PROGRESSIVE LIGHT-CONFIDENCE REPRESENTATION LEARNING FOR HIDDEN SCREEN RECONSTRUCTION FROM WALL LIGHT PATTERNS

Anonymous authors
Paper under double-blind review

ABSTRACT

Can an ordinary camera recover what a hidden screen displays by observing only the light it casts onto a wall? The task is governed by two coupled difficulties: (i) screen emission, ambient illumination, secondary transport, wall texture, and sensor noise mix in the wall pattern, so screen-consistent evidence varies across space; (ii) diffuse propagation attenuates high-frequency structure, leaving fine detail weakly represented and increasing reconstruction ambiguity. We introduce Progressive Light-Confidence Representation Learning (PLRL), which addresses these difficulties through ordered restoration and confidence-calibrated detail recovery. Its Physics-Guided Progressive Learning (PPRL) organizes wall evidence into coarse, intermediate, and fine scales and reconstructs layout, structure, and detail through bounded residual updates. Its Light-Propagation Confidence Representation (LPCR) combines scale compatibility, edge state, and image-wide context to identify reliable detail and calibrate its contribution. Across four screen-content domains, PLRL recovers the hidden display's brightness distribution, regional layout, and major contours, achieving the lowest LPIPS on every domain when compared with 24 restoration methods. Experiments spanning illumination, wall response, distance, and viewing changes establish the initial feasibility of recovering principal screen content from wall light captured by an ordinary camera.

1 INTRODUCTION

When a screen's displayed content is hidden from direct view, its light can still reach a nearby wall. Backes et al. (2008) recovered display content from optical reflections on nearby objects. Here, we ask whether spatially mixed illumination on a diffuse wall can reveal the hidden display's principal content to an ordinary camera. In their diffuse-wall experiments, Backes et al. (2009) recovered coarse shapes of individual 10-cm-high letters using a privacy-filtered display and calibrated deconvolution. We extend this wall-based recovery to regional brightness, layout, and major contours across four screen-content domains (Figures 1 and 3), connecting perception under occlusion with visual information exposure.

Diffuse transport mixes light from different screen regions and strongly attenuates fine structure, leaving broad brightness variations as a foundation for recovery. Ambient illumination, secondary reflections, wall texture, and sensor noise also introduce local variation that can resemble screen detail. The two difficulties are linked: recovering the layout provides a structural reference for interpreting weak local evidence, and assessing that evidence determines how strongly it should refine the screen. PLRL follows this connection by coupling progressive structural recovery with confidence-guided detail refinement.

PLRL couples progressive recovery and spatial confidence (Figure 2). PPRL uses bounded residual updates to recover layout and structure before predicting conditioned detail. LPCR combines scale compatibility, edge state, and image-wide context to calibrate that detail. Their combination recovers principal screen regions and contours, achieving the lowest LPIPS on all four domains among the 24 restoration baselines.

The main contributions are:

(1) We demonstrate the initial feasibility of reconstructing hidden screen content from wall light captured by an ordinary camera, recovering regional brightness, layout, and major contours across four screen-content domains.

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.

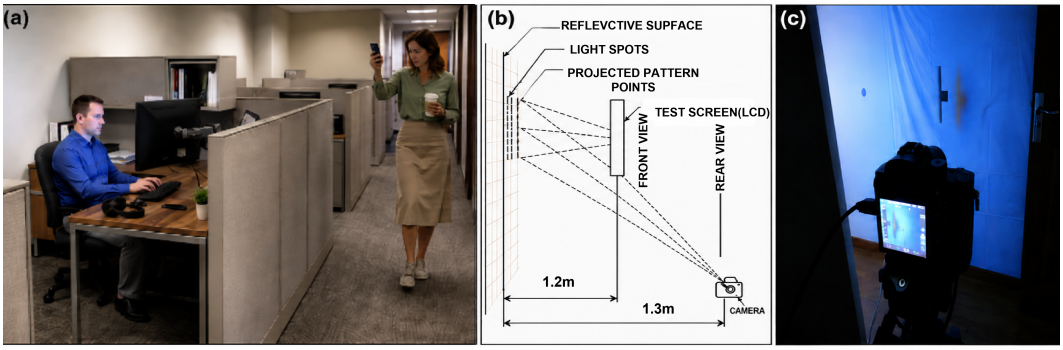


Figure 1: Overview of hidden screen reconstruction from wall light patterns. (a), (b), and (c) show the rendered scene, sensing configuration schematic, and real-world acquisition scene, respectively.

(2) We propose Physics-Guided Progressive Learning (PPRL), which organizes scale-dependent wall evidence into bounded residual updates, enabling coarse-to-fine reconstruction with detail prediction conditioned on recovered structure.
(3) We introduce Light-Propagation Confidence Representation (LPCR), which estimates spatial detail confidence from scale compatibility and edge state together with image-wide context, using this confidence to selectively weight detail corrections.

2 Related Work

Indirect Visual Recovery and Light-Mediated Imaging. Reflection, scattering, and secondary illumination can carry visual information beyond direct view, including display content leaked through environmental reflections Backes et al. (2008). Active NLOS systems use controlled illumination and transient measurements Velten et al. (2012); O'Toole et al. (2018); Lindell et al. (2019); passive approaches use shadows, penumbras, and occlusion boundaries Bouman et al. (2017); Saunders et al. (2019); Yedidia et al. (2019). Saunders et al. (2019) reconstructed two-dimensional hidden scenes with an ordinary digital camera, and Sharma et al. (2021) inferred occupancy and activities from changes in wall illumination. Aittala et al. (2019) jointly estimated hidden video and indirect light transport from temporal variation in observed sequences, typically containing 500–1,500 images. PLRL learns the wall-to-screen mapping from paired images and uses scale-dependent recovery with spatial detail confidence to reconstruct static displayed content from a wall-light image, including document regions and foreground-window placement (Figure 3).

Learning-Based Restoration and Progressive Inverse Reconstruction. CNN-based methods learn restoration mappings for super-resolution and image restoration Zhang et al. (2018a); Zamir et al. (2021); Transformer and hybrid architectures model long-range interactions and adaptive aggregation Wang et al. (2022); Potlapalli et al. (2023). Task-specific methods address diverse degradations Li et al. (2022) and low-light enhancement Cai et al. (2023). Laplacian-pyramid and multi-stage methods organize successive image estimates Lai et al. (2017); Zamir et al. (2021), while curriculum and self-paced learning order training by example difficulty Bengio et al. (2009); Kumar et al. (2010). Multiscale representations and context reasoning provide complementary recovery cues Finder et al. (2024); Dosovitskiy et al. (2021); Wang et al. (2018); Chen et al. (2019). PromptIR encodes degradation with learned prompts Potlapalli et al. (2023), while AdaIR mines and modulates low- and high-frequency features Cui et al. (2025). PLRL couples recovery order with spatial confidence in proposed detail, reaching 0.1678 LPIPS on Screen-2000 compared with 0.3804 for PromptIR and 0.3873 for AdaIR in our wall-to-screen comparison (Table 1).

3 Method

3.1 Overall Architecture

PLRL formulates hidden-screen recovery as a physics-constrained inverse problem, combining a screen structural prior with the transport of emitted light through wall reflection and camera acquisition.
For a color channel c, the registered wall observation is modeled as

$$x_c(r) = \int_{\Omega_w} C_{\xi,c}(r,u) \left(\int_{\Omega_s} T_{\xi,c}(u,\rho) y_c(\rho)\, d\rho \right) du + a_c(r) + n_c(r). \tag{1}$$

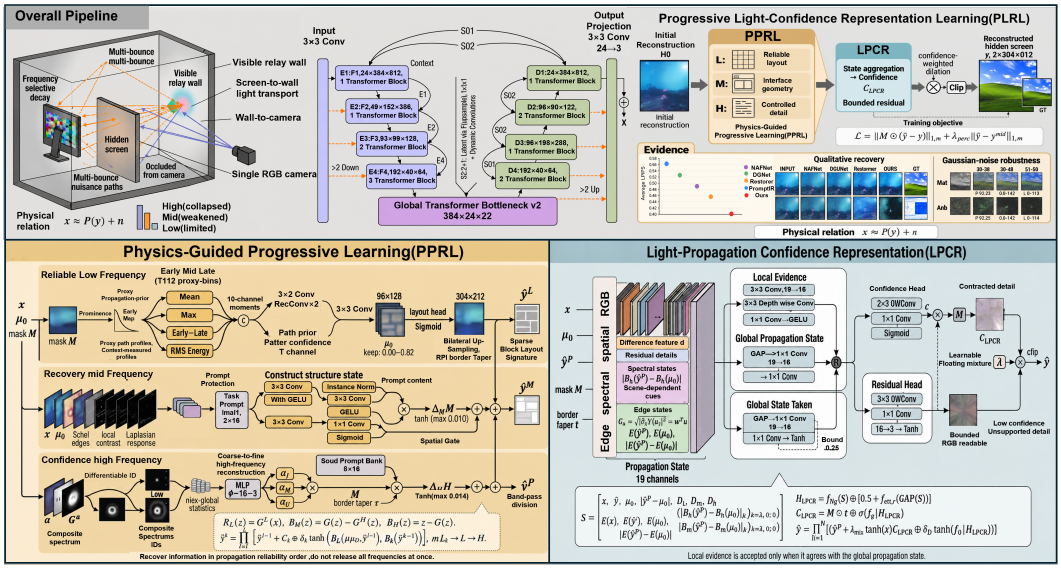


Figure 2: Physical observation model and PLRL data flow from spatially anisotropic wall responses through scale decomposition to confidence-calibrated reconstruction.

where y denotes the hidden screen radiance field, $T_{\xi,c}(u,\rho)$ represents the screen-to-wall transport response, and $C_{\xi,c}(r,u)$ describes the projection and sensor integration of wall-reflected light at camera coordinate r. The latent variable ξ collects the scene-dependent physical factors, including geometry, reflectance properties, and visibility conditions. The terms $a_c(r)$ and $n_c(r)$ represent ambient illumination and sensor noise, respectively.

After discretization, the continuous propagation process can be written as

$$x = A_\xi y + b + n, \qquad A_\xi = U_\xi \Sigma_\xi V_\xi^{\mathsf{T}}, \qquad \Sigma_\xi = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_N). \tag{2}$$

where A_ξ denotes the spatially non-uniform transport operator. The singular values in Σ_ξ characterize the preservation capability of different propagation modes. Small singular values identify strongly attenuated transport modes, making their recovery sensitive to observation noise. This unequal preservation motivates establishing the screen's broad organization before refining weaker structure.

To recover the hidden screen while maintaining consistency with the physical transport process, we formulate reconstruction as a constrained optimization problem:

$$\hat{y} = \underset{z}{\mathrm{argmin}} \left\{ \|A_\xi z - x + b\|_2^2 + \lambda_1 \|\nabla z\|_1 + \lambda_2 \|C(z) - C(y)\|_2^2 \right\}. \tag{3}$$

The first term enforces fidelity with the observed propagation process, the second term preserves spatial structural continuity, and the third term introduces a propagation-aware structural constraint. These constraints motivate the subsequent progressive reconstruction and confidence-guided detail calibration modules.

3.2 Physics-Guided Progressive Learning

Diffuse transport preserves low-frequency layout and illumination more reliably than high-frequency detail, which is more sensitive to diffusion and sensor noise. PPRL follows this scale-dependent reliability through a physically guided, coarse-to-fine recovery trajectory.

The observation is projected into a continuous scale space to separate propagation information by frequency:

$$X_\tau = \sum_{k=1}^{K} \alpha_k^\tau \phi_k, \tag{4}$$

where ϕ_k denotes the scale-space basis and α_k^τ represents the corresponding transport response coefficient at scale τ. Based on this representation, the scale projection operator $\mathcal{M}_\tau(\cdot)$ extracts information with different propagation stability. Specifically, the low-frequency and middle-frequency

components are obtained as $x_l = \mathcal{M}_{\tau_l}(x)$ and $x_m = \mathcal{M}_{\tau_m}(x) - \mathcal{M}_{\tau_l}(x)$, respectively, while the remaining high-frequency component is determined by the complementary relation $x_h = x - x_l - x_m$. To model transport-induced coupling across scales, PPRL describes recovery as the continuous evolution of a hidden state $Z(t) \in \mathcal{H}$:

$$\frac{\partial Z(t)}{\partial t} = \mathcal{A}_\theta(Z(t)) + \mathcal{B}_\theta(X_\tau) + \nabla_Z \Phi(Z(t), X_\tau), \tag{5}$$

where $\mathcal{A}_\theta(\cdot)$ represents the propagation-driven state transition, $\mathcal{B}_\theta(\cdot)$ injects scale-dependent evidence, and $\Phi(\cdot)$ denotes the recovery potential induced by the physical prior. This formulation describes recovery as a gradual transition from broad layout to finer structure and detail. The three learned updates in Equations equation 7–equation 9 organize reconstruction into these successive stages.

To establish interactions between spatial positions under different propagation conditions, we introduce a cross-scale propagation operator. For two input states, the operator is defined as

$$\Theta_a(U, V)(r) = \int_\Omega \frac{\exp[-\chi(U(r), V(v))]}{\int_\Omega \exp[-\chi(U(r), V(u))] du} V(v)\, dv. \tag{6}$$

where $\chi(\cdot, \cdot)$ measures the propagation-state discrepancy between different locations. This operator aggregates spatially consistent evidence and suppresses unreliable responses caused by severe transport attenuation.

In the first stage, stable low-frequency propagation information is utilized to recover the global screen layout, while additional scale evidence is introduced to correct spatial distortion:

$$l(r) = \Pi\left[x_l(r) + \alpha \Theta_l\left(x_l(r), x_m(r), x_h(r), \Theta_a(x_l, x_m, x_h)\right)\right]. \tag{7}$$

Here, $\Pi(\cdot)$ clips the output to the valid image range. In the bounded formulation of Equation equation A.11, the correction operators include $\Theta_l = \tanh r_l$, $\Theta_s = \tanh r_s$, and $\Theta_h = \tanh r_d$, with r_l, r_s, r_d denoting the learned correction fields.

After obtaining a stable layout representation, the second stage further restores mid-frequency structural information. The structure recovery process is written as

$$s(r) = \Pi\left[l(r) + \beta \Theta_s\left(l(r), x_m(r), \Theta_a(l, x_m)\right)\right]. \tag{8}$$

High-frequency components are the most vulnerable to transport degradation. PPRL predicts their residual from the recovered structure together with the high-frequency wall evidence, so the detail proposal refines an already organized screen estimate:

$$d(r) = \gamma \Theta_h\left(s(r), x_h(r), \Theta_a(s, x_h)\right). \tag{9}$$

The preliminary reconstruction is $p(r) = \Pi(s(r) + d(r))$, combining the recovered structure with its conditioned detail residual.

We characterize changes between successive image estimates using $Z_0 = x_l$, $Z_1 = l$, $Z_2 = s$, and $Z_3 = p$:

$$\mathcal{R}_{\text{prog}} = \sum_{i=1}^{3} \eta_i \|Z_i - Z_{i-1}\|_\infty, \tag{10}$$

where $\eta_i \geq 0$ weight the stage changes. Equation equation A.12 bounds each update by α, β, or γ, keeping later refinements connected to the established layout and structure. Progressive supervision is specified separately by $\mathcal{L}_{\text{prog}}$ in Equation equation A.14.

