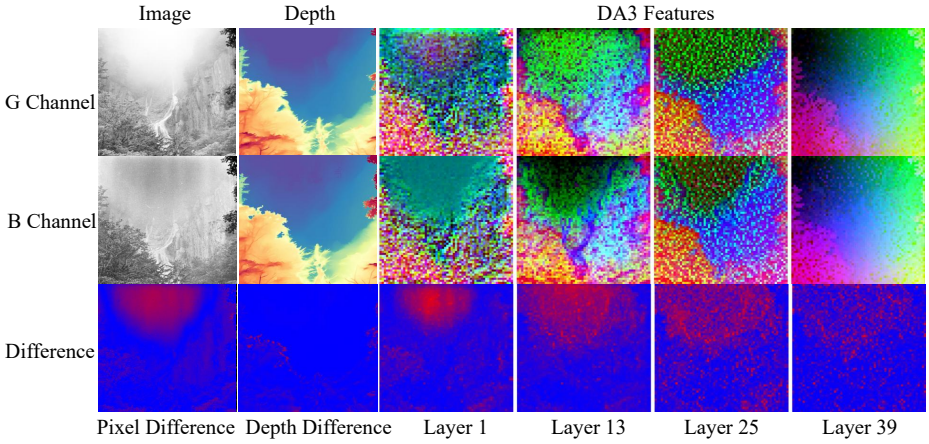

Figure 3: Comparison of cross-channel discrepancies in image, depth, and intermediate DA3 feature spaces. The top two rows are to the G- and degraded-B-channel's pixel images, depth predictions and feature maps, and the bottom row shows their normalized differences. Direct pixel differences respond strongly to inherent appearance variations, whereas predicted depths remain similar despite substantial degradation. Intermediate features retain clearer degradation-related discrepancies for localization.

where $V_{\mathrm{R}}, V_{\mathrm{G}}, V_{\mathrm{B}}$ are the replicated channel views, and Φ_{L_i} denotes joint DA3 processing up to layer L_i, followed by token extraction and spatial reshaping. Joint processing allows DA3's cross-view attention to exchange information among the spatially aligned channel observations, so that their feature maps are formed under a shared cross-view context rather than inferred independently. The resulting feature maps $D_{\mathrm{R}}, D_{\mathrm{G}}, D_{\mathrm{B}_d} \in \mathbb{R}^{C\times h\times w}$ encode the corresponding channel-wise geometric features, where C is the DA3 feature dimension and $h \times w$ is the spatial resolution of the feature maps. For a spatial location q, $D_{\mathrm{R}}(q)$, $D_{\mathrm{G}}(q)$, and $D_{\mathrm{B}_d}(q) \in \mathbb{R}^C$ denote the corresponding token features. Since the three channel views are spatially aligned, token features at the same q correspond to the same image region and can therefore be directly compared.

To leverage the structural information preserved in the R and G channels, we then combine their feature maps into a reference representation:

$$D_{\mathrm{RG}} = (1-w)D_{\mathrm{R}} + wD_{\mathrm{G}}, \qquad w = \sigma(\rho), \tag{2}$$

where ρ is a learnable scalar, and σ is the sigmoid function, yielding a global mixing weight shared across spatial locations. Using this reference, we compute the cosine distance between the blue-channel token feature and the corresponding reference feature at each spatial location:

$$U(q) = \mathcal{N}_{\mathrm{mm}} \left(1 - \frac{D_{\mathrm{B}_d}(q)^\top D_{\mathrm{RG}}(q)}{\|D_{\mathrm{B}_d}(q)\|_2 \|D_{\mathrm{RG}}(q)\|_2 + \epsilon} \right), \tag{3}$$

where ϵ ensures numerical stability and $\mathcal{N}_{\mathrm{mm}}$ denotes per-image min–max normalization over the patch grid. The resulting cross-channel geometric inconsistency map $U \in [0,1]^{1\times h\times w}$ provides spatial restoration guidance, with larger values indicating stronger local inconsistency. It measures relative discrepancy within an image rather than a calibrated degradation probability.

The choice of DA3 extraction layer aims to balance cross-channel information exchange and sensitivity to local differences. Shallow features are coarse-grained and lack sufficient inter-channel interaction, hindering precise degradation localization, while deeper layers progressively homogenize channel representations and weaken local discrepancies. As shown in Fig. 3, this trade-off is reflected in the feature visualizations and normalized difference maps, with further analysis in Appendix B. We therefore choose $L_i = 13$ to retain degradation-relevant local structure and stable restoration guidance.

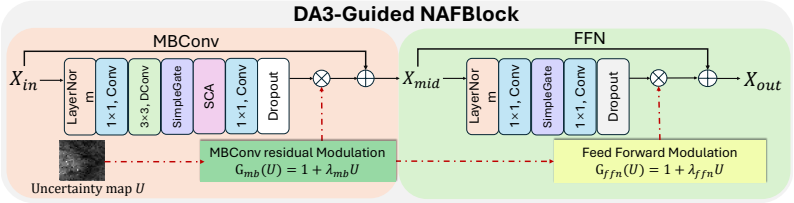

Figure 4: DA3-guided NAFBlock. The uncertainty map modulates the residual updates of the MBConv and FFN branches while preserving the identity paths.

3.2 DA3-GUIDED NAFBLOCK

The uncertainty map localizes blue-channel degradation, but it must be translated into spatially adaptive corrections. Since NAFBlock performs feature refinement through residual corrections, a natural way to exploit the uncertainty map is to spatially control the strength of these corrections. The standard NAFBlock (Chen et al., 2022) consists of two successive residual branches (MBConv branch and a feed-forward network (FFN)). However, neither of them is explicitly conditioned on degradation. We therefore introduce the *DA3-guided NAFBlock* (Fig. 4), which uses the uncertainty map to modulate the residual corrections while leaving the identity paths unchanged. In this way, the block adapts how strongly each spatial region is modified according to its estimated degradation.

Given an input feature X_{in} and an uncertainty map U aligned with restoration stage s, let $\mathcal{M}$ and $\mathcal{F}$ denote the MBConv and FFN transformations, respectively. We construct two unit-centered spatial gains and apply them to the corresponding residual updates:

$$\begin{aligned} G_{\mathrm{mb}}(U) &= 1 + \lambda_{\mathrm{mb}} U, & G_{\mathrm{ffn}}(U) &= 1 + \lambda_{\mathrm{ffn}} U, \\ X_{\mathrm{mid}} &= X_{\mathrm{in}} + \beta \odot \mathcal{M}(X_{\mathrm{in}}) \odot G_{\mathrm{mb}}(U), & X_{\mathrm{out}} &= X_{\mathrm{mid}} + \gamma \odot \mathcal{F}(X_{\mathrm{mid}}) \odot G_{\mathrm{ffn}}(U), \end{aligned} \tag{4}$$

Here, β and γ are the channel-wise residual scales inherited from NAFBlock, whereas λ_{mb} and λ_{ffn} independently control the spatial modulation of the two branches. This factorization retains the original channel-wise scaling while introducing DA3-derived spatial guidance, which is broadcast across feature channels. Because the gains are centered at one, U modulates only the residual corrections rather than masking the propagated features or identity paths. When $U = 0$, or when both modulation strengths are zero, the standard NAFBlock update is recovered.

As shown in the bottom-left of Fig. 2, guided blocks are used at the bottleneck and throughout the decoder of the NAF-based UNet, while the encoder retains standard NAFBlocks. At each decoder stage, U is expanded to the feature resolution by repeating each patch score over its corresponding spatial region.

3.3 DA3-GUIDED CROSS-VIEW COLOR MAPPING

Although the restoration network can substantially recover the corrupted structure of the B channel, the restored negative remains in the negative color domain and differs visually from a printed photo. We therefore introduce a reference-guided color mapping module to transform the restored negative into the printed-color domain, as shown in the bottom right of Fig. 2. Our observation is that a reference printed photo may share geometric or semantic correspondences with the input negative. These correspondences provide localized cues for how printed-domain colors should be assigned to different regions of the restored negative. Instead of applying a single global color adjustment, we again leverage DA3 (Lin et al., 2025) to estimate cross-view correspondences and use the appearance of corresponding reference regions to guide color mapping of the restored negative.

Let $T_{\mathrm{neg}} \in \mathbb{R}^{N_{\mathrm{neg}} \times C}$ and $T_{\mathrm{ref}} \in \mathbb{R}^{N_{\mathrm{ref}} \times C}$ denote the DA3 token sequences of the restored negative and reference print after spatial flattening, where each token is ℓ_2-normalized along the feature dimension. For each restored-negative token $T_{\mathrm{neg},i} \in \mathbb{R}^C$, we compute its cosine similarities with all reference tokens $T_{\mathrm{ref},j} \in \mathbb{R}^C$ and normalize them across j using a temperature τ. This produces the cross-view affinity matrix $\mathbf{M}$:

$$M_{ij} = \frac{\exp\left(T_{\mathrm{neg},i}^{\top} T_{\mathrm{ref},j}/\tau\right)}{\sum_{k=1}^{N_{\mathrm{ref}}} \exp\left(T_{\mathrm{neg},i}^{\top} T_{\mathrm{ref},k}/\tau\right)}, \qquad \mathbf{M} \in \mathbb{R}^{N_{\mathrm{neg}} \times N_{\mathrm{ref}}}, \tag{5}$$

where each row of $\mathbf{M}$ defines a correspondence distribution over the reference tokens.

Let $F_{\mathrm{ref}} \in \mathbb{R}^{N_{\mathrm{ref}} \times C_{\mathrm{naf}}}$ and $F_{\mathrm{neg}} \in \mathbb{R}^{N_{\mathrm{neg}} \times C_{\mathrm{naf}}}$ denote the intermediate feature maps from the shared NAF encoder for the reference and restored-negative images, respectively, where C_{naf} is the feature dimension. These features encode local appearance cues for correspondence-guided color mapping. We aggregate correspondence-weighted reference features onto the restored-negative view as

$$Z_{\mathrm{ref}} = \mathbf{M} F_{\mathrm{ref}} \in \mathbb{R}^{N_{\mathrm{neg}} \times C_{\mathrm{naf}}}. \tag{6}$$

The resulting feature Z_{ref} provides correspondence-aware appearance information for each restored-negative location. When no dominant correspondence exists, the affinity distribution becomes diffuse, so Z_{ref} aggregates broader reference appearance and provides a more global color cue instead of forcing an unreliable local match. Then the transferred reference appearance features are concatenated with the current restored-negative features and passed through an MLP to predict a residual appearance adjustment. We denote the updated feature as F'_{neg}:

$$F'_{\mathrm{neg}} = F_{\mathrm{neg}} + \mathrm{MLP}\left([F_{\mathrm{neg}}; Z_{\mathrm{ref}}]\right), \tag{7}$$

where $[\cdot;\cdot]$ denotes channel-wise concatenation. This residual formulation transfers reference appearance information while preserving the spatial structure of the restored negative. We apply the modulation at the bottleneck and two decoder stages.