3.3 Light-Propagation Confidence Representation

PPRL's recovered details remain spatially uncertain: attenuated regions may mix useful structure with unreliable residuals. LPCR estimates spatial confidence to calibrate these details according to their propagation reliability.

Given the intermediate restoration states generated by PPRL, the local propagation state is constructed by jointly considering the observation, progressive restoration results, scale consistency, and structural variations:

$$\zeta(r) = [x(r), l(r), s(r), p(r), D_l(r), D_m(r), D_h(r), \nabla s(r), \nabla p(r)]. \tag{11}$$

where $D_l(r)$, $D_m(r)$, and $D_h(r)$ denote the consistency errors of different restoration scales. These terms describe the reliability variation between different progressive stages and provide additional information for evaluating local propagation uncertainty.

To measure the relationship between different spatial positions under the propagation state space, the state difference is defined as $\Delta_\zeta(r, v) = \zeta(r) - \zeta(v)$. Based on this difference, LPCR constructs a propagation-aware spatial probability kernel:

$$\Gamma_\zeta(r, v) = \frac{\exp\left[-\frac{1}{2}\Delta_\zeta^{\mathsf{T}}(r, v)\Sigma_\zeta^{-1}\Delta_\zeta(r, v) - \mu\|r - v\|_2^2\right]}{\int_\Omega \exp\left[-\frac{1}{2}\Delta_\zeta^{\mathsf{T}}(r, u)\Sigma_\zeta^{-1}\Delta_\zeta(r, u) - \mu\|r - u\|_2^2\right] du}. \tag{12}$$

where Σ_ζ represents the uncertainty covariance of the propagation state and μ controls the spatial locality constraint. The kernel evaluates spatial relationships through the propagation state, allowing contextual evidence to guide confidence in each proposed detail.

Based on the propagation kernel, LPCR aggregates non-local contextual evidence through

$$\mathcal{P}_{\text{conf}}(r) = \int_\Omega \Gamma_\zeta(r, v)\Phi(\zeta(v), \bar{\zeta})\, dv, \tag{13}$$

where $\Phi(\cdot)$ extracts propagation-related contextual information and $\bar{\zeta}$ denotes the global propagation state obtained from the spatial statistics of ζ.

The confidence value expresses learned support for a proposed screen detail. It combines the wall observation with the current reconstruction to assess how strongly that detail should contribute to the recovered image. An ideal score decreases with structural error. Let the local error be $e(r) = \|\nabla\hat{y}(r) - \nabla y(r)\|_2$, then the ideal confidence target is defined as $q^\star(r) = \exp[-\kappa e(r)]$.

Paired screen targets define the ideal reliability signal. Conditional on the observation, the ideal confidence field is characterized by the risk objective:

$$q^{\text{opt}} = \underset{q}{\operatorname{argmin}}\, \mathbb{E}_{\xi\sim p(\xi|x)}\left[\rho\left(q(r) - e^{-\kappa e(r)}\right)\right] + \lambda_q \int_\Omega \|\nabla q(r)\|_2^2 dr. \tag{14}$$

where $p(\xi|x)$ represents the distribution of possible propagation states conditioned on the observation. The first term aligns the estimated confidence with the latent reliability target, while the second term enforces spatial smoothness of the confidence field.

At inference, a learned feedforward readout predicts confidence from the local propagation state, contextual evidence, and global transport prior:

$$q(r) = \sigma\left(\Theta_q(\zeta(r), \mathcal{P}_{\text{conf}}(r), \bar{\zeta})\right), \tag{15}$$

where $\sigma(\cdot)$ constrains the confidence value into the range $[0, 1]$.

3.3.1 Confidence-Guided Residual Calibration

To avoid amplifying unreliable PPRL residuals, LPCR calibrates detail by minimizing a propagation-constrained energy:

$$\begin{aligned}\hat{d} = \underset{\delta}{\operatorname{argmin}} \quad & \frac{1}{2}\int_\Omega \omega_q(r)\|\delta(r) - d(r)\|_2^2 dr \\ & + \frac{\lambda_d}{2}\int_\Omega\int_\Omega \Gamma_\zeta(r, v)\|\delta(r) - \delta(v)\|_2^2 dr\, dv \\ & + \lambda_s \int_\Omega \Psi(\nabla\delta(r), \nabla p(r), \zeta(r)) dr.\end{aligned} \tag{16}$$

where $\omega_q(r) = (1 - q(r))/(q(r) + \epsilon)$ adaptively adjusts the residual constraint according to the local propagation confidence. Reliable regions with high confidence preserve informative details, while uncertain regions receive stronger regularization to suppress unreliable responses.

Finally, the hidden-screen reconstruction is obtained by combining the confidence-calibrated residual with the recovered structural component:

$$\hat{y}(r) = \Pi(s(r) + \hat{d}(r)). \tag{17}$$

4 Experiments and Results

4.1 Experimental Setup

We evaluate hidden-screen reconstruction on four paired datasets: Screen-Figure, Screen-2000, Screen-Game, and OS-Lock. Each pair contains a wall-light observation and its screen target on a 512×384 RGB canvas. We train a separate model for each domain and evaluate PSNR, SSIM, and LPIPS within the valid target region. Appendices A.3–A.4 describe the data and training settings.

4.2 Comparison with State-of-the-Art

Table 1 compares PLRL with 24 external restoration methods. PLRL achieves the highest PSNR and SSIM and the lowest LPIPS across all four screen-content domains, ranking first in all 12 dataset-metric comparisons and supporting recovery of hidden-screen visual organization from diffuse wall patterns.

Table 1: Comparison with 24 external restoration methods on Screen-Figure, Screen-2000, Screen-Game, and OS-Lock.

Method	Source	Screen-Figure-1250			Screen-2000			Screen-Game-1250			OS-Lock-1399		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
CNN-Based General Image Restoration													
UNet	MICCAI'15	11.815	0.5658	0.6675	9.184	0.2401	0.8061	10.285	0.4715	0.7581	12.180	0.4874	0.7618
DnCNN	TIP'17	14.204	0.5738	0.6639	8.883	0.3466	0.7245	8.738	0.4275	0.7258	11.818	0.4787	0.7221
NAFNet	ECCV'22	17.729	0.6477	0.6216	21.617	0.7585	0.3794	18.160	0.6587	0.5759	15.242	0.5619	0.6637
DGUNet	CVPR'22	18.038	0.6952	0.6229	18.962	0.6960	0.4818	17.888	0.6596	0.5950	16.547	0.6236	0.6514
DExNet	TPAMI'25	18.602	0.7208	0.6214	20.005	0.7323	0.4382	17.629	0.6586	0.5954	15.518	0.6084	0.6520
ConvIR	TPAMI'24	18.540	0.7151	0.6293	20.271	0.7283	0.4495	17.642	0.6515	0.6013	14.911	0.5925	0.6867
IRNeXt	ICML'23	17.479	0.6679	0.6326	19.956	0.7234	0.4542	18.136	0.6681	0.5735	16.134	0.6082	0.6785
DeepSN-Net	TPAMI'25	14.401	0.4993	0.6920	10.489	0.2954	0.7787	10.868	0.4440	0.7390	12.836	0.4228	0.7711
MIRNetv2	TPAMI'22	18.340	0.7133	0.6536	18.237	0.6917	0.4923	17.679	0.6559	0.6120	15.211	0.6042	0.6575
OKNet	AAAI'24	19.245	0.7211	0.5964	20.584	0.7384	0.4445	17.892	0.6625	0.5756	14.664	0.5553	0.6781
Transformer/Hybrid General Image Restoration													
Restormer	CVPR'22	18.921	0.7096	0.5508	22.601	0.7774	0.3449	17.368	0.6438	0.5676	16.754	0.6071	0.6329
Uformer	CVPR'22	17.586	0.6187	0.6260	19.204	0.7068	0.4639	17.547	0.6470	0.6142	15.149	0.5652	0.6743
SFHformer	ECCV'24	13.205	0.5970	0.7134	8.817	0.3940	0.8066	10.010	0.4620	0.7598	12.461	0.5095	0.7404
PromptIR	NeurIPS'23	18.757	0.6913	0.6198	21.564	0.7552	0.3804	17.784	0.6599	0.5708	15.342	0.5936	0.6374
AdaIR	ICLR'25	17.930	0.6690	0.6228	21.714	0.7621	0.3873	17.436	0.6394	0.5754	15.168	0.5729	0.6455
AST	CVPR'24	17.891	0.6386	0.6324	19.391	0.7050	0.4668	17.456	0.6453	0.6158	15.072	0.5366	0.6920
Task-Specific and All-in-One Restoration													
OneRestore	ECCV'24	12.529	0.5863	0.7044	10.409	0.4642	0.7142	10.313	0.5062	0.7474	12.237	0.5221	0.7212
DocRes	CVPR'24	18.772	0.7184	0.6108	21.145	0.7490	0.3999	17.528	0.6472	0.5795	16.357	0.6094	0.6467
DarkIR	CVPR'25	4.185	0.0384	0.9290	9.221	0.1749	0.8673	4.465	0.0409	0.9361	8.787	0.0949	0.8040
Retinexformer	ICCV'23	17.895	0.6905	0.6186	19.567	0.6732	0.4554	17.925	0.6556	0.5924	14.600	0.5586	0.6506
DehazeFormer	TIP'23	18.917	0.7086	0.6157	19.366	0.6931	0.4794	17.774	0.6494	0.6074	14.873	0.5375	0.6687
HVI-CIDNet	CVPR'25	19.415	0.7210	0.5965	19.692	0.7178	0.4724	17.774	0.6555	0.5979	15.606	0.5758	0.6714
DEA-Net	TIP'24	18.179	0.7145	0.6425	20.651	0.7500	0.4240	17.729	0.6579	0.5879	15.962	0.6069	0.6742
Pixel2Pixel	TPAMI'25	15.476	0.6145	0.6379	10.324	0.4816	0.6637	11.140	0.4502	0.6917	12.405	0.5178	0.7303
PLRL	**Ours**	20.336	0.7457	0.4712	27.341	0.8682	0.1678	18.378	0.6743	0.4736	17.088	0.6297	0.6273

Figure 3 compares PLRL with 11 external methods on nine Screen-2000 examples using identical observations and targets. Across interfaces, documents, desktops, and dialogs, PLRL recovers the major screen extent, window positions, and global color organization. Appendix perturbation and fixed-backbone comparisons extend this evidence across the tested conditions.

4.3 Ablation Study

Table 2 shows complementary gains from establishing screen structure and controlling its refinement. Full PLRL ranks first in all 12 metric-dataset combinations. PPRL raises OS-Lock PSNR from 15.123 to 17.088 dB and reduces Screen-Figure LPIPS from 0.620 to 0.471. LPCR raises OS-Home PSNR from 13.409 to 14.950 dB and Screen-Game SSIM from 0.596 to 0.674. Figure 4 locates these gains in the arrangement of major screen regions and their boundaries.

LapSRN reconstructs residual subbands progressively, and MPRNet develops restoration through successive stages Lai et al. (2017); Zamir et al. (2021). In our PPRL replacement comparison, the task-specific recovery order reaches 0.6743 SSIM on Screen-Game, compared with 0.6037 and 0.6043 for the LapSRN- and MPRNet-based replacements (Table 2). The LPCR replacement comparison uses SRMD and PromptIR Zhang et al. (2018a); Potlapalli et al. (2023).

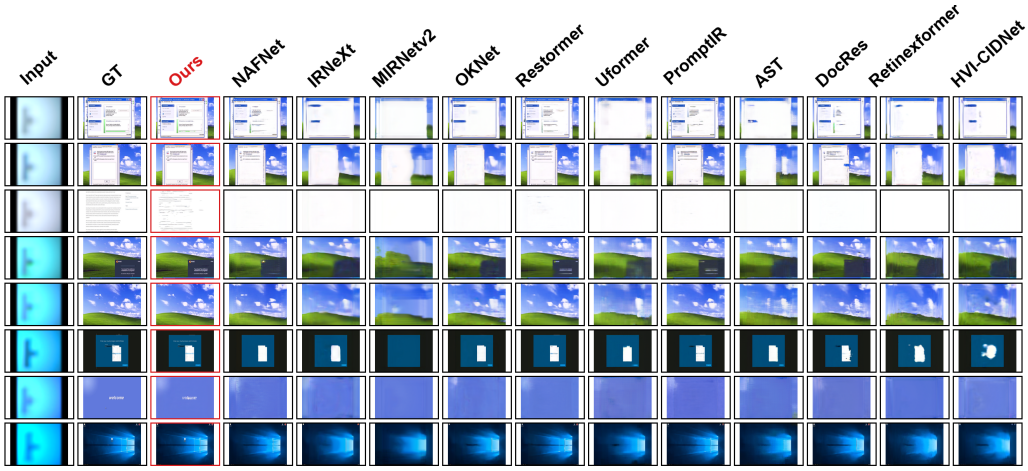


Figure 3: Nine Screen-2000 examples comparing the wall-light observation, target, PLRL, and 11 external restoration methods; the PLRL column is enclosed by a red border.

Table 2: Joint ablation of PPRL and LPCR on OS-Lock, OS-Home, Screen-Figure, and Screen-Game with the same training setup.

Method	Model	OS-Lock			OS-Home			Screen-Figure			Screen-Game		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
PPRL Module Ablations													
PPRL ablation	—	15.123	0.5633	0.6688	13.383	0.5149	**0.7371**	**18.852**	0.7004	**0.6197**	16.693	0.5956	0.6347
PPRL → LPR	LapSRN	14.944	0.5561	0.6638	13.595	0.5194	0.7392	17.927	0.7031	0.6472	16.771	0.6037	0.6318
PPRL → MSPR	MPRNet	**15.650**	0.5700	0.6666	**13.882**	0.5248	0.7499	18.046	**0.7072**	0.6419	16.720	0.6043	**0.6294**
LPCR Module Ablations													
LPCR ablation	—	14.939	0.5475	0.6731	13.409	0.5094	0.7532	18.677	0.6962	0.6289	16.708	0.5960	0.6377
LPCR → DMC	SRMD	15.197	**0.5767**	0.6689	13.363	0.5207	0.7509	18.190	0.7057	0.6334	**17.024**	**0.6054**	0.6308
LPCR → DPC	PromptIR	15.375	0.5484	0.6756	13.777	0.5284	0.7462	18.497	0.6969	0.6368	16.736	0.5991	0.6392
Ours	**PLRL**	**17.088**	**0.6217**	**0.6273**	**14.950**	**0.5493**	**0.7002**	**19.336**	**0.7094**	**0.4712**	**18.378**	**0.6743**	**0.4736**

Figure 4 visualizes the configurations evaluated in Table 2, together with an additional joint PPRL+LPCR ablation condition. The red-bordered Ours column is placed immediately after the ground truth for direct comparison. Across the representative hidden-screen examples, PLRL preserves the target layout, major color regions, and visible interface structure; the corresponding variants show stronger blur, color drift, or incomplete structure. These visual comparisons align with the metric gains in Table 2.