3.4 TRAINING OBJECTIVE

Given I^d_{neg} and its negative-restoration target I^{gt}_{neg}, the restoration network predicts $\widehat{B}$ and constructs the three-channel output $I^r_{\mathrm{neg}} = [R, G, \widehat{B}]$. It is trained with

$$\mathcal{L}_{\mathrm{restoration}} = \mathcal{L}_1\left(I^r_{\mathrm{neg}}, I^{gt}_{\mathrm{neg}}\right) + \lambda_{\mathrm{fft}} \mathcal{L}_{\mathrm{fft}}\left(I^r_{\mathrm{neg}}, I^{gt}_{\mathrm{neg}}\right). \tag{8}$$

Both terms supervise all three RGB channels of I^r_{neg} against I^{gt}_{neg}; the unchanged R/G channels are included on the loss through the merged RGB output. Here, $\mathcal{L}_1$ is the mean pixel-wise ℓ_1 loss, while $\mathcal{L}_{\mathrm{fft}}$ is the mean ℓ_1 distance between the real and imaginary components of the two-dimensional Fourier transforms of the two images. We jointly optimize the restoration parameters θ and the red–green mixture parameter ρ, while keeping DA3 frozen.

The color mapping network is supervised using the target-view ground-truth print I^{gt}_{print}:

$$\mathcal{L}_{\mathrm{color}} = \mathcal{L}_1\left(\widehat{I}_{\mathrm{print}}, I^{gt}_{\mathrm{print}}\right) + \lambda_{\mathrm{fft}} \mathcal{L}_{\mathrm{fft}}\left(\widehat{I}_{\mathrm{print}}, I^{gt}_{\mathrm{print}}\right). \tag{9}$$

We use $\lambda_{\mathrm{fft}} = 0.125$. The intermediate feature residuals receive no direct supervision and are optimized only through the loss on $\widehat{I}_{\mathrm{print}}$.

Table 1: Quantitative comparison with state-of-the-art image restoration methods and geometric-guidance ablation (bottom block). Best and second-best results are highlighted in **bold** and underline, respectively.

	Blue Channel				Overall			
Method	LPIPS ↓	PSNR ↑	SSIM ↑	FID ↓	LPIPS ↓	PSNR ↑	SSIM ↑	FID ↓
Restormer (Zamir et al., 2022)	0.0381	27.7210	0.9773	22.5711	0.0280	32.4922	0.9924	11.4283
MambaIR (Guo et al., 2024)	0.0495	25.6927	0.9626	26.3474	0.0386	30.4639	0.9875	23.8103
UniRestore (Chen et al., 2025)	0.1288	25.0592	0.7667	44.6613	0.0754	29.8304	0.9222	19.7145
RDBM (Wang et al., 2026b)	0.0725	26.6570	0.9543	32.1204	0.0373	31.4282	0.9848	16.0440
AirNet (Li et al., 2022)	0.0570	27.2181	0.9656	31.7917	0.0349	31.9893	0.9885	14.0222
PromptIR (Potlapalli et al., 2023)	0.1005	26.8679	0.9218	38.4411	0.0429	31.6391	0.9739	14.6260
AdaIR Cui et al. (2025)	0.0393	27.6194	0.9771	22.6578	0.0284	32.3906	0.9924	13.1839
R2R (Wang et al., 2026a)	0.0315	28.8594	0.9820	**20.8261**	0.0231	33.6306	0.9940	9.6919
w/o Geometric Guidance	0.0336	28.0243	0.9817	24.3138	0.0250	32.7955	0.9939	11.5307
Pixel-level Guidance	0.0316	28.3090	0.9822	21.9238	0.0240	33.0802	0.9941	11.0381
DINO Guidance	0.0306	28.9843	0.9828	22.6598	0.0229	33.7555	0.9943	10.0952
Ours (DA3 Guidance)	**0.0295**	**29.5605**	**0.9833**	21.3199	**0.0210**	**34.3317**	**0.9944**	**8.9825**

4 EXPERIMENTS

This section describes the implementation and training setup (Sec. 4.1), compares our method with state-of-the-art approaches (Sec. 4.2) for negative-film restoration and reference-conditioned color mapping, and analyzes the contributions of the proposed components through ablation studies (Sec. 4.3).

4.1 IMPLEMENTATION DETAILS

We train our framework in PyTorch on a single NVIDIA A100 GPU in two sequential stages, without joint optimization and with DA3 frozen throughout. The restoration network predicts only $\widehat{B}$, which is combined with the unchanged R/G channels of I^{d}_{neg} to form I^{r}_{neg} for subsequent color-mapping training. For restoration, we use an align-scale NAFNet with width 32 and layer-13 DA3 features for geometric guidance, trained for 5,800 iterations on 504×504 crops using AdamW with an initial learning rate of 2×10^{-4}. We train on BlueNeg (Liu et al., 2025), adding synthetic spatially smooth B-channel fading to augment scarce border-fading cases, yielding 6,800 fixed training samples. For reference-conditioned print color mapping, layer-19 DA3 features estimate cross-view affinity between the restored negative and reference print. We train on 1,326 pairs, each comprising a restored negative I^{r}_{neg}, its target-view printed ground truth I^{gt}_{print}, and a reference print I^{ref}_{print} from the same film roll. The reference depicts either the same scene from a different viewpoint or a different scene from the roll. Full architectural configurations, optimization settings, and augmentation details are provided in Appendix A.

4.2 COMPARISON

Negative-film Restoration. We compare with the general IR methods Restormer (Zamir et al., 2022), MambaIR (Guo et al., 2024), UniRestore (Chen et al., 2025), and RDBM (Wang et al., 2026b), and all-in-one IR methods AirNet (Li et al., 2022), PromptIR (Potlapalli et al., 2023), AdaIR (Cui et al., 2025), and R2R (Wang et al., 2026a). For fairness, we retrain supervised baselines on the same BlueNeg degraded-negative/pseudo-ground-truth training pairs as our method.

Our method performs best on nearly all B-channel and RGB metrics (Tab. 1), demonstrating the effectiveness of our DA3 guidance in localizing channel-specific corruption for targeted restoration. Compared with the strongest baseline, R2R, our method improves both B-channel and RGB PSNR by 0.70 dB and reduces overall LPIPS by approximately 9.1%. In Fig. 5, baselines struggle with large, gradually degraded regions in the first row, often only attenuating the yellowish corruption or spreading it to nearby areas. For high-frequency degradation in the second and third rows, several baselines leave noticeable local pattern artifacts, while UniRestore exhibits local distortions and loses fine structural details, as its restoration relies on priors from a pretrained generative model. Our method more effectively removes both continuous degradation over large regions and local artifacts while preserving fine structures.

Reference-conditioned Print Color Mapping. We compare with photorealistic style-transfer methods PhotoWCT^2 (Chiu & Gurari, 2022) and CAP-VSTNet (Wen et al., 2023), spatially adaptive SA-LUT (Gong et al., 2025), flow-matching-based ColorFM (He et al., 2026), and reference-based

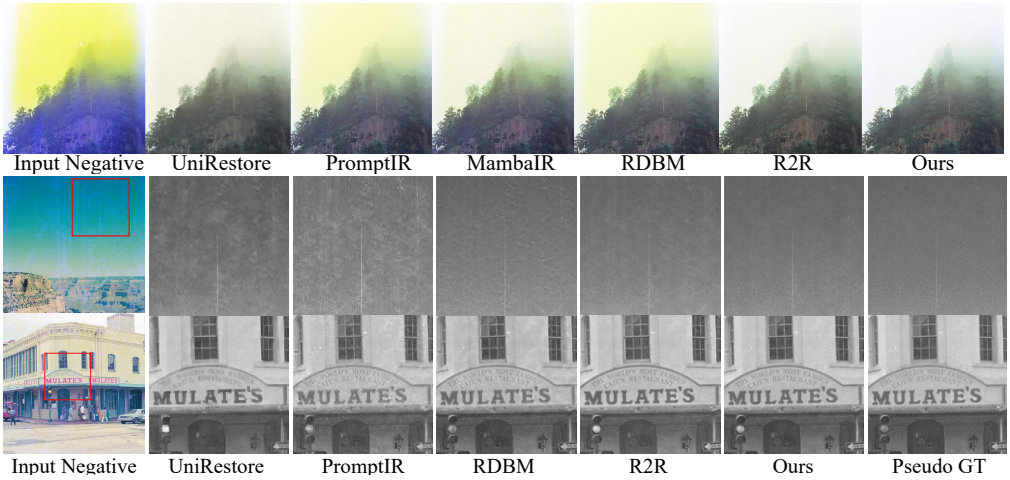

Figure 5: Qualitative comparison of image restoration methods. The top panel compares restored RGB results, while the bottom panel compares restoration details in selected regions.

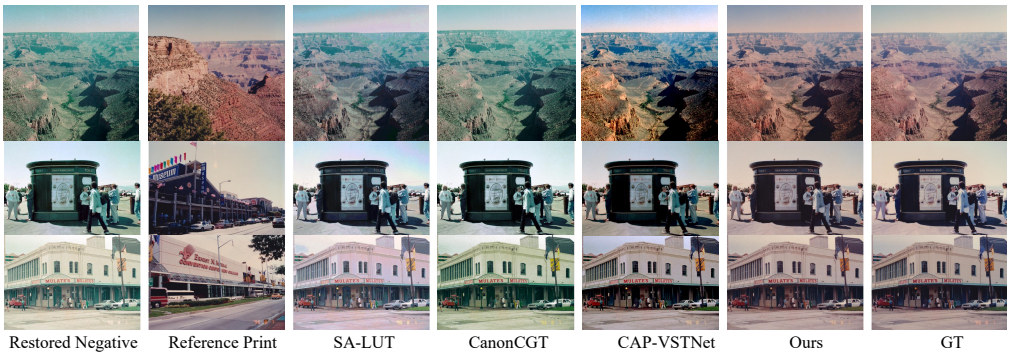

Figure 6: Qualitative comparison of color-mapping methods across three representative scenes.

Table 2: Quantitative comparison with other color-mapping methods and correspondence ablation (bottom block). Best and second-best results are shown in **bold** and underlined, respectively. * denotes results using official pretrained weights, as they require additional supervision unavailable in the BlueNeg dataset for training.

Method	SSIM ↑	PSNR ↑	LPIPS ↓	ΔE ↓	H-Corr ↑	H-Chi ↓
PhotoWCT2 (Chiu & Gurari, 2022)	0.7911	15.9856	0.3713	22.6376	0.0986	0.6795
CAP-VSTNet (Wen et al., 2023)	0.8235	22.5275	0.2441	10.9357	**0.6421**	**0.2146**
SA-LUT (Gong et al., 2025)	0.7364	18.1527	0.3268	18.9029	0.3569	0.3431
ColorFM (He et al., 2026)	0.8007	21.6873	0.2742	11.9347	0.5256	0.2963
Deep Preset * (Ho & Zhou, 2021)	0.6666	14.5346	0.4462	26.6997	0.0228	0.4774
CanonCGT* (Ko et al., 2026)	0.6830	16.4892	0.4349	23.5894	0.0069	0.4634
Uniform Affinity	0.8779	24.1279	0.2095	9.7074	0.5341	0.2982
Ours	**0.8883**	**24.4418**	**0.2008**	**9.6288**	0.5833	0.2629

photographic adjustment and color-grading methods such as Deep Preset (Ho & Zhou, 2021) and CanonCGT (Ko et al., 2026). We use official pretrained weights for Deep Preset and CanonCGT because BlueNeg lacks their required supervision, and retrain the remaining applicable baselines on our training split with default settings.