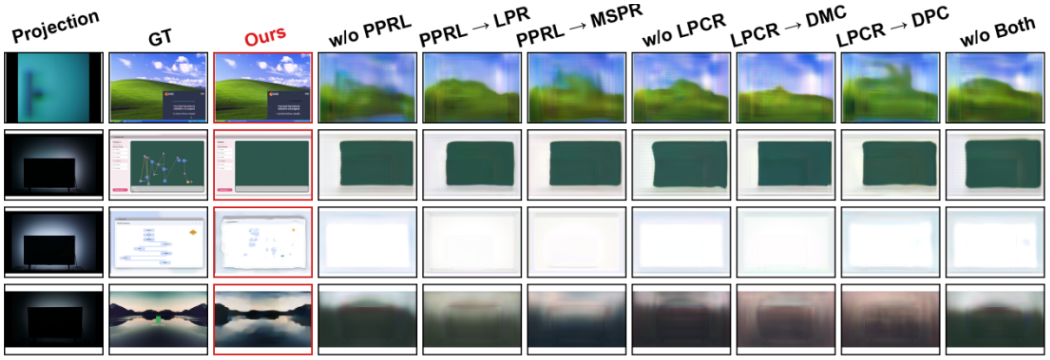


Figure 4: Qualitative visualization of the joint PPRL and LPCR ablation. The columns show the wall-light observation, ground truth, PLRL, and the corresponding ablation variants.

Table 3 shows that the complete LPCR design ranks first in all 12 metric-dataset combinations. Adding the full propagation state lifts Screen-Game PSNR from 16.674 dB to 18.378 dB. WTConv extracts multiband features through wavelet convolutions, while ViT represents images as patch sequences Finder et al. (2024); Dosovitskiy et al. (2021). In our propagation-state comparison, LPCR reaches 0.6743 SSIM on Screen-Game versus 0.5903 for the WTConv-based alternative, 0.5910 for the ViT-based alternative, and 0.5902 for second-order covariance, with higher PSNR and lower LPIPS across the four domains. Non-local networks aggregate information across positions, while

graph-based global reasoning models interactions in a learned reasoning space Wang et al. (2018); Chen et al. (2019). LPCR reaches 0.6217 SSIM on OS-Lock, compared with 0.5620 for non-local consistency and 0.5618 for MPN-COV, and lowers the graph alternative's OS-Home LPIPS from 0.7506 to 0.7002. LPCR leads the channel, non-local, and graph context comparisons, supporting context that assesses proposed detail through wall evidence and the current reconstruction together. The table also includes covariance (MPN-COV) and channel-recalibration (SENet) alternatives Li et al. (2017; 2018); Hu et al. (2018).

Table 3: Internal LPCR ablation on OS-Lock, OS-Home, Screen-Figure, and Screen-Game with the same training setup.

Method	Model	OS-Lock			OS-Home			Screen-Figure			Screen-Game		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Propagation State Representation													
Propagation-state ablation	—	14.943	0.5602	0.6720	**13.885**	**0.5273**	**0.7368**	**18.669**	0.6957	**0.6323**	**16.674**	**0.6053**	**0.6366**
Wavelet subbands	WTConv	**15.346**	0.5544	**0.6676**	13.568	0.5150	0.7485	18.021	**0.7132**	0.6657	16.603	0.5903	0.6401
Patch tokens	ViT	15.085	0.5598	0.6697	13.666	0.5258	0.7547	18.015	0.7099	0.6645	16.564	0.5910	0.6408
Second-order covariance	MPN-COV	14.701	0.5618	0.6712	13.781	0.5272	0.7498	18.124	0.7057	0.6707	16.624	0.5902	0.6434
Image-Wide Context													
Image-wide-context ablation	—	14.769	0.5569	0.6718	13.685	0.5203	0.7499	18.603	0.6974	0.6352	16.564	0.5984	0.6404
Channel recalibration	SENet	14.814	0.5483	0.6738	12.542	0.4898	0.7530	18.024	0.7062	0.6371	16.579	0.5736	0.6452
Non-local consistency	Non-local Net	15.023	**0.5620**	0.6775	13.723	0.5206	0.7565	18.427	0.6957	0.6339	16.535	0.5851	0.6384
Graph reasoning	GloRe	14.353	0.5481	**0.6676**	13.591	0.5182	0.7506	18.200	0.7092	0.6517	15.957	0.5855	0.6485
Ours	**PLRL**	**17.088**	**0.6217**	**0.6273**	**14.950**	**0.5493**	**0.7002**	**19.336**	**0.7147**	**0.4712**	**18.378**	**0.6743**	**0.4736**

4.4 Robustness under Brightness Changes

Table 4 shows that PLRL leads across the tested display-brightness range from -100 to $+100$ nit relative to the 300-nit nominal condition. PLRL ranks first in all 15 condition-metric comparisons. At the nominal 300-nit screen brightness (zero shift), PLRL obtains 27.341 dB PSNR, 0.868 SSIM, and 0.168 LPIPS, compared with the strongest competing values of 21.614 dB, 0.758, and 0.379 from NAFNet. At the -100-nit condition, PLRL obtains 10.019 dB, 0.458, and 0.666, maintaining the leading measurements in this suite.

Table 4: Screen-2000 robustness under display-brightness shifts of -100, -50, 0, $+50$, and $+100$ nit relative to the 300-nit nominal condition.

Method	Source	-100 **nit**			-50 **nit**			**0 nit**			$+50$ **nit**			$+100$ **nit**		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	9.564	0.445	0.761	10.336	0.458	0.741	10.409	0.464	0.714	10.134	0.466	0.706	9.678	0.464	0.702
DocRes	CVPR'24	9.237	0.409	0.766	13.778	0.499	0.615	21.138	0.748	0.400	16.266	0.620	0.527	13.060	0.543	0.628
Pixel2Pixel	TPAMI'25	**9.962**	0.439	**0.704**	10.336	0.469	0.674	10.325	0.482	0.664	10.206	0.486	0.663	10.042	0.484	0.671
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	9.216	0.408	0.780	12.933	0.472	0.645	**21.614**	**0.758**	**0.379**	14.737	0.566	0.567	11.730	0.505	0.660
DGUNet	CVPR'22	9.228	0.423	0.770	12.472	0.508	0.669	18.961	0.695	0.482	13.640	0.574	0.606	11.052	0.532	0.678
Uformer	CVPR'22	8.969	0.406	0.784	13.524	**0.552**	0.626	19.201	0.706	0.464	15.023	0.633	0.545	12.820	**0.564**	0.616
MIRNetv2	TPAMI'22	8.828	0.414	0.760	12.039	0.531	0.656	18.235	0.691	0.492	13.372	0.577	0.600	11.139	0.540	0.660
OKNet	AAAI'24	9.530	0.429	0.777	12.913	0.521	0.659	20.582	0.738	0.445	13.776	0.597	0.615	10.733	0.522	0.703
AST	CVPR'24	9.327	0.427	0.793	13.120	0.527	0.660	19.389	0.704	0.467	14.829	0.621	0.560	12.506	0.554	0.635
ConvIR	TPAMI'24	8.861	0.410	0.775	11.413	0.500	0.679	20.266	0.728	0.449	14.363	0.593	0.586	10.934	0.522	0.686
DnCNN	TIP'17	8.461	0.331	0.743	8.859	0.353	0.726	8.881	0.346	0.724	8.612	0.332	0.740	7.914	0.302	0.764
UNet	MICCAI'15	8.583	0.360	0.805	8.918	0.289	0.808	9.184	0.240	0.806	8.981	0.236	0.787	8.564	0.247	0.774
Models for Related Degradations																
DExNet	TPAMI'25	9.245	0.410	0.755	13.139	0.513	0.636	20.018	0.731	0.438	**16.580**	0.640	**0.520**	**13.345**	0.548	**0.609**
HVI-CIDNet	CVPR'25	9.201	0.410	0.773	10.671	0.452	0.709	19.690	0.717	0.472	13.606	0.558	0.622	11.559	0.505	0.686
Retinexformer	ICCV'23	9.639	0.394	0.743	13.133	0.467	0.643	19.563	0.672	0.455	14.215	0.562	0.584	11.606	0.504	0.674
DehazeFormer	TIP'23	8.972	0.404	0.784	12.034	0.493	0.685	19.357	0.692	0.479	13.581	0.566	0.623	11.124	0.526	0.688
DEA-Net	TIP'24	8.935	0.413	0.786	**13.831**	0.470	0.635	20.645	0.749	0.424	16.528	**0.651**	0.537	13.159	0.528	0.639
IRNeXt	ICML'23	9.154	0.431	0.768	11.943	0.517	0.659	19.952	0.723	0.454	15.009	0.615	0.569	11.587	0.534	0.665
SFHformer	ECCV'24	9.679	0.399	0.808	9.438	0.437	0.801	8.817	0.394	0.807	8.228	0.353	0.821	7.865	0.315	0.830
DarkIR	CVPR'25	8.432	0.189	0.866	8.914	0.184	0.869	9.221	0.175	0.867	9.358	0.166	0.859	9.373	0.158	0.850
PLRL	**Ours**	**10.019**	**0.458**	**0.666**	**15.312**	**0.570**	**0.553**	**27.341**	**0.868**	**0.168**	**20.272**	**0.712**	**0.343**	**14.254**	**0.566**	**0.568**

Figure 5 plots the same brightness sweep as Table 4, with absolute screen brightness values of 200, 250, 300, 350, and 400 nit corresponding to the five relative shifts. PLRL reaches its strongest values near the nominal 300-nit condition and leads the comparison toward both endpoints. Its PSNR, SSIM, and LPIPS curves preserve the favorable direction at every displayed brightness, showing PLRL's leading measurements across the tested luminance changes.

Figure 5: Screen-2000 PSNR, SSIM, and LPIPS curves across absolute screen brightness from 200 to 400 nit; 300 nit is the nominal condition.

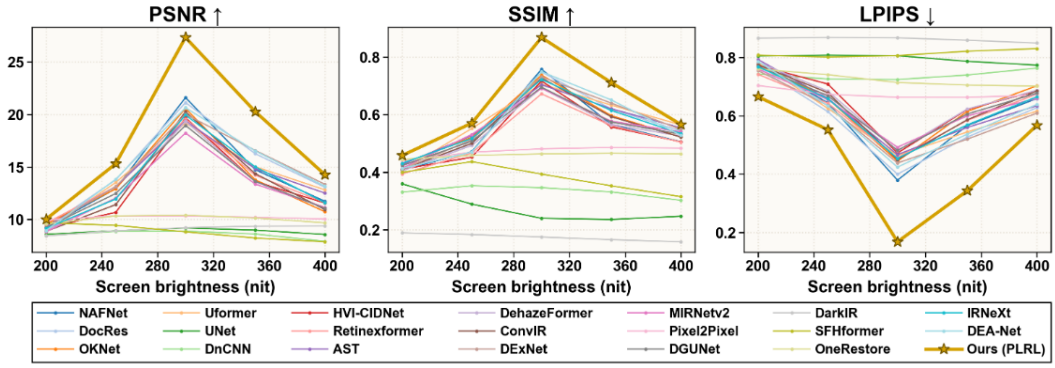
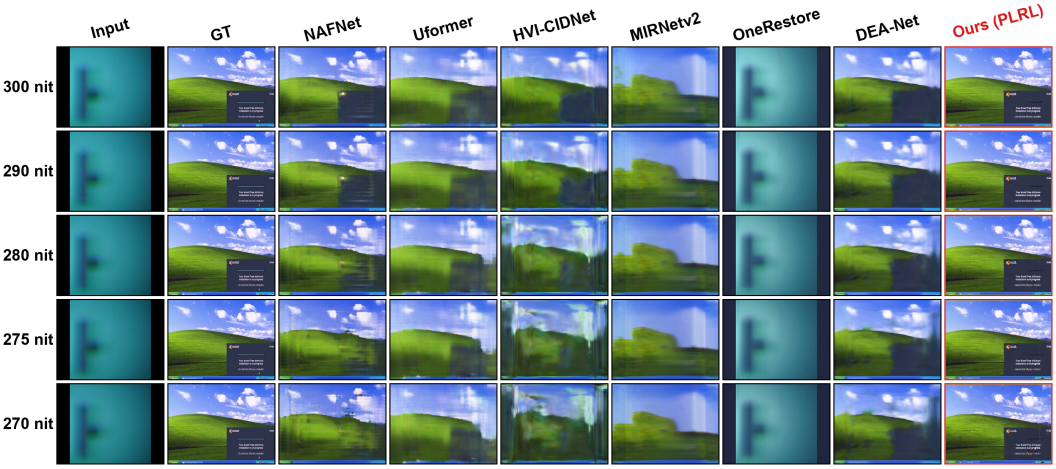


Figure 6 shows representative Screen-2000 reconstructions at 300, 290, 280, 275, and 270 nit. PLRL preserves the dominant desktop layout and large window structure across this local brightness range, while the displayed baselines show stronger blur, color mixing, or layout distortion. The preserved desktop and window arrangement makes the recovered spatial organization visible over 270–300 nit, complementing the broader 200–400-nit comparison in Figure 5.

Figure 6: Representative Screen-2000 reconstructions under local brightness changes. The columns show the input, ground truth, six restoration baselines, and PLRL (Ours), with the PLRL (Ours) column outlined in red.



5 Conclusion

We establish the feasibility of recovering hidden screen content from wall light captured by an ordinary camera while the screen remains outside its view. PLRL organizes recovery from layout to structure and detail, then uses learned spatial confidence to determine how strongly proposed details refine that estimate. Across four screen-content domains, it ranks first in all 12 dataset-metric comparisons against 24 restoration methods, with the highest PSNR and SSIM and the lowest LPIPS on each domain. The reconstructed images reveal brightness distribution, regional layout, and major screen contours. Component ablations connect these gains to the complementary roles of progressive recovery and confidence-guided refinement. Experiments spanning brightness, wall response, sensor noise, camera distance, and viewing direction extend the evidence across the evaluated observation conditions. These results establish a learned route from diffuse wall illumination to the principal visual organization of a hidden display, with an explicit recovery order and spatial control over the contribution of proposed detail.

REFERENCES

Miika Aittala, Prafull Sharma, Lukas Murmann, Adam Yedidia, Gregory Wornell, Bill Freeman, and Fredo Durand. Computational mirrors: Blind inverse light transport by deep matrix factorization. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/5a2afca61e35f45a7dd44ca46e0225f4-Paper.pdf.

Michael Backes, Markus Dürmuth, and Dominique Unruh. Compromising reflections, or how to read LCD monitors around the corner. In *IEEE Symposium on Security and Privacy*, pp. 158–169, 2008. doi: 10.1109/SP.2008.25.

Michael Backes, Tongbo Chen, Markus Duermuth, Hendrik P.A. Lensch, and Martin Welk. Tempest in a Teapot: Compromising Reflections Revisited. In *2009 30th IEEE Symposium on Security and Privacy*, pp. 315–327, 2009. doi: 10.1109/sp.2009.20.

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the International Conference on Machine Learning*, pp. 41–48, 2009. doi: 10.1145/1553374.1553380.

Katherine L. Bouman, Vickie Ye, Adam B. Yedidia, Frédo Durand, Gregory W. Wornell, Antonio Torralba, and William T. Freeman. Turning corners into cameras: Principles and methods. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2289–2297, 2017. doi: 10.1109/ICCV.2017.249.

Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12504–12513, 2023. doi: 10.1109/ICCV51070.2023.01149.

Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple Baselines for Image Restoration. In *Computer Vision – ECCV 2022*, pp. 17–33, 2022. doi: 10.1007/978-3-031-20071-7_2.

Yunpeng Chen, Marcus Rohrbach, Zhicheng Yan, Shuicheng Yan, Jiashi Feng, and Yannis Kalantidis. Graph-based global reasoning networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 433–442, 2019. doi: 10.1109/CVPR.2019.00052.

Zixuan Chen, Zewei He, and Zhe-Ming Lu. DEA-Net: Single Image Dehazing Based on Detail-Enhanced Convolution and Content-Guided Attention. *IEEE Transactions on Image Processing*, 33, pp. 1002–1015, 2024. doi: 10.1109/tip.2024.3354108.

Yuning Cui, Wenqi Ren, Sining Yang, Xiaochun Cao, and Alois Knoll. IRNeXt: Rethinking convolutional network design for image restoration. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 6545–6564. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/cui23d.html.

Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Revitalizing Convolutional Network for Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 9423–9438, 2024a. doi: 10.1109/tpami.2024.3419007.

Yuning Cui, Wenqi Ren, and Alois Knoll. Omni-Kernel Network for Image Restoration. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(2), pp. 1426–1434, 2024b. doi: 10.1609/aaai.v38i2.27907.

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. URL https://openreview.net/forum?id=M5t0WvjfCg.

Xin Deng, Chenxiao Zhang, Lai Jiang, Jingyuan Xia, and Mai Xu. DeepSN-Net: Deep Semi-Smooth Newton Driven Network for Blind Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(4):2632–2646, 2025. doi: 10.1109/tpami.2024.3525089.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=YicbFdNTTy.

Daniel Feijoo, Juan C. Benito, Alvaro Garcia, and Marcos V. Conde. DarkIR: Robust Low-Light Image Restoration. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10879–10889, 2025. doi: 10.1109/cvpr52734.2025.01016.

Shahaf E. Finder, Roy Amoyal, Eran Treister, and Oren Freifeld. Wavelet convolutions for large receptive fields. In *European Conference on Computer Vision*, pp. 363–380, 2024. doi: 10.1007/978-3-031-72949-2_21.

Spyros Gidaris, Praveer Singh, and Nikos Komodakis. Unsupervised representation learning by predicting image rotations. In *International Conference on Learning Representations*, 2018. URL https://arxiv.org/abs/1803.07728.

Yu Guo, Yuan Gao, Yuxu Lu, Huilin Zhu, Ryan Wen Liu, and Shengfeng He. OneRestore: A Universal Restoration Framework for Composite Degradation. In *Computer Vision – ECCV 2024*, pp. 255–272, 2024. doi: 10.1007/978-3-031-72655-2_15.

G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006. doi: 10.1126/science.1127647.

Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141, 2018. doi: 10.1109/CVPR.2018.00745.

Jun-Jie Huang, Tianrui Liu, Zihan Chen, Xinwang Liu, Meng Wang, and Pier Luigi Dragotti. A lightweight deep exclusion unfolding network for single image reflection removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(6):4957–4973, 2025. doi: 10.1109/TPAMI.2025.3548148.

Q. Huynh-Thu and M. Ghanbari. Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44(13):800–801, 2008. doi: 10.1049/el:20080522.

Xingyu Jiang, Xiuhui Zhang, Ning Gao, and Yue Deng. When fast fourier transform meets transformer for image restoration. In *European Conference on Computer Vision*, pp. 381–402, 2024. doi: 10.1007/978-3-031-72995-9_22.

M. Pawan Kumar, Benjamin Packer, and Daphne Koller. Self-paced learning for latent variable models. In *Advances in Neural Information Processing Systems*, volume 23, pp. 1189–1197, 2010.

Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5835–5843, 2017. doi: 10.1109/CVPR.2017.618.

Chen-Yu Lee, Saining Xie, Patrick Gallagher, Zhengyou Zhang, and Zhuowen Tu. Deeply-supervised nets. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 38, pp. 562–570, 2015. URL https://proceedings.mlr.press/v38/lee15a.html.

Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17452–17462, 2022. URL https://openaccess.thecvf.com/content/CVPR2022/html/Li_All-in-One_Image_Restoration_for_Unknown_Corruption_CVPR_2022_paper.html.

Peihua Li, Jiangtao Xie, Qilong Wang, and Zilin Gao. Towards faster training of global covariance pooling networks by iterative matrix square root normalization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 947–955, 2018. doi: 10.1109/CVPR.2018.00105.

Peihua Li, Jiangtao Xie, Qilong Wang, and Wangmeng Zuo. Is second-order information helpful for large-scale visual recognition? In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2070–2078, 2017. doi: 10.1109/ICCV.2017.228. URL https://openaccess.thecvf.com/content_ICCV_2017/html/Li_Is_Second-Order_Information_ICCV_2017_paper.html.

David B. Lindell, Gordon Wetzstein, and Matthew O'Toole. Wave-based non-line-of-sight imaging using fast f-k migration. *ACM Transactions on Graphics*, 38(4):116, 2019. doi: 10.1145/3306346.3322937. URL https://www.computationalimaging.org/publications/nlos-fk/.

Yidi Liu, Dong Li, Xueyang Fu, Xin Lu, Jie Huang, and Zheng-Jun Zha. UHD-processer: Unified UHD image restoration with progressive frequency learning and degradation-aware prompts. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23121–23130, 2025. doi: 10.1109/CVPR52734.2025.02153.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Bkg6RiCqY7.

Qing Ma, Junjun Jiang, Xiong Zhou, Pengwei Liang, Xianming Liu, and Jiayi Ma. Pixel2Pixel: A Pixelwise Approach for Zero-Shot Single Image Denoising. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(6):4614–4629, 2025. doi: 10.1109/tpami.2025.3546870.

Chong Mou, Qian Wang, and Jian Zhang. Deep Generalized Unfolding Networks for Image Restoration. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17378–17389, 2022. doi: 10.1109/cvpr52688.2022.01688.

Mehdi Noroozi and Paolo Favaro. Unsupervised learning of visual representations by solving jigsaw puzzles. In *European Conference on Computer Vision*, pp. 69–84, 2016. doi: 10.1007/978-3-319-46466-4_5.

Matthew O'Toole, David B. Lindell, and Gordon Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338–341, 2018. doi: 10.1038/nature25489.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman Khan, and Fahad Shahbaz Khan. PromptIR: Prompting for all-in-one image restoration. In *Advances in Neural Information Processing Systems*, volume 36, pp. 71275–71293, 2023. doi: 10.52202/075280-3121.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241, 2015. doi: 10.1007/978-3-319-24574-4_28.

Charles Saunders, John Murray-Bruce, and Vivek K. Goyal. Computational periscopy with an ordinary digital camera. *Nature*, 565(7740):472–475, 2019. doi: 10.1038/s41586-018-0868-6.

Prafull Sharma, Miika Aittala, Yoav Y. Schechner, Antonio Torralba, Gregory W. Wornell, William T. Freeman, and Frédo Durand. What you can learn by staring at a blank wall. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2330–2339, October 2021.

Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision Transformers for Single Image Dehazing. *IEEE Transactions on Image Processing*, 32, pp. 1927–1941, 2023. doi: 10.1109/tip.2023.3256763.

Kenichiro Tanaka, Yasuhiro Mukaigawa, and Achuta Kadambi. Polarized non-line-of-sight imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2133–2142, June 2020. doi: 10.1109/cvpr42600.2020.00221.

Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G. Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, 3:745, 2012. doi: 10.1038/ncomms1747.

Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the International Conference on Machine Learning*, pp. 1096–1103, 2008. doi: 10.1145/1390156.1390294.

Xiaolong Wang, Ross Girshick, Abhinav Gupta, and Kaiming He. Non-local neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7794–7803, 2018. doi: 10.1109/CVPR.2018.00813.

Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general U-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17683–17693, 2022. doi: 10.1109/cvpr52688.2022.01716. URL https://openaccess.thecvf.com/content/CVPR2022/html/Wang_Uformer_A_General_U-Shaped_Transformer_for_Image_Restoration_CVPR_2022_paper.html.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. doi: 10.1109/TIP.2003.819861.

Qingsen Yan, Yixu Feng, Cheng Zhang, Guansong Pang, Kangbiao Shi, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. HVI: A new color space for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5678–5687, 2025. doi: 10.1109/CVPR52734.2025.00533.

Adam B. Yedidia, Manel Baradad, Christos Thrampoulidis, William T. Freeman, and Gregory W. Wornell. Using unknown occluders to recover hidden scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12231–12239, 2019. doi: 10.1109/CVPR.2019.01251.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14821–14831, 2021. doi: 10.1109/cvpr46437.2021.01458. URL https://openaccess.thecvf.com/content/CVPR2021/html/Zamir_Multi-Stage_Progressive_Image_Restoration_CVPR_2021_paper.html.

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 (CVPR)*, pp. 5728–5739, June 2022. doi: 10.1109/cvpr52688.2022.00564.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning Enriched Features for Fast Image Restoration and Enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1934–1948, 2023. doi: 10.1109/tpami.2022.3167175.

Jiaxin Zhang, Dezhi Peng, Chongyu Liu, Peirong Zhang, and Lianwen Jin. DocRes: A generalist model toward unifying document image restoration tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15654–15664, June 2024. doi: 10.1109/cvpr52733.2024.01482.

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. doi: 10.1109/tip.2017.2662206.

Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3262–3271, 2018a. doi: 10.1109/CVPR.2018.00344.

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*, pp. 586–595, 2018b. doi: 10.1109/CVPR.2018.00068.

Shihao Zhou, Duosheng Chen, Jinshan Pan, Jinglei Shi, and Jufeng Yang. Adapt or perish: Adaptive sparse transformer with attentive feature refinement for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2952–2963, June 2024. doi: 10.1109/cvpr52733.2024.00285.

APPENDIX

A.1 SUPPLEMENTARY SCENE INFORMATION

The acquisition scenes in Figure 7 pair an indirectly observed wall-light image with the screen content that produced it. A self-emissive display illuminates a nearby white wall, and an ordinary RGB camera records the reflected light while the displayed image is outside its direct view. Tanaka et al. (2020) used a camera-side polarizer and measured wall reflectance to improve two-dimensional reconstruction from indirect light. PLRL learns the inverse mapping from paired images and recovers principal screen regions using the RGB wall observations recorded by a Sony A7S II camera (Figures 3 and 7). The wall spreads and mixes light from different screen regions, producing broad color and intensity variations rather than a directly readable image of the display.

Two observation configurations are represented in the acquisition scenes and dataset examples. In the occluder configuration, an intervening T-shaped board casts a shadow onto the illuminated wall, as visible in the Screen-2000 observations. In the screen-back configuration, the camera faces the rear of the display, recording its dark silhouette together with the surrounding wall response. The Screen-Game, Screen-Figure, and OS-Lock examples retain this screen-back geometry. The input is therefore the recorded observation, which may include a shadow or the rear outline of the display; the paired target is the original screen image. Figure 7 shows both configurations, and Appendix A.3 presents actual input–target pairs for each evaluated domain.

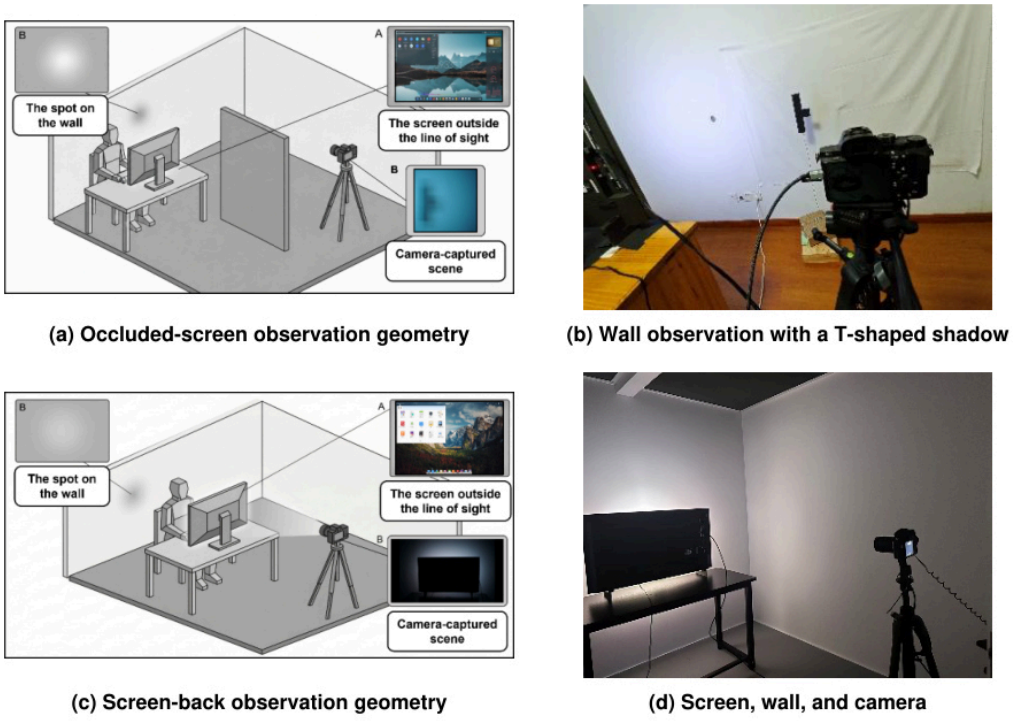


Figure 7: Acquisition configurations for the screen datasets. Top: an occluded-screen schematic and a photograph of the camera recording a wall patch with a T-shaped shadow. Bottom: the screen-back schematic and a real scene in which the camera observes the rear of the display and the illuminated wall. The displayed content supplies the paired target. These views explain the different observation appearances in Screen-2000 and the Screen-Game, Screen-Figure, and OS-Lock examples; the schematics and photographs illustrate the two acquisition configurations used to obtain paired wall-light observations and screen images.

The robustness material further documents changes in camera position, ambient illumination, and wall finish. The corresponding photographs are included alongside the ambient-light, wall-texture, and camera-distance discussions in Appendices A.9, A.10, and A.14. The photographs illustrate variations in camera placement, ambient illumination, and wall finish; the corresponding experiment descriptions specify the conditions and fixed checkpoints used for quantitative evaluation.

A.2 Mechanism and Formula Derivation

Screen light becomes spatially mixed before it reaches the camera, which makes light transport the starting point for explaining PLRL. For screen coordinate $\boldsymbol{\rho} \in \Omega_s$ and wall coordinate $\mathbf{u} \in \Omega_w$, the wall signal is

$$L_{w,c}(\mathbf{u}) = \int_{\Omega_s} T_{\xi,c}(\mathbf{u}, \boldsymbol{\rho}) y_c(\boldsymbol{\rho})\, d\boldsymbol{\rho} + a_{w,c}(\mathbf{u}). \tag{A.1}$$

The kernel in Equation equation A.1 describes how each screen region contributes to the light observed on the wall. Here $y_c : \Omega_s \to [0, 1]$ is the normalized hidden-screen signal in channel c, $T_{\xi,c}$ is the scene-dependent transport kernel, and $a_{w,c}$ is the ambient contribution on the wall. Absolute radiometric gain is absorbed into $T_{\xi,c}$. The state ξ collects geometry, visibility, attenuation, and reflectance, explaining how the same screen content can produce different wall responses as the surroundings change. Paired wall and screen images teach PLRL to recover screen content from these responses.