Our method performs best on most metrics (Tab. 2). Localized cross-image correspondences help account for spatial differences between the reference and target beyond global color statistics. Compared with CAP-VSTNet, our method improves PSNR by 1.91 dB while remaining competitive on H-Corr and H-Chi. In Fig. 6, baselines struggle to transfer the reference's reddish-brown mountain colors to corresponding input regions. CanonCGT and SA-LUT show color-distribution mismatches and occasional halo-like artifacts, while CAP-VSTNet over-brightens local regions and alters luminance relationships. Our method transfers colors more faithfully while preserving input structure and luminance. Additional results are provided in Appendix C.2.



Figure 7: Qualitative ablation of geometric guidance for restoration and reference-conditioned color mapping. DA3 guidance improves removal of large-area fading and local artifacts while preserving fine structure, while geometry-aware affinity yields more faithful reference-conditioned colors than uniform affinity.

4.3 Ablation Studies

Negative-film Restoration We ablate the geometry-aware guidance to assess its contribution to negative-film restoration. The results are shown in the bottom block of Tab. 1. First, we remove the three-channel DA3 (Lin et al., 2025) descriptor extraction and cross-channel geometric inconsistency estimation, reducing the model to a standard U-Net composed solely of NAFBlocks. We further introduce a pixel-guided variant to determine whether the improvement arises merely from explicit inter-channel differences rather than from geometry-aware representations. Specifically, we compute $G - B_d$ and $R - B_d$ at the pixel level, combine them using the same learnable weighting scheme as in the full model, and encode the resulting features using NAFBlocks at the corresponding stage. The encoded features are then normalized to produce a pixel-derived guidance map. In addition, we replace DA3 descriptors with DINOv2 (Oquab et al., 2024) features while keeping the remaining guidance pipeline unchanged, in order to examine whether generic visual representations provide comparable cross-channel degradation cues.

As shown in Tab. 1 and Fig. 7, removing DA3 guidance substantially weakens the model's ability to localize and correct large, gradually varying degradation, while the pixel-guided variant also yields lower performance. Although pixel-level channel differences provide direct degradation cues, they are confounded by intrinsic color variations and therefore cannot reliably distinguish degradation from scene content. DINO-based guidance performs better than pixel-level guidance but remains inferior to DA3. We attribute this to the channel-heterogeneous nature of negative-film degradation: geometry-oriented DA3 features provide more discriminative and spatially consistent cross-channel correspondences, making structural inconsistencies caused by degradation easier to identify.

Color Mapping For color mapping, we ablate geometric priors for correspondence estimation by replacing each token's affinity distribution in Eq. 5 with a uniform distribution, i.e., $M_{ij} = 1/N_{\text{ref}}$. This assigns equal weight to all reference tokens, eliminating selective aggregation from corresponding reference regions. As shown in the bottom block of Tab. 2, this variant reduces color fidelity and weakens agreement with corresponding regions in the reference image. For example, in Fig. 7, the Uniform Affinity variant tints mountain regions an unnatural blue because it cannot selectively retrieve appearance from geometrically corresponding reference tokens. These results support the role of geometry-aware correspondences in guiding spatially coherent color mapping.

5 Conclusion

In this paper, we present a geometry-guided framework for negative-film restoration and print-like appearance reconstruction. Our key insight is that channel-heterogeneous deterioration disrupts the geometric consistency naturally shared across the R, G, and B channels. By measuring cross-channel discrepancies in intermediate DA3 (Lin et al., 2025) features, our method derives spatial guidance to modulate restoration residuals according to local B-channel degradation. We further exploit geometric and semantic associations between the restored negative and a reference printed photo for coherent reference-conditioned print color mapping. Extensive experiments demonstrate that our approach outperforms existing restoration and color-mapping pipelines, highlighting the potential of pretrained geometric priors for recovering both the structure and appearance of deteriorated film. Despite these gains, leveraging DA3 introduces additional computational and memory overhead, motivating more lightweight geometric priors in future work.

AI USE STATEMENT

In this work, we used generative AI tools to assist with language editing and polishing, literature search, coding, and drafting portions of the manuscript.

We did not use generative AI tools to generate synthetic datasets or to establish or prove mathematical claims.

All AI-assisted outputs were carefully reviewed and verified by the authors. In particular, references identified with AI assistance were independently checked against their original sources to ensure their existence, relevance, and accurate representation. AI-assisted drafts and language revisions were reviewed to ensure that they faithfully reflect the actual methodology, experimental setup, results, and conclusions of this work. AI-assisted code was also inspected and validated by the authors before use.

We take full responsibility for the final content of this work, including all text, claims, code, and other artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide detailed descriptions of the proposed restoration and reference-guided color mapping frameworks in the method section (Sec. 3). Dataset construction, preprocessing, evaluation protocols, and implementation details are documented in the experiment section (Sec. 4). We also describe the training and adaptation procedures used for the compared baselines to ensure a consistent evaluation setting. These materials are intended to provide sufficient information for reproducing the main experimental results reported in this work.

REFERENCES

Vittoria Bruni and Domenico Vitulano. A generalized model for scratch detection. *IEEE transactions on image processing*, 13(1):44–50, 2004.

Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE transactions on image processing*, 25(11):5187–5198, 2016.

Weiwei Cai, Huaidong Zhang, Xuemiao Xu, Shengfeng He, Kun Zhang, and Jing Qin. Contextual-assisted scratched photo restoration. *IEEE Transactions on circuits and systems for video technology*, 33(10):5458–5469, 2023.

Weiwei Cai, Xuemiao Xu, Jiajia Xu, Huaidong Zhang, Haoxin Yang, Kun Zhang, and Shengfeng He. Hierarchical damage correlations for old photo restoration. *Information Fusion*, 107:102340, 2024.

Rong-Chi Chang, Yun-Long Sie, Su-Mei Chou, and Timothy K Shih. Photo defect detection for image inpainting. In *Seventh IEEE International Symposium on Multimedia (ISM'05)*, pp. 5–pp. IEEE, 2005.

I-Hsiang Chen, Wei-Ting Chen, Yu-Wei Liu, Yuan-Chun Chiang, Sy-Yen Kuo, and Ming-Hsuan Yang. Unirestore: Unified perceptual and task-oriented image restoration model using diffusion prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17969–17979, 2025.

Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. *arXiv preprint arXiv:2204.04676*, 2022.

Tai-Yin Chiu and Danna Gurari. Photowct 2: Compact autoencoder for photorealistic style transfer resulting from blockwise training and skip connections of high-frequency residuals. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2978–2987. IEEE, 2022.

Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4621–4630. IEEE, 2021.

Marcos V Conde, Gregor Geigle, and Radu Timofte. Instructir: High-quality image restoration following human instructions. In *European Conference on Computer Vision*, pp. 1–21. Springer, 2024.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Shahbaz Khan. AdaIR: Adaptive all-in-one image restoration via frequency mining and modulation. In *The Thirteenth International Conference on Learning Representations*, 2025.

Zerui Gong, Zhonghua Wu, Qingyi Tao, Qinyue Li, and Chen Change Loy. Sa-lut: spatial adaptive 4d look-up table for photorealistic style transfer. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 18294–18303. IEEE, 2025.

Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. Mambair: A simple baseline for image restoration with state-space model. In *European conference on computer vision*, pp. 222–241. Springer, 2024.

Yuhang He, Kai Zhang, Xiaoming Li, Du Chen, and Jian Yang. Colorfm: An optimization-to-learning framework for color transfer via flow matching. *arXiv preprint arXiv:2607.07119*, 2026.

Man M Ho and Jinjia Zhou. Deep preset: Blending and retouching photos with color style transfer. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 2112–2120. IEEE, 2021.

Zhanghan Ke, Yuhao Liu, Lei Zhu, Nanxuan Zhao, and Rynson WH Lau. Neural preset for color style transfer. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14173–14182. IEEE, 2023.

Jinwon Ko, Keunsoo Ko, and Chang-Su Kim. Canoncgt: Reference-based color grading via canonical pivot representation. *arXiv preprint arXiv:2606.01638*, 2026.

Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *2017 IEEE international conference on computer vision (ICCV)*, pp. 4780–4788. Ieee, 2017.

Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17431–17441. IEEE, 2022.

Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *2021 IEEE/CVF international conference on computer vision workshops (ICCVW)*, pp. 1833–1844. IEEE, 2021.

Haotong Lin, Sili Chen, Junhao Liew, Donny Y Chen, Zhenyu Li, Guang Shi, Jiashi Feng, and Bingyi Kang. Depth anything 3: Recovering the visual space from any views. *arXiv preprint arXiv:2511.10647*, 2025.

Jingbo Lin, Zhilu Zhang, Wenbo Li, Renjing Pei, Hang Xu, Hongzhi Zhang, and Wangmeng Zuo. Unirestorer: Universal image restoration via adaptively estimating image degradation at proper granularity. In *International Conference on Learning Representations*, volume 2026, pp. 2369–2409, 2026.

Hanyuan Liu, Chengze Li, Minshan Xie, Zhenni Wang, Jiawen Liang, Chi-Sing Leung, and Tien-Tsin Wong. Blueneg: A 35mm negative film dataset for restoring channel-heterogeneous deterioration. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13119–13128, 2025. doi: 10.1109/ICCV51701.2025.01219.

Fujun Luan, Sylvain Paris, Eli Shechtman, and Kavita Bala. Deep photo style transfer. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6997–7005. IEEE, 2017.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Photo-realistic image restoration in the wild with controlled vision-language models. *arXiv preprint arXiv:2404.09732*, 2024.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. Dinov2: Learning robust visual features without supervision, 2024. URL `https://arxiv.org/abs/2304.07193`.

Francois Pitie, Anil C Kokaram, and Rozenn Dahyot. N-dimensional probability density function transfer and its application to color transfer. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pp. 1434–1439. IEEE, 2005.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in neural information processing systems*, 36:71275–71293, 2023.

Erik Reinhard, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer graphics and applications*, 21(5):34–41, 2001.

Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3932–3941. IEEE, 2019.

Ziyu Wan, Bo Zhang, Dongdong Chen, Pan Zhang, Dong Chen, Jing Liao, and Fang Wen. Bringing old photos back to life. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2744–2754. IEEE, 2020.

Ziyu Wan, Bo Zhang, Dong Chen, Pan Zhang, Fang Wen, and Jing Liao. Old photo restoration via deep latent space translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):2071–2087, 2022.