The camera adds a second stage of spatial mixing as it projects and integrates the reflected wall light. Its observation is

$$x_c(\mathbf{r}) = \int_{\Omega_w} C_{\xi,c}(\mathbf{r}, \mathbf{u}) L_{w,c}(\mathbf{u})\, d\mathbf{u} + n_c(\mathbf{r}). \tag{A.2}$$

The camera-domain quantities specify the evidence available for reconstruction. The coordinate $\mathbf{r}$ indexes a registered camera pixel, $C_{\xi,c}$ describes projection and sensor integration, and n_c is measurement noise. The ambient term in the main-text model follows the same projection, giving $a_c(\mathbf{r}) = \int_{\Omega_w} C_{\xi,c}(\mathbf{r}, \mathbf{u}) a_{w,c}(\mathbf{u})\, d\mathbf{u}$. Thus, $a_{w,c}$ describes ambient light on the wall and a_c describes its contribution in the camera image. The normalized observation $x \in [0, 1]^{C \times H \times W}$ enters PLRL, and the paired hidden-screen image y supplies its training target.

Combining these two stages exposes the screen signal and the disturbances that reconstruction must separate. After discretization,

$$x = A_\xi y + b + n. \tag{A.3}$$

The term $A_\xi y$ contains the hidden-screen information carried through reflection and acquisition. The matrix A_ξ maps the screen grid to the observation grid, b is the sampled camera-domain ambient contribution, and n is discretized noise. PLRL must recover screen structure from their mixture. The next step is to determine why some structures provide a stronger foundation for this recovery than others; the singular system of A_ξ makes that difference explicit.

Light transport preserves different screen components with different strengths. For the singular modes of A_ξ,

$$\begin{aligned} A_\xi v_i = \sigma_i u_i, \qquad A_\xi^{\mathsf{T}} u_i = \sigma_i v_i, \\ \langle x - b, u_i \rangle = \sigma_i \langle y, v_i \rangle + \langle n, u_i \rangle. \end{aligned} \tag{A.4}$$

The strength σ_i determines how clearly a screen component appears above the noise. The unit vectors u_i and v_i are the observation-domain and screen-domain modes, and $\sigma_i \geq 0$ measures their coupling. A small σ_i leaves little screen signal in the corresponding observed component, where noise can then have a large relative effect. Under diffuse propagation, this unequal preservation motivates reconstruction by spatial scale: broad layout offers a stronger starting point, and the recovered layout can guide the interpretation of weaker structures.

Recovering a weakly transmitted component also magnifies errors in the observation, making control over detail corrections necessary. For the minimum-norm perturbation $\delta y^\dagger = A_\xi^\dagger \delta x$,

$$\begin{aligned} \delta y^\dagger &= A_\xi^\dagger \delta x, \\ \langle \delta y^\dagger, v_i \rangle &= \frac{\langle \delta x, u_i \rangle}{\sigma_i} \quad (\sigma_i > 0), \\ \|\delta y^\dagger\|_2^2 &= \sum_{i:\sigma_i > 0} \frac{|\langle \delta x, u_i \rangle|^2}{\sigma_i^2}. \end{aligned} \tag{A.5}$$

The inverse factor $1/\sigma_i$ explains why small wall-image errors can become large reconstructed details. Here $A_\xi^\dagger$ is the Moore-Penrose pseudoinverse and δx is an observation perturbation. As a transport mode weakens, the same perturbation receives greater amplification in the recovered image. PPRL addresses this need for controlled refinement through bounded updates: each stage corrects an established estimate, and the structure recovered so far guides the next correction.

Regularizing the inverse provides a simple reference for how weak components can be stabilized. A classical quadratic estimate is

$$y_\mu = (A_\xi^\top A_\xi + \mu I)^{-1} A_\xi^\top (x - b), \qquad \mu > 0. \tag{A.6}$$

The parameter μ controls the stability of this estimate, with I denoting the identity and y_μ the regularized screen image. This example motivates the two forms of control used in PLRL's learned inverse. Bounded progressive updates establish layout and structure, after which a spatial confidence field controls the contribution of proposed detail. We now express that construction directly in the image domain, beginning with the scale differences that determine the recovery order.

PPRL first separates the wall observation into scales so that broad layout and weaker local variation can play distinct roles. The deterministic image-domain form of the main-text scale projection $\mathcal{M}_\tau$ is

$$\begin{aligned} (\mathcal{M}_\tau u)_c(\mathbf{r}) &= \sum_{\mathbf{v}\in\Omega} h_\tau(\mathbf{r},\mathbf{v}) u_c(\mathbf{v}), \\ h_\tau(\mathbf{r},\mathbf{v}) \geq 0, &\qquad \sum_{\mathbf{v}\in\Omega} h_\tau(\mathbf{r},\mathbf{v}) = 1. \end{aligned} \tag{A.7}$$

The normalized weights make each scale response comparable across the observation and reconstruction. The image $u \in [0,1]^{C\times H\times W}$ can be the wall observation, a candidate reconstruction, or the screen target on the common registered grid Ω. The weights include boundary extension, operate independently on each channel, and preserve constant values at image edges. Applying this operation to the relevant images gives PPRL a shared way to describe their broad and local spatial variations.

Two scales divide those variations into the three components needed for progressive screen recovery. For $\tau_l > \tau_m > 0$,

$$\mathcal{B}_l = \mathcal{M}_{\tau_l}, \qquad \mathcal{B}_m = \mathcal{M}_{\tau_m} - \mathcal{M}_{\tau_l}, \qquad \mathcal{B}_h = I - \mathcal{M}_{\tau_m}. \tag{A.8}$$

Each component has a specific role in recovering the hidden display. The operators $\mathcal{B}_l$, $\mathcal{B}_m$, and $\mathcal{B}_h$ describe broad layout, intermediate structure, and fine variation, respectively. These fixed transforms are used during both training and inference. By exposing the scale differences within the wall pattern, they provide the inputs from which PPRL learns to establish the main screen regions before refining their boundaries and detail.

The three components preserve the complete wall observation through an exact partition. They satisfy

$$\mathcal{B}_l + \mathcal{B}_m + \mathcal{B}_h = I, \qquad x_b = \mathcal{B}_b x, \quad b \in \{l, m, h\}. \tag{A.9}$$

Every observed component remains available to the learned recovery stages. Each $x_b \in \mathbb{R}^{C\times H\times W}$ contributes to $x = x_l + x_m + x_h$, so the partition organizes the wall evidence by scale while retaining its full image-domain content. The learned inverse determines how each part contributes to the screen estimate. This separation gives the network an explicit recovery order while allowing all three components to inform its corrections.

The learned correction operators turn these scale components into successive changes to the screen estimate. At the composite-operator level,

$$\mathbf{z}_x = (x_l, x_m, x_h), \qquad r_j = \mathcal{R}_{j,\theta}(\mathbf{z}_x), \quad j \in \{l, s, d\}. \tag{A.10}$$

The corrections retain the spatial form needed to refine screen regions and boundaries. Each $r_j \in \mathbb{R}^{C\times H\times W}$ is an image-shaped field, θ denotes learned parameters, and j assigns its layout, structure, or detail role. The structure and detail operators include the dependence on preceding

estimates described in the main text. Since those estimates are themselves functions of $\mathbf{z}_x$, Equation equation A.10 collects the full dependence on the wall observation into a compact expression.

PPRL applies the corrections in a sequence that gives weak detail an established structure to refine. The image estimates are

$$\begin{aligned} l &= \Pi(x_l + \alpha \tanh r_l), \\ s &= \Pi(l + \beta \tanh r_s), \\ d &= \gamma \tanh r_d, \qquad p = \Pi(s + d). \end{aligned} \tag{A.11}$$

The first two stages establish the spatial organization that the final stage uses. The images l, s, and p are the layout estimate, structure estimate, and candidate reconstruction in $[0,1]^{C\times H\times W}$; d is the proposed detail residual. The projection Π clips values to the valid image range, and the positive coefficients α, β, and γ set the correction radii. The resulting sequence moves from the wall pattern's broad support to a structured screen candidate with a separately controlled detail contribution.

Bounded updates keep each refinement tied to the image already recovered. Since Π is nonexpansive and $|\tanh t| \le 1$,

$$\|l - x_l\|_\infty \le \alpha, \qquad \|s - l\|_\infty \le \beta, \qquad \|d\|_\infty \le \gamma, \qquad \|p - s\|_\infty \le \gamma. \tag{A.12}$$

These inequalities specify the maximum contribution of every recovery stage. The infinity norm measures the largest absolute component across channels and pixels, and the bounds apply during training and inference. In particular, the final proposal stays within the detail radius γ of the recovered structure. This is the image-domain control motivated by weak transport modes: the model can refine a screen boundary while retaining an explicit limit on how far the correction moves the existing estimate.

Paired screen targets teach the early stages which layout and structure the wall evidence should recover. The corresponding target components are

$$t_l = \mathcal{M}_{\tau_l} y, \qquad t_m = (\mathcal{M}_{\tau_m} - \mathcal{M}_{\tau_l}) y. \tag{A.13}$$

The target components connect the recovery order to the actual hidden-screen image. Both t_l and t_m have the screen image's spatial shape, with t_l representing its broad layout and t_m its intermediate structure. These components assign a concrete recovery objective to each supervised stage. At inference, the learned stages construct l and s from the wall observation, carrying forward the layout and structural relationships learned from the paired examples.

Progressive supervision trains each intermediate estimate to recover its assigned spatial scale. Using the mean absolute norm,

$$\begin{aligned} \|u\|_{1,\mathrm{m}} &= \frac{1}{N_u} \sum_{k=1}^{N_u} |u_k|, \\ \mathcal{L}_{\mathrm{prog}} &= \lambda_l \|\mathcal{M}_{\tau_l} l - t_l\|_{1,\mathrm{m}} \\ &\quad + \lambda_s \|(\mathcal{M}_{\tau_m} - \mathcal{M}_{\tau_l}) s - t_m\|_{1,\mathrm{m}}. \end{aligned} \tag{A.14}$$

This loss gives layout and structure their own learning signals before detail calibration begins. Here N_u is the number of scalar entries in u, and $\lambda_l, \lambda_s \ge 0$ weight the mean absolute errors for layout and structure. The main-text $\mathcal{R}_{\mathrm{prog}}$ measures changes between successive recovery states, while $\mathcal{L}_{\mathrm{prog}}$ measures agreement with target scale components. Together, these descriptions specify the progression and its supervision. Once a candidate has been formed, the remaining question concerns which of its local details agree with the available wall evidence.

LPCR begins that assessment by comparing the observation and candidate at each scale. Its compatibility maps are

$$D_b(\mathbf{r}) = \frac{1}{C} \sum_{c=1}^{C} \left| (\mathcal{B}_b x)_c(\mathbf{r}) - (\mathcal{B}_b p)_c(\mathbf{r}) \right|, \qquad b \in \{l, m, h\}. \tag{A.15}$$

The maps describe where the wall pattern and recovered screen candidate differ in broad and local variation. Each nonnegative $D_b \in \mathbb{R}_{\ge 0}^{H\times W}$ averages channelwise absolute differences. The registered wall observation and screen candidate occupy distinct image domains, so LPCR learns to

interpret these differences as cross-domain compatibility cues together with the restoration states. Their role is to inform the assessment of a proposed correction through the relation between observed light and recovered screen structure.

Edge information then reveals how the proposed detail changes local screen structure. We define

$$\begin{aligned} Y(u)(\mathbf{r}) &= \mathbf{w}^{\mathsf{T}} u(\mathbf{r}), \\ G_u(\mathbf{r}) &= \sqrt{|\partial_1 Y(u)(\mathbf{r})|^2 + |\partial_2 Y(u)(\mathbf{r})|^2 + \varepsilon}, \\ \mathbf{g}(\mathbf{r}) &= [G_x, G_s, G_p, |G_p - G_s|](\mathbf{r}). \end{aligned} \tag{A.16}$$

The edge state places each local correction in the context of the boundary it modifies. The fixed vector $\mathbf{w} \in \mathbb{R}^C$ converts color to luminance, ∂_1 and ∂_2 are fixed spatial derivatives, and $\varepsilon > 0$ stabilizes the magnitude. The four channels of $\mathbf{g}$ record edge magnitudes in the observation, structure, and candidate, together with the change $|G_p - G_s|$ introduced by detail. This final cue identifies where refinement alters edge strength, complementing the scale differences used to assess its compatibility.

Image-wide context helps LPCR interpret a local response within the screen's overall brightness and organization. Retaining the main-text state ζ, we denote the detailed luminance-based encoding by ζ_{c} and its spatial mean by $\bar{\zeta}_{\mathrm{c}}$:

$$\begin{aligned} \zeta(\mathbf{r}) &= [x(\mathbf{r}), l(\mathbf{r}), s(\mathbf{r}), p(\mathbf{r}), \\ &\qquad D_l(\mathbf{r}), D_m(\mathbf{r}), D_h(\mathbf{r}), \nabla s(\mathbf{r}), \nabla p(\mathbf{r})], \\ \zeta_{\mathrm{c}}(\mathbf{r}) &= [x(\mathbf{r}), s(\mathbf{r}), p(\mathbf{r}), D_l(\mathbf{r}), \\ &\qquad D_m(\mathbf{r}), D_h(\mathbf{r}), \mathbf{g}(\mathbf{r})], \\ \bar{\zeta}_{\mathrm{c}} &= \frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} \zeta_{\mathrm{c}}(\mathbf{r}). \end{aligned} \tag{A.17}$$

These states collect the evidence needed to judge a detail correction in relation to the complete reconstruction. The state ζ combines wall evidence, layout, structure, candidate, scale compatibility, and structural gradients, as specified in Section 3.3. The detailed encoding ζ_{c} contains x, s, p, the three compatibility maps, and the four scalar edge cues from Equation equation A.16. Its spatial mean provides a reference for the overall illumination and reconstruction pattern. Both encodings are formed from quantities available during inference, allowing local confidence to depend on the observed image as a whole.

Confidence learning connects those observable cues to the agreement between recovered and target edges. For a fixed distribution of the edge-magnitude target in the detailed training objective,

$$\begin{aligned} q^{\dagger}(\mathbf{r}) \in \operatorname*{arg\,min}_{a \in [0,1]} \mathbb{E}\Big[& \\ & |a - \exp(-\kappa |G_{\hat{y}}(\mathbf{r}) - G_y(\mathbf{r})|)| \;\Big|\; \zeta_{\mathrm{c}}|_{\mathcal{N}(\mathbf{r})}, \bar{\zeta}_{\mathrm{c}} \Big], \\ q(\mathbf{r}) = q_\varphi(\mathbf{r}) \approx q^{\dagger}(\mathbf{r}), & \qquad 0 < q(\mathbf{r}) < 1. \end{aligned} \tag{A.18}$$

The learned score uses local and global recovery evidence to estimate how strongly detail should contribute. The scalar a is a candidate confidence value, $\mathcal{N}(\mathbf{r})$ is the local neighborhood, and the expectation is over paired training examples conditioned on local compatibility and image-wide context. The learned readout q_φ predicts this bounded score from the restoration state. A higher score admits a larger fraction of the detail proposal into the screen estimate, linking the confidence prediction directly to the reconstruction.

Local and image-wide information enter the confidence readout through complementary operators. In the detailed formulation, the image-wide map $\Gamma_\varphi(\bar{\zeta}_{\mathrm{c}})$ and scalar readout Φ_φ express the dependence on global context and local state. Section 3.3 describes nonlocal aggregation using the spatial kernel $\Gamma_\zeta(\mathbf{r}, \mathbf{v})$ and context $\mathcal{P}_{\mathrm{conf}}$; Γ_φ here is the learned map of the image-wide summary. These definitions specify how broader evidence informs the spatial confidence field q_φ, whose values weight the proposed detail at each position.

The training target gives a high confidence score to reconstructed edges whose strength agrees closely with the hidden screen. The edge-magnitude target is

$$q^{\star}_{\mathrm{mag}}(\mathbf{r}) = \exp[-\kappa\,\mathrm{sg}(|G_{\hat{y}}(\mathbf{r}) - G_y(\mathbf{r})|)]\,. \tag{A.19}$$

This target turns edge agreement into a direct learning signal for detail calibration. Here $q^{\star}_{\mathrm{mag}} \in (0,1]^{H\times W}$, $\kappa > 0$ controls the decay, and sg stops gradients through the edge-magnitude error. The main-text target $q^{\star} = \exp(-\kappa e)$ uses the vector-gradient discrepancy $e(\mathbf{r}) = \|\nabla\hat{y}(\mathbf{r}) - \nabla y(\mathbf{r})\|_2$, while $q^{\star}_{\mathrm{mag}}$ uses the absolute difference of scalar luminance-edge magnitudes. Close agreement in these magnitudes produces a high target score; a larger discrepancy lowers the requested detail contribution. Paired supervision thereby teaches LPCR to relate the observable restoration state to the accuracy of recovered edge strength.

Absolute-error regression gives the ideal confidence score a precise interpretation through its training targets. For the fixed target distribution,

$$q^{\dagger}(\mathbf{r}) \in \mathrm{med}\Big[\exp(-\kappa|G_{\hat{y}}(\mathbf{r}) - G_y(\mathbf{r})|) \;\Big|\; \zeta_{\mathrm{c}}|_{\mathcal{N}(\mathbf{r})}, \bar{\zeta}_{\mathrm{c}}\Big]. \tag{A.20}$$

The conditional median characterizes the score that best matches edge agreement under this objective. The operator med denotes the set of conditional medians of the training target. Thus, for a fixed target distribution, Equation equation A.20 identifies the optimum of confidence regression given the local compatibility state and image-wide summary. This interpretation supplies the link from training evidence to the next operation: using the predicted score to determine how much detail enters the recovered screen.

The multiplicative readout uses that score to refine the screen while retaining the structure already established by PPRL. It is

$$\begin{aligned} \hat{d}_{\mathrm{mult}} &= q \odot d, \qquad \hat{y} = \Pi(s + \hat{d}_{\mathrm{mult}}), \\ \|\hat{y} - s\|_{\infty} &\le \|q \odot d\|_{\infty} \le \gamma. \end{aligned} \tag{A.21}$$

Confidence now has a concrete effect on the reconstructed image at every location. Its value at each pixel is shared across color channels, and $\odot$ denotes elementwise multiplication. The bound follows from $0 < q < 1$ and PPRL's detail radius γ. The final image therefore combines the established structure with a locally weighted correction computed from the wall observation. Regions assigned greater confidence receive more of the proposed refinement, and lower-confidence regions retain a closer relationship to the structural estimate.

For fixed confidence, the same readout can be understood as choosing a residual under a confidence-dependent amplitude penalty. The exact pointwise identity is

$$q \odot d = \arg\min_{\delta} \left\{ \frac{1}{2}\|\delta - d\|_2^2 + \frac{1}{2}\left\| \sqrt{\frac{1-q}{q}} \odot \delta \right\|_2^2 \right\}. \tag{A.22}$$

The two terms explain how confidence balances refinement against its amplitude. The first keeps the calibrated residual near the PPRL proposal, and the second penalizes its size through $(1-q)/q$. Lower confidence gives stronger shrinkage toward zero detail, while higher confidence passes more of the proposal. Equation equation A.22 characterizes the explicit readout $\hat{d}_{\mathrm{mult}}$ in Equation equation A.21. The main-text $\hat{d}$ is defined by a spatially coupled energy with confidence-weighted proposal fidelity, nonlocal consistency, and structural regularization; the pointwise expression here makes the multiplicative formulation's amplitude control explicit.

The edge-agreement target links this amplitude control to the error in recovered edge strength. Specifically,

$$e_{\nabla} = |G_{\hat{y}} - G_y|, \qquad \lambda^{\star} = \frac{1 - q^{\star}_{\mathrm{mag}}}{q^{\star}_{\mathrm{mag}}} = \exp(\kappa\,\mathrm{sg}(e_{\nabla})) - 1. \tag{A.23}$$

The induced penalty increases as reconstructed and target edge magnitudes move farther apart. Here e_{∇} is their discrepancy map, and $\lambda^{\star}$ is the corresponding detail penalty in Equation equation A.22. Close edge agreement supports a fuller correction; larger discrepancies request stronger shrinkage. This relation closes the chain from the wall observation to calibrated detail: the restoration state

predicts confidence, paired edge agreement trains that prediction, and confidence sets the strength of the final refinement.

The full training objective brings progressive structure learning and confidence-calibrated refinement together. Its detailed form is

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{\text{prog}} &+ \lambda_r \|\hat{y} - y\|_{1,\text{m}} \\ &+ \lambda_e \|G_{\hat{y}} - G_y\|_{1,\text{m}} + \lambda_q \|q - q^{\star}_{\text{mag}}\|_{1,\text{m}}.\end{aligned} \tag{A.24}$$

Each loss contributes to a specific part of hidden-screen recovery. The nonnegative weights λ_r, λ_e, and λ_q balance reconstruction, edge, and confidence errors, while $\mathcal{L}_{\text{prog}}$ supplies layout and structure supervision. During inference, the scale transforms, learned corrections, compatibility state, confidence score, and final readout all operate from the wall observation. The following experiments examine the practical outcome of that sequence: how much principal screen content PLRL recovers as the strength, reliability, and spatial distribution of the wall evidence change.

A.3 Dataset Description

The main reconstruction comparison uses four paired datasets: **Screen-2000**, **Screen-Game**, **Screen-Figure**, and **OS-Lock**. Each pair consists of a camera observation of screen-induced wall light and the corresponding screen image. Table 5 summarizes their content and sample counts. Figures 8–11 show five examples per domain, with the observation above its ground truth (GT). These panels use the actual test-set pairs and contain no model predictions.

Table 5: The four datasets in the main comparison. Counts refer to paired observations and screen targets. Training and test counts follow the evaluation data view used by the reported runs.

Dataset	Screen content	Total	Train	Test
Screen-2000	Desktops, documents, application windows, and dialogs	2,496	2,245	251
Screen-Game	Game boards, puzzles, and graphical game interfaces	1,250	1,000	250
Screen-Figure	Charts, plots, and structured diagrams	1,250	1,000	250
OS-Lock	Operating-system lock and login interfaces	1,399	1,120	279

Screen-2000. This collection covers heterogeneous computer-screen content, including desktop backgrounds, document pages, terminals, application windows, and system dialogs. Figure 8 illustrates the variation from broad desktop color regions to bright document areas and dark foreground windows. The wall observations contain a diffuse color field with an occluder shadow. Screen-2000 contains 2,496 paired examples. We combine its original 2,121 training pairs and 124 validation pairs into a training set of 2,245 pairs and use the remaining 251 pairs for testing.

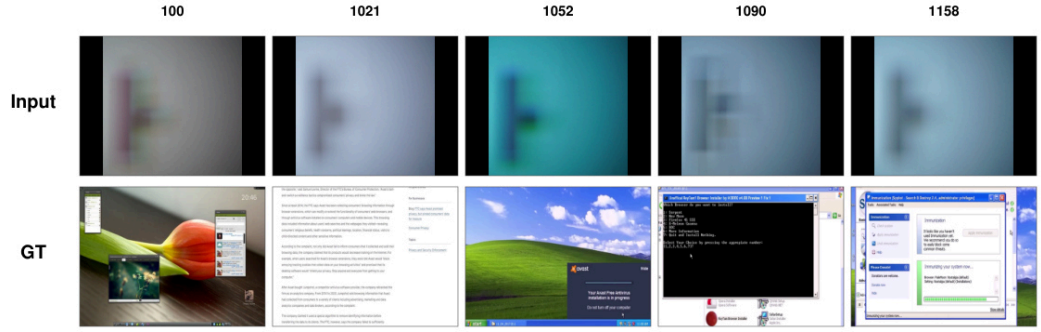


Figure 8: Screen-2000 examples. The upper row shows the wall-light observations and the lower row shows the corresponding screen images: a desktop, a document, a system notification, a terminal window, and a system-settings dialog. Each column is one test pair, identified above the images. Broad screen brightness and color variations remain visible in the diffuse observations, while the window boundaries and text are strongly mixed.

Screen-Game. Screen-Game contains 1,250 pairs, with 1,000 used for training and 250 for testing. Its targets include symbol-matching puzzles, colored-object layouts, connected-node games, and board-game interfaces, as shown in Figure 9. These examples combine a large play area with small pieces, thin links, or compact control panels. The dataset therefore tests whether recovery preserves the play-area layout and the positions of major graphical elements despite the spatial mixing in the wall response.

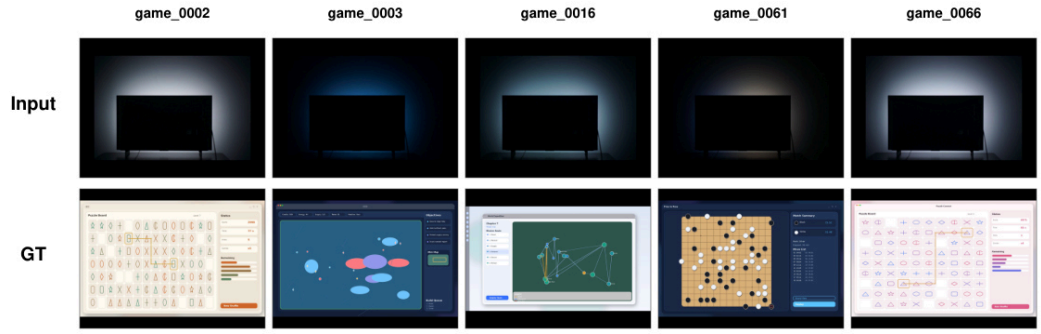


Figure 9: Screen-Game examples, with camera observations in the upper row and paired GT images in the lower row. The five columns show different puzzle, colored-object, connected-node, and board-game layouts. The observation includes the dark rear outline of the display and its surrounding wall light. The targets expose the difference between recovering the broad game layout and resolving individual pieces or links.

Screen-Figure. Screen-Figure contains 1,250 pairs, divided into 1,000 training and 250 test examples. The screen content includes bar charts, scatter plots, hierarchy diagrams, flowcharts, and timeline-style charts. Figure 10 shows representative examples. Large colored chart regions coexist with thin axes, connecting lines, and small labels. This domain evaluates how well reconstruction preserves chart regions, their spatial arrangement, and the main diagram structure. Exact chart-value recovery and label legibility require separate evaluation.

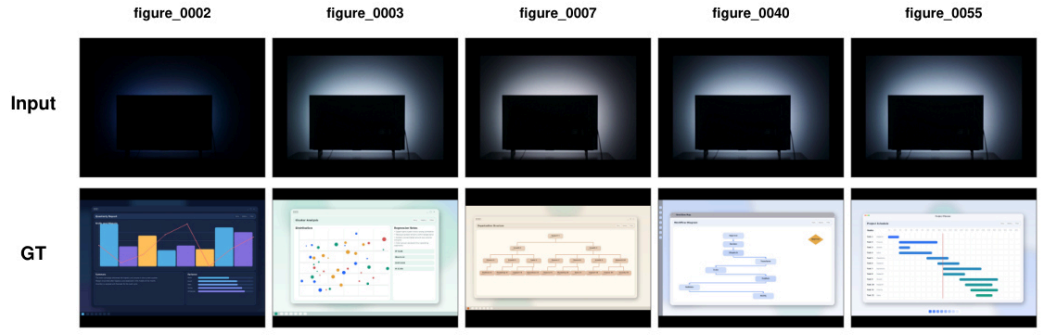


Figure 10: Screen-Figure examples. Each upper-row wall observation is paired with the lower-row chart or diagram displayed on the screen. The examples include a bar chart, scatter plot, hierarchy, flowchart, and timeline. The camera views retain the screen-back silhouette; the GT images reveal the fine lines and small annotations that are much less apparent in the reflected light.

OS-Lock. OS-Lock contains 1,399 pairs, with 1,120 training and 279 test examples. It covers operating-system lock and login screens with different wallpapers, clocks, account icons, and authentication-panel layouts. Figure 11 illustrates both bright natural-image backgrounds and darker, compact login interfaces. The dataset evaluates the appearance and placement of wallpapers, clocks, account icons, and login panels.

Preprocessing and split usage. Inputs and targets are represented as RGB images on a common 512×384 canvas. Aspect-preserving target placement can introduce padding; the valid target bounding box is stored with each sample and is used for metric evaluation. The panels retain the complete



Figure 11: OS-Lock examples, with wall observations above their paired lock or login screens. The examples vary in background brightness, wallpaper, clock position, and login-panel organization. As in Screen-Game and Screen-Figure, the camera records the back of the display with the surrounding reflected light. The examples show the visible organization of operating-system lock and login screens.

preprocessed canvas without brightness enhancement or removal of the screen-back silhouette. A separate model is trained for each domain. Appendix A.4 describes checkpoint selection and the resulting test-set estimates.

Additional domain used in ablations. The module-ablation suite in Tables 2, 3, and 26 uses OS-Lock, OS-Home, Screen-Figure, and Screen-Game. OS-Home contains 2,906 training pairs and 726 test pairs of operating-system home or desktop screens, extending the ablation analysis to this additional content domain.

A.4 Detailed Experimental Setup

Testing feasibility across different screen contents requires paired observations that connect wall light to the display that produced it. Each example pairs a wall-light observation with a hidden-screen target at 512×384 pixels, and we train a separate PLRL model for each domain. Appendix A.3 gives the content, paired examples, and split counts for the four main datasets and the additional OS-Home ablation domain. In particular, the Screen-2000 training view contains 2,245 pairs after merging its source training and validation partitions.

The external comparison measures how published restoration architectures handle the same wall-to-screen recovery problem. Each method retains its published architecture core, uses an RGB input-output interface, and is trained separately on each evaluated dataset. This comparison examines how effectively the adapted architectures recover screen content from diffuse wall observations.

The external architectures include U-Net (Ronneberger et al., 2015), DnCNN (Zhang et al., 2017), NAFNet (Chen et al., 2022), DGUNet (Mou et al., 2022), ConvIR (Cui et al., 2024a), IRNeXt (Cui et al., 2023), DeepSN-Net (Deng et al., 2025), MIRNetv2 (Zamir et al., 2023), OKNet (Cui et al., 2024b), Restormer (Zamir et al., 2022), SFHformer (Jiang et al., 2024), AdaIR (Cui et al., 2025), AST (Zhou et al., 2024), OneRestore (Guo et al., 2024), DocRes (Zhang et al., 2024), DarkIR (Feijoo et al., 2025), DehazeFormer (Song et al., 2023), DEA-Net (Chen et al., 2024), and Pixel2Pixel (Ma et al., 2025), alongside the methods cited in the main text.