Baoyuan Wang, Yizhou Yu, Tien-Tsin Wong, Chun Chen, and Ying-Qing Xu. Data-driven image color theme enhancement. *ACM Transactions on Graphics*, 29(6), 2010.

Chenxu Wang, Kai Zhang, and Jian Yang. Retrieve-to-restore: Efficient all-in-one image restoration with a retrieval-based degradation bank. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1277–1287, 2026a.

Cong Wang, Jinshan Pan, Wei Wang, Jiangxin Dong, Mengzhu Wang, Yakun Ju, and Junyang Chen. Promptrestorer: A prompting image restoration method with degradation perception. *Advances in Neural Information Processing Systems*, 36:8898–8912, 2023.

Hebaixu Wang, Jing Zhang, Haoyang Chen, Haonan Guo, Di Wang, Jiayi Ma, and Bo Du. Residual diffusion bridge model for image restoration. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 8375–8386, 2026b.

Linfeng Wen, Chengying Gao, and Changqing Zou. Cap-vstnet: Content affinity preserved versatile style transfer. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18300–18309. IEEE, 2023.

Runsheng Xu, Zhengzhong Tu, Yuanqi Du, Xiaoyu Dong, Jinlong Li, Zibo Meng, Jiaqi Ma, Alan Bovik, and Hongkai Yu. Pik-fix: Restoring and colorizing old photos. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 1724–1734. IEEE, 2023.

Jaejun Yoo, Youngjung Uh, Sanghyuk Chun, Byeongkyu Kang, and Jung-Woo Ha. Photorealistic style transfer via wavelet transforms. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9035–9044. IEEE, 2019.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 14816–14826. IEEE, 2021.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728–5739, 2022.

Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26 (7):3142–3155, 2017.

Yuzhi Zhao, Lai-Man Po, Tingyu Lin, Xuehui Wang, Kangcheng Liu, Yujia Zhang, Wing-Yin Yu, Pengfei Xian, and Jingjing Xiong. Legacy photo editing with learned noise prior. In *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 2102–2111. IEEE, 2021.

Dian Zheng, Xiao-Ming Wu, Shuzhou Yang, Jian Zhang, Jian-Fang Hu, and Wei-Shi Zheng. Selective hourglass mapping for universal image restoration based on diffusion model. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 25445–25455. IEEE, 2024.

A Implementation Details

A.1 Restoration Architecture

The restoration network is implemented as an align-scale NAFNet with base width 32. It contains 2 NAFBlocks at the full-resolution encoder stage and 2 NAFBlocks at the half-resolution encoder stage, followed by 8 NAFBlocks at the DA3-aligned $1/14$ bottleneck. The decoder contains 2 NAFBlocks at the half-resolution stage and another 2 NAFBlocks at the full-resolution stage.

We extract geometric guidance from layer 13 of the frozen DA3 (Lin et al., 2025) backbone and adopt the once-persist strategy throughout the restoration network. For each guided block in the bottleneck and decoder, the uncertainty map U is expanded to the corresponding feature resolution by repeating each patch score over its spatial region. The channel-wise residual scales β and γ inherited from NAFBlock are initialized to zero. The modulation strengths λ_{mb} and λ_{ffn} are learnable scalar gains applied to the MBConv and feed-forward residual branches, respectively. For red–green feature aggregation, the mixing parameter is initialized with $\rho = 0$, corresponding to $w = \sigma(\rho) = 0.5$.

The restoration network predicts only the blue channel $\widehat{B}$. The restored negative I^r_{neg} is constructed by combining $\widehat{B}$ with the unchanged R/G channels from the degraded input I^d_{neg}.

A.2 Optimization and Data Augmentation

All models are implemented in PyTorch and trained. The restoration stage is trained for 5,800 iterations with batch size 4 and patch size 504×504. We use AdamW with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and zero weight decay. The initial learning rate is 2×10^{-4}. We apply a 200-iteration warm-up followed by cosine annealing to 10^{-8}. The restoration objective combines an ℓ_1 reconstruction loss with a Fourier-domain loss:

$$\mathcal{L} = \mathcal{L}_1 + 0.125\,\mathcal{L}_{\mathrm{FFT}}. \tag{10}$$

We use BlueNeg (Liu et al., 2025) as the restoration training dataset. Because pseudo-ground-truth examples exhibiting strong border fading are relatively scarce, we augment the training set with synthetic B-channel fading. Specifically, half of the pseudo-ground-truth images are randomly selected for degradation synthesis. For each selected image, we apply a broad and spatially smooth mask to the B channel. The mask is generated using either an elliptical or blob-shaped region, with its spatial radius sampled from $[0.25, 1.0]$ times the shorter image dimension and its degradation strength sampled from $[0.88, 1.0]$. The spatial falloff is randomly selected from cosine, linear, Gaussian, and power profiles. This synthesis procedure results in a fixed training set of 6,800 restoration samples.

During training, we further apply online geometric augmentation. Horizontal and vertical flips are independently applied with probability 0.5, and a rotation index $k \in 0, 1, 2, 3$ is uniformly sampled, corresponding to rotations of $0°$, $90°$, $180°$, and $270°$, respectively.

A.3 Reference-conditioned Print Color Mapping

The color-mapping stage is trained after the restoration network has converged, with no joint optimization between stages and with DA3 frozen throughout. Each training sample consists of a restored negative I^r_{neg}, its target-view printed ground truth I^{gt}_{print}, and a reference printed photo I^{ref}_{print} from the same film roll; the reference may depict the same scene from another viewpoint or a different scene on the roll. Images from the same roll share film stock and laboratory processing, so they still provide consistent print-domain color statistics even when scene content differs. We train on 1,326 such pairs.

B Additional Analysis of Geometric Guidance

We further analyze how the DA3 extraction layer affects cross-channel geometric guidance and downstream restoration quality. As discussed in Sec. 3.1, shallow descriptors lack sufficient cross-view interaction, whereas deeper layers progressively suppress localized cross-channel discrepancies. Layer 12 lies immediately before the first cross-view attention, so its inconsistency maps are

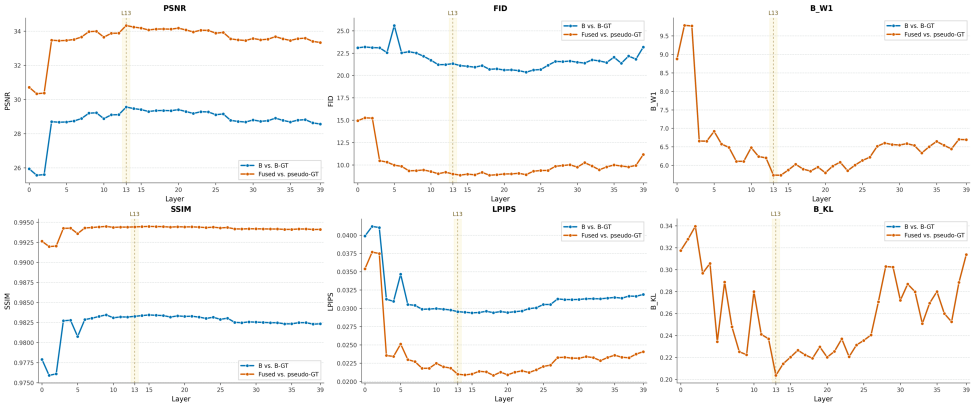

Figure 8: Restoration performance under geometric guidance extracted from different DA3 layers. Layer 13 provides the best overall trade-off between cross-channel sensitivity and stable spatial guidance.

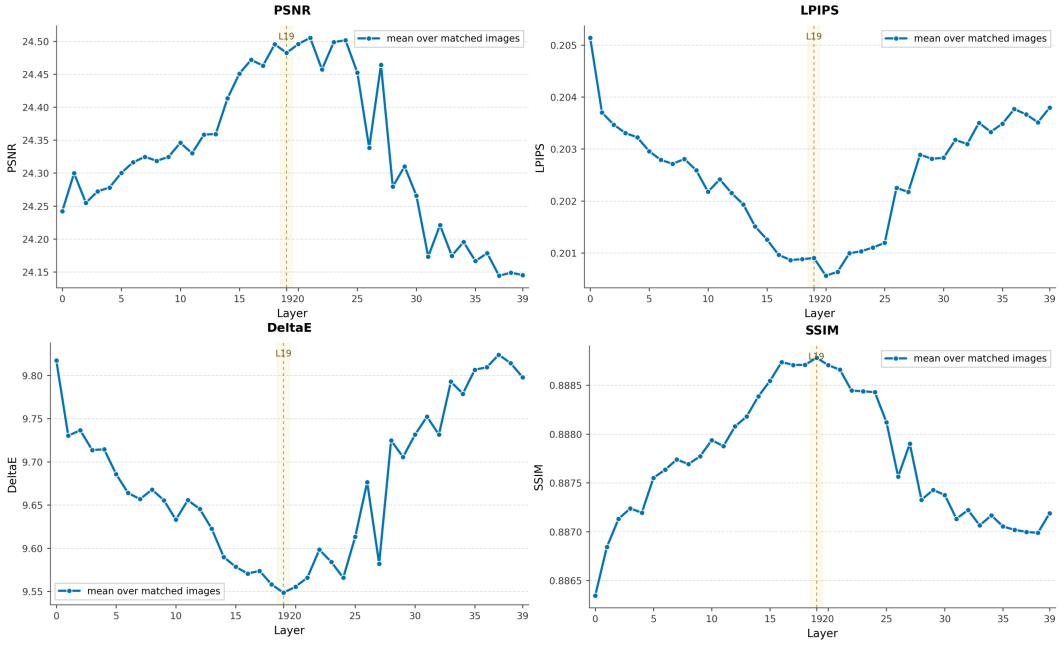

Figure 9: Color-mapping performance under cross-view affinity extracted from different DA3 layers. Layer 19 provides the best overall trade-off for reference-conditioned print color mapping.

highly sensitive to local differences; layer 13 follows this interaction and yields slightly smoother but more stable guidance. Fig. 8 reports restoration metrics when geometric guidance is extracted from different intermediate layers. Although layer 12 can produce sharper inconsistency responses, layer 13 consistently achieves better quantitative performance, suggesting that overly sensitive guidance is less compatible with the NAF-based restoration backbone. We therefore adopt layer 13 for restoration guidance. We apply the same layer analysis to reference-conditioned color mapping (Fig. 9). Layer 19 provides the best overall trade-off among the evaluated layers, and we adopt it for cross-view affinity.

C Additional Results

C.1 Negative-film Restoration

Fig. 10 shows additional negative-film restoration comparisons on representative test scenes, including large-area fading and high-frequency B-channel degradation.

C.2 Reference-conditioned Print Color Mapping

Fig. 11 shows additional reference-conditioned color-mapping comparisons on representative scenes from the test set, including cases with viewpoint change and cross-scene references from the same film roll.



Figure 10: Additional qualitative comparisons for negative-film restoration.



Figure 11: Additional qualitative comparisons for reference-conditioned print color mapping.