Three complementary metrics connect the recovered screen to its target through pixel values, structure, and appearance. Peak Signal-to-Noise Ratio (PSNR; Huynh-Thu & Ghanbari, 2008) measures pixel fidelity in decibels, Structural Similarity Index Measure (SSIM; Wang et al., 2004) measures structural similarity, and AlexNet-based Learned Perceptual Image Patch Similarity (LPIPS; Zhang et al., 2018b) measures perceptual feature distance. Evaluation uses the valid target region; higher PSNR and SSIM and lower LPIPS indicate closer agreement. We use the displayed table precision for numerical comparisons and rankings. Tables 22 and 23 then examine how learning behavior changes with training duration and sample size.

The implementation follows the two difficulties identified in the physical model. PPRL uses $\mathcal{M}_{\tau_l}$ and $\mathcal{M}_{\tau_m}$ to construct coarse, intermediate, and fine evidence, then applies learned layout, structure, and detail corrections. LPCR combines scale compatibility, edge state, and image-wide context to predict q, which weights the detail proposal before addition to the structure estimate. The

confidence-target decay is $\kappa = 4.000$. Scale separation addresses unequal attenuation, and the spatial score addresses differences in the reliability of local detail.

The training settings specify how these recovery stages and confidence predictions are learned. PLRL contains 964,586 trainable parameters and uses AdamW (Loshchilov & Hutter, 2019) for 10 epochs with learning rate 5×10^{-5}, weight decay 1×10^{-4}, two gradient accumulation steps, and gradient clipping at 1.000. The batch size is 8 for OS-Home and 12 for the other datasets. Layout, structure, reconstruction, edge, and confidence losses receive weights of 0.350, 0.450, 1.000, 0.080, and 0.050, respectively. For the reported runs, checkpoint selection monitored LPIPS on the test split, and the selected checkpoints were evaluated on that same split. No separate held-out validation set was used; the tables give point estimates under this protocol.

We train the external-comparison, main-ablation, and supplementary-ablation suites independently. Table 1, Tables 2 and 3, and Table 26, respectively, compare the complete PLRL model with its alternatives from the corresponding training suite. Each suite's PLRL row provides the reference for assessing the alternatives within that setting.

A.5 Detailed External Comparison

PLRL's advantage across screen domains spans pixel fidelity, structural similarity, and perceptual appearance. In Table 1, PLRL ranks first in all 12 dataset-metric comparisons, with the highest PSNR and SSIM and the lowest LPIPS on all four datasets. Restormer captures long-range interactions for image restoration Zamir et al. (2022). In our wall-to-screen comparison, PLRL raises Screen-2000 PSNR from Restormer's 22.601 to 27.341 dB and reduces LPIPS from 0.3449 to 0.1678, a 51.3% reduction in perceptual distance (Table 1). Its SSIM is 0.8682, compared with Restormer's 0.7774. On Screen-Game, PLRL reaches 18.378 dB, 0.6743, and 0.4736, compared with the strongest competing values of 18.160 dB from NAFNet, 0.6681 from IRNeXt, and 0.5676 from Restormer. On OS-Lock, PLRL has the highest PSNR of 17.088 dB and lowest LPIPS of 0.6273; its SSIM of 0.6297 exceeds the strongest competing value of 0.6236 from DGUNet. On Screen-Figure, PLRL achieves the highest PSNR of 20.336 dB and SSIM of 0.7457, compared with the strongest competing values of 19.415 dB from HVI-CIDNet and 0.7211 from OKNet, and lowers LPIPS from Restormer's 0.5508 to 0.4712.

The advantage in perceptual recovery directly serves the goal of revealing the hidden screen's principal visual organization. PPRL establishes the broad layout before refining its structure, giving the detail proposal an image with meaningful spatial relationships. LPCR then evaluates the proposal through the wall evidence and the reconstruction state. Together, these operations are designed to preserve the screen regions that remain informative after diffuse transport weakens fine detail. Tables 2 and 3 examine their individual contributions, and Figure 3 shows the resulting organization in reconstructed screens. The metric gains thus lead to a concrete task outcome: closer coarse appearance across different forms of displayed content.

The next question is whether this recovery survives changes in the light that carries the screen information. The supplementary experiments vary display brightness, acquisition perturbations, ambient illumination, wall response, noise, and camera position. Screen-2000 supplies the principal scene comparisons, with additional OS-Lock and Screen-Game evaluations where reported. The measurements show how recovery changes under each perturbation, and the accompanying figures identify the brightness regions, layouts, and contours that remain visible. This sequence follows the light path from screen emission through reflection to camera acquisition before examining the learning choices that support reconstruction.

A.6 Additional Display-Brightness Robustness

Display brightness changes the strength of the screen information reaching the wall, making it a direct test of recovery from weak light. The main text evaluates this change on Screen-2000, and the supplementary experiments extend it to OS-Lock and Screen-Game. Each domain uses its trained model across brightness shifts of -100, -50, 0, +50, and +100 nit from the nominal 300-nit setting. The resulting 200-to-400-nit range tests whether the main screen structure remains recoverable as its contribution to the observed light becomes weaker or stronger.

OS-Lock. PLRL retains the best reconstruction measurements across the complete OS-Lock brightness range in Table 6. It leads all 15 condition-metric comparisons, reaching 15.537 dB

PSNR, 0.614 SSIM, and 0.659 LPIPS at 200 nit. At the nominal 300 nit, the values are 17.088 dB, 0.623, and 0.627; at 400 nit, they are 15.973 dB, 0.612, and 0.635. The advantage at both endpoints shows that the relative improvement extends to screen emission levels on either side of the nominal condition.

Table 6: OS-Lock results under display-brightness shifts of -100, -50, 0, $+50$, and $+100$ nit relative to the 300-nit nominal condition.

Method	Source	-100 nit			-50 nit			0 nit			+50 nit			+100 nit		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	12.165	0.526	0.724	12.209	0.524	0.722	12.237	0.522	0.721	12.252	0.519	0.720	12.249	0.516	0.720
DocRes	CVPR'24	15.052	0.598	**0.664**	15.985	0.607	**0.650**	16.357	0.609	**0.647**	16.218	0.607	**0.646**	15.842	0.603	**0.648**
Pixel2Pixel	TPAMI'25	12.307	0.522	0.728	12.365	0.520	0.729	12.405	0.517	0.730	12.426	0.514	0.732	12.432	0.512	0.733
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	14.505	0.569	0.669	15.078	0.567	0.664	15.241	0.561	0.664	15.105	0.556	0.665	14.839	0.550	0.668
DGUNet	CVPR'22	15.267	0.594	0.677	**16.202**	**0.615**	0.659	**16.547**	**0.622**	0.652	**16.369**	**0.617**	0.650	**15.958**	**0.606**	0.650
Uformer	CVPR'22	14.604	0.560	0.689	15.199	0.566	0.677	15.149	0.565	0.674	14.780	0.558	0.676	14.263	0.549	0.681
MIRNetv2	TPAMI'22	14.061	0.590	0.675	14.886	0.600	0.665	15.211	0.604	0.658	15.133	0.603	0.654	14.893	0.599	0.654
OKNet	AAAI'24	13.311	0.549	0.692	14.140	0.555	0.682	14.663	0.555	0.678	14.805	0.552	0.676	14.697	0.547	0.677
AST	CVPR'24	14.372	0.551	0.692	14.885	0.545	0.690	15.071	0.536	0.692	14.978	0.525	0.696	14.706	0.514	0.700
ConvIR	TPAMI'24	13.721	0.581	0.704	14.476	0.590	0.694	14.910	0.592	0.687	14.990	0.590	0.683	14.815	0.585	0.682
DnCNN	TIP'17	11.691	0.482	0.725	11.762	0.480	0.723	11.819	0.478	0.722	11.866	0.476	0.722	11.898	0.474	0.722
UNet	MICCAI'15	12.075	0.493	0.762	12.121	0.490	0.762	12.179	0.487	0.762	12.229	0.483	0.761	12.265	0.478	0.761
Models for Related Degradations																
DExNet	TPAMI'25	14.583	0.595	0.664	15.276	0.605	0.655	15.518	0.608	0.652	15.474	0.607	0.651	15.259	0.603	0.651
HVI-CIDNet	CVPR'25	14.558	0.565	0.693	15.341	0.573	0.678	15.605	0.575	0.671	15.553	0.573	0.671	15.327	0.566	0.674
Retinexformer	ICCV'23	13.499	0.538	0.667	14.350	0.552	0.656	14.600	0.558	0.651	14.472	0.558	0.650	14.163	0.555	0.652
DehazeFormer	TIP'23	13.730	0.524	0.685	14.593	0.533	0.673	14.873	0.537	0.669	14.812	0.535	0.669	14.581	0.529	0.672
DEA-Net	TIP'24	14.814	0.590	0.701	15.729	0.602	0.683	15.962	0.606	0.674	15.823	0.603	0.672	15.452	0.597	0.674
IRNeXt	ICML'23	**15.276**	**0.604**	0.701	16.091	0.609	0.687	16.134	0.607	0.679	15.766	0.601	0.676	15.217	0.592	0.676
SFHformer	ECCV'24	12.405	0.524	0.738	12.451	0.517	0.739	12.461	0.509	0.740	12.443	0.500	0.742	12.403	0.491	0.745
DarkIR	CVPR'25	8.672	0.090	0.817	8.730	0.092	0.810	8.787	0.095	0.804	8.843	0.097	0.800	8.896	0.099	0.796
PLRL	**Ours**	**15.537**	**0.614**	**0.659**	**16.703**	**0.623**	**0.638**	**17.088**	**0.623**	**0.627**	**16.692**	**0.619**	**0.628**	**15.973**	**0.612**	**0.635**

The brightness experiment changes how strongly the screen contributes to the wall observation while preserving the recovery objective. PLRL leads all three metrics throughout the OS-Lock 200–400-nit range in Table 6, with SSIM remaining between 0.612 and 0.623. The result connects brightness tolerance to the screen's structural organization: different emission strengths still provide useful evidence for its main regions.

Recovery is strongest near nominal brightness, while structural similarity remains close across much of the OS-Lock sweep. Figure 12 shows that PSNR peaks at 300 nit and that SSIM stays between 0.612 and 0.623 from 250 to 400 nit. LPIPS reaches 0.627 at 300 nit and remains close at 350 nit, with a value of 0.628. These distinct trajectories show how brightness affects pixel accuracy and recovered appearance with different sensitivities. The nominal setting supplies the most favorable measured combination, and nearby settings retain much of that structural and perceptual agreement.

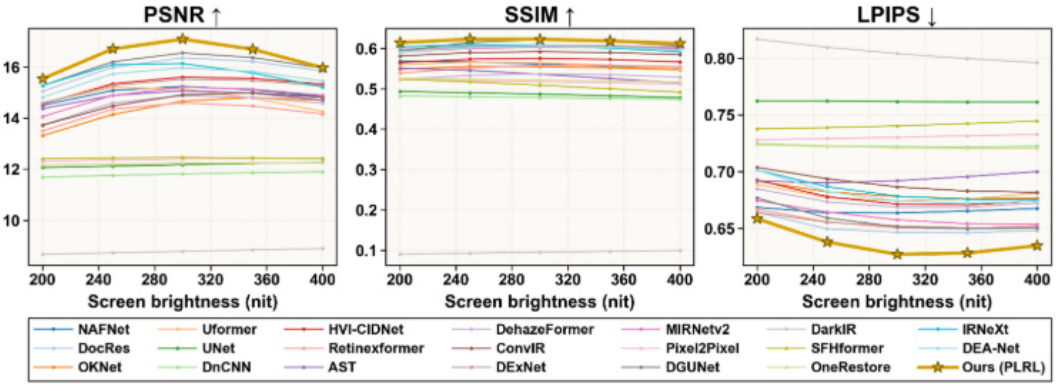


Figure 12: OS-Lock PSNR, SSIM, and LPIPS curves across absolute screen brightness from 200 to 400 nit.

The perceptual curve shows that weaker screen emission still supports a useful coarse estimate. In Figure 12, LPIPS rises to 0.659 at the dimmest condition while remaining the lowest among the compared methods. The retained ranking links the brightness sweep to the feasibility question: the dimmer wall pattern continues to carry information from which PLRL recovers the screen's broad appearance. Its scale-ordered estimate provides a structural reference as the input weakens, while confidence calibration addresses the more uncertain detail. Figure 13 shows the regional divisions through which that coarse appearance can be inspected directly.

The supplementary Screen-Game example in Figure 13 retains its main spatial divisions as display brightness decreases. Across the displayed 300, 275, 250, 225, and 200 nit settings, PLRL recovers the large pale central region and the narrow right-side panel, most clearly near 300 nit. Their divisions remain visible as the display dims, and the lower colored strip supplies an additional reference for the recovered layout. The columns place the wall observation and hidden-screen target alongside restoration baselines and PLRL, making the location and extent of these screen regions the basis for comparison.

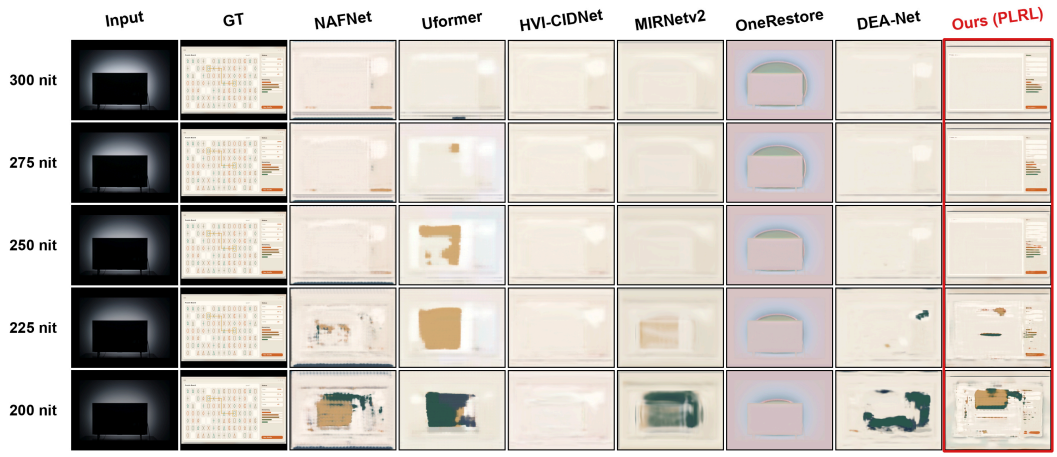


Figure 13: Supplementary Screen-Game reconstructions at absolute display brightness levels of 300, 275, 250, 225, and 200 nit. The columns compare the wall-light observation, hidden-screen target, restoration baselines, and PLRL. The large workspace and right-side panel provide the main spatial cues across the displayed brightness levels.

The dimmer examples in Figure 13 show why layout must provide the foundation for subsequent refinement. The large regions remain the clearest recovered features, while their local appearance changes more strongly toward 200 nit. Broad spatial organization thus remains useful as the finer wall evidence weakens. This relation is the motivation for PPRL's progression from layout to structure and LPCR's separate control over detail. The visual example establishes what coarse recovery looks like under reduced emission; the Screen-Game measurements next quantify recovery across this additional content domain.

Screen-Game. PLRL also leads all 15 brightness condition-metric comparisons on Screen-Game in Table 7. At 200 nit, it obtains 16.620 dB PSNR, 0.624 SSIM, and 0.628 LPIPS. The strongest competing values are 14.178 dB from DocRes, 0.585 SSIM from DExNet, and 0.634 LPIPS from HVI-CIDNet. At the opposite endpoint of 400 nit, PLRL reaches 17.087 dB, 0.641, and 0.601. Its relative advantage therefore persists across the measured range for this second screen-content distribution.

The Screen-Game comparison extends the brightness result from changing light strength to changing screen organization. At 300 nit, PLRL achieves 0.474 LPIPS, compared with the strongest competing value of 0.574 from IRNeXt. Its simultaneous advantages in PSNR, SSIM, and LPIPS show that the improvement includes pixel fidelity, structural agreement, and perceptual features. These measurements support using an ordered structural estimate together with confidence-calibrated detail across the two supplementary domains. The next step is to follow how far the recovered appearance moves from its nominal result as brightness approaches either endpoint.

The Screen-Game curves place the best measured reconstruction around the nominal 300-nit setting. In Figure 14, PLRL reaches 18.378 dB PSNR and 0.674 SSIM there; PSNR becomes 16.620 dB at 200 nit and 17.087 dB at 400 nit. The unequal changes in the two directions show that display brightness interacts with the learned recovery and the scene response. This matters for wall-light reconstruction because the screen signal is observed after reflection and camera acquisition, so the relation between emission strength and recovered image quality is shaped by the complete observation path.

PLRL's perceptual advantage persists even as recovered appearance changes toward the brightness endpoints. Figure 14 shows LPIPS increasing from 0.474 at 300 nit to 0.628 at 200 nit and 0.601 at

Table 7: Screen-Game results under display-brightness shifts of -100, -50, 0, $+50$, and $+100$ nit relative to the 300-nit nominal condition.

Method	Source	-100 nit			-50 nit			0 nit			+50 nit			+100 nit		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.371	0.519	0.765	10.359	0.513	0.757	10.313	0.506	0.747	10.272	0.501	0.742	10.219	0.496	0.740
DocRes	CVPR'24	**14.178**	0.568	0.658	17.194	0.634	**0.592**	17.528	0.646	0.579	**17.407**	0.641	**0.583**	**16.401**	**0.611**	**0.607**
Pixel2Pixel	TPAMI'25	10.080	0.461	0.709	10.645	0.460	0.697	11.140	0.450	0.692	11.358	0.434	0.696	11.264	0.415	0.704
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	12.688	0.520	0.672	**17.425**	0.635	0.599	**18.160**	0.658	0.576	17.329	0.638	0.595	15.214	0.589	0.643
DGUNet	CVPR'22	12.424	0.549	0.685	14.887	0.613	0.629	17.888	0.659	0.595	16.872	0.631	0.627	14.255	0.572	0.693
Uformer	CVPR'22	11.677	0.560	0.697	15.841	0.617	0.643	17.547	0.646	0.614	15.962	0.620	0.645	13.827	0.581	0.687
MIRNetv2	TPAMI'22	13.404	0.581	0.688	17.315	**0.643**	0.624	17.679	0.655	0.612	17.165	**0.644**	0.625	15.074	0.608	0.675
OKNet	AAAI'24	12.005	0.532	0.685	16.588	0.625	0.602	17.891	0.662	0.576	16.448	0.629	0.620	13.758	0.571	0.688
AST	CVPR'24	12.232	0.571	0.687	15.784	0.619	0.642	17.455	0.645	0.616	16.067	0.623	0.640	14.110	0.586	0.678
ConvIR	TPAMI'24	11.839	0.559	0.704	16.379	0.631	0.627	17.641	0.651	0.601	16.072	0.621	0.642	13.788	0.582	0.694
DnCNN	TIP'17	8.681	0.421	0.739	8.749	0.426	0.730	8.738	0.427	0.726	8.703	0.417	0.726	8.607	0.404	0.729
UNet	MICCAI'15	10.170	0.478	0.765	10.282	0.476	0.759	10.285	0.471	0.758	10.256	0.465	0.758	10.245	0.459	0.758
Models for Related Degradations																
DExNet	TPAMI'25	13.243	**0.585**	0.656	16.894	0.639	0.609	17.629	0.658	0.595	16.897	0.639	0.613	15.178	0.601	0.655
HVI-CIDNet	CVPR'25	12.059	0.496	**0.634**	16.955	0.622	0.604	17.773	0.655	0.598	16.975	0.635	0.601	14.568	0.582	0.616
Retinexformer	ICCV'23	11.553	0.489	0.699	15.373	0.588	0.632	17.925	0.655	0.592	17.080	0.633	0.615	14.658	0.570	0.657
DehazeFormer	TIP'23	10.732	0.479	0.719	15.665	0.603	0.631	17.774	0.649	0.607	15.941	0.610	0.647	13.409	0.561	0.700
DEA-Net	TIP'24	10.964	0.509	0.743	16.318	0.623	0.627	17.727	0.657	0.588	16.796	0.635	0.619	14.437	0.590	0.674
IRNeXt	ICML'23	12.780	0.555	0.664	16.535	0.629	0.602	18.135	**0.667**	**0.574**	16.551	0.633	0.616	14.128	0.586	0.679
SFHformer	ECCV'24	10.011	0.481	0.769	9.996	0.470	0.763	10.010	0.461	0.760	9.860	0.453	0.762	9.570	0.444	0.767
DarkIR	CVPR'25	4.259	0.035	0.961	4.364	0.038	0.948	4.465	0.041	0.936	4.559	0.043	0.926	4.647	0.044	0.917
PLRL	**Ours**	**16.620**	**0.624**	**0.628**	**17.649**	**0.648**	**0.520**	**18.378**	**0.674**	**0.474**	**17.628**	**0.652**	**0.520**	**17.087**	**0.641**	**0.601**

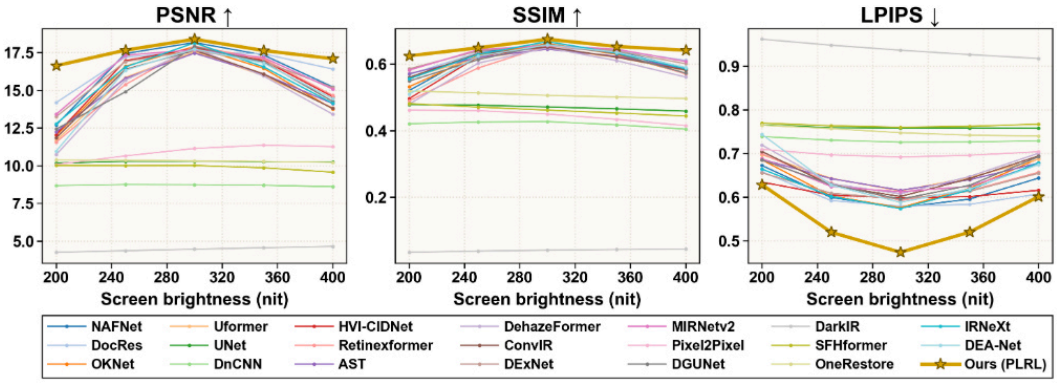


Figure 14: Screen-Game PSNR, SSIM, and LPIPS curves across absolute screen brightness from 200 to 400 nit.

400 nit, with PLRL retaining the best value at every measured setting. This pattern defines the useful outcome precisely: the method recovers coarse screen content over the tested range, and the amount of recoverable structure varies with the available wall signal. The brightness experiment therefore supports the initial feasibility claim while making the role of input strength visible through the full trajectory.

Figure 15 revisits the interface example in Figure 13 to show which spatial relations remain visible across the brightness changes. Its rows again cover 300 to 200 nit, and the PLRL outputs preserve the large pale workspace, narrow right-side panel, and lower colored boundary. These divisions carry the screen's main organization even as the small items within the workspace become strongly simplified. The repeated panel connects the numerical brightness comparison to a concrete screen layout whose main regions can be compared at each displayed emission level.

Regional separation gives PLRL's result in Figure 15 its advantage over several displayed alternatives. Those outputs merge the workspace and surrounding regions into almost uniform areas, while PLRL retains more of the target's distinction between the central workspace and its surrounding controls. This visible distinction matches the task's emphasis on the main displayed organization and the aggregate advantage in Table 7. The brightness analysis has now varied how strongly the screen illuminates the wall. We next disturb the spatial alignment of that light during acquisition, introducing a different challenge for the recovered boundaries.

A.7 Camera-Shake Robustness

Camera shake disturbs the spatial evidence from which the hidden screen must be reconstructed. We test amplitudes of 0, 1, 2, 3, 4, and 6 pixels on Screen-2000 while keeping the model and

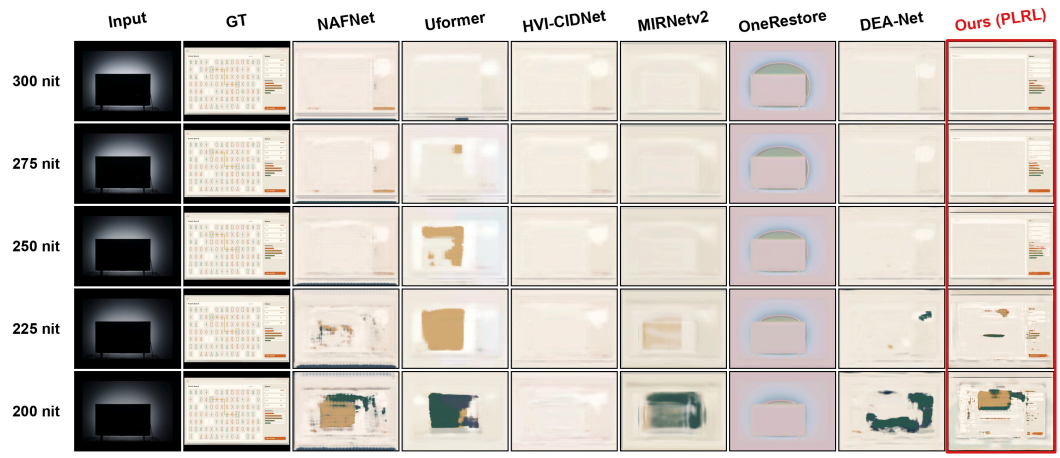


Figure 15: Supplementary Screen-Game reconstructions at absolute display brightness levels of 300, 275, 250, 225, and 200 nit. This panel revisits the example in Figure 13; the large workspace, right-side panel, and lower colored boundary show the retained coarse organization.

other acquisition settings fixed. Table 8 compares reconstruction measurements, Figure 16 follows their changes with shake amplitude, and Figure 17 shows the corresponding screen appearance. This experiment asks whether the broad layout can continue to support recovery as the wall pattern becomes spatially perturbed.

PLRL retains the best measurements at every tested shake amplitude in Table 8. It leads all 18 condition-metric comparisons, including the largest 6-pixel perturbation, where it reaches 23.268 dB PSNR, 0.788 SSIM, and 0.267 LPIPS. NAFNet supplies the strongest competing values at that amplitude: 18.960 dB, 0.701, and 0.441. The remaining separation at the strongest tested shake shows that PLRL recovers closer screen structure and appearance under the tested spatial disturbance.

Table 8: Screen-2000 results under camera-shake amplitudes of 0, 1, 2, 3, 4, and 6 pixels.

Method	Source	0 px			1 px			2 px			3 px			4 px			6 px		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																			
OneRestore	ECCV'24	10.409	0.464	0.714	10.419	0.467	0.723	10.426	0.467	0.730	10.434	0.467	0.735	10.442	0.467	0.738	10.460	0.467	0.741
DocRes	CVPR'24	21.138	0.748	0.400	20.341	0.730	0.420	18.562	0.686	0.456	17.376	0.652	0.487	16.425	0.621	0.514	14.766	0.551	0.573
Pixel2Pixel	TPAMI'25	10.325	0.482	0.664	10.300	0.482	0.672	10.260	0.481	0.681	10.223	0.480	0.687	10.193	0.479	0.690	10.152	0.477	0.693
General-Purpose End-to-End Image Restoration Frameworks																			
NAFNet	ECCV'22	**21.614**	**0.758**	**0.379**	**21.450**	**0.755**	**0.383**	**21.141**	**0.750**	**0.390**	**20.728**	**0.741**	**0.400**	**20.224**	**0.731**	**0.412**	**18.960**	**0.701**	**0.441**
DGUNet	CVPR'22	18.961	0.695	0.482	18.948	0.695	0.484	18.851	0.689	0.492	18.462	0.676	0.506	17.986	0.663	0.520	17.194	0.642	0.544
Uformer	CVPR'22	19.201	0.706	0.464	18.596	0.695	0.479	17.767	0.679	0.498	17.029	0.661	0.515	16.418	0.639	0.531	15.350	0.601	0.556
MIRNetv2	TPAMI'22	18.235	0.691	0.492	18.235	0.691	0.493	18.235	0.691	0.494	18.227	0.691	0.495	18.191	0.688	0.497	17.782	0.673	0.511
OKNet	AAAI'24	20.582	0.738	0.445	20.376	0.729	0.455	19.203	0.698	0.489	17.630	0.655	0.532	16.585	0.624	0.565	15.484	0.586	0.608
AST	CVPR'24	19.389	0.704	0.467	18.385	0.683	0.493	17.088	0.656	0.527	16.381	0.633	0.551	15.382	0.607	0.572	13.734	0.559	0.607
ConvIR	TPAMI'24	20.266	0.728	0.449	19.950	0.708	0.472	19.256	0.685	0.492	18.785	0.672	0.505	18.396	0.663	0.514	17.874	0.650	0.526
DnCNN	TIP'17	8.881	0.346	0.724	8.797	0.347	0.731	8.801	0.347	0.734	8.797	0.346	0.736	8.785	0.346	0.737	8.763	0.345	0.738
UNet	MICCAI'15	9.184	0.240	0.806	9.180	0.242	0.809	9.182	0.243	0.811	9.182	0.243	0.813	9.187	0.243	0.815	9.209	0.243	0.817
Models for Related Degradations																			
DExNet	TPAMI'25	20.018	0.731	0.438	19.888	0.727	0.443	19.482	0.718	0.454	18.754	0.702	0.469	17.899	0.682	0.488	16.702	0.646	0.521
HVI-CIDNet	CVPR'25	19.690	0.717	0.472	17.773	0.665	0.528	17.187	0.648	0.545	16.693	0.633	0.559	16.296	0.621	0.573	15.701	0.602	0.595
Retinexformer	ICCV'23	19.563	0.672	0.455	16.199	0.567	0.544	10.948	0.402	0.684	9.651	0.366	0.722	9.353	0.357	0.732	9.200	0.354	0.739
DehazeFormer	TIP'23	19.357	0.692	0.479	17.724	0.649	0.520	14.749	0.567	0.603	12.464	0.503	0.672	11.138	0.468	0.713	10.173	0.440	0.744
DEA-Net	TIP'24	20.645	0.749	0.424	20.061	0.738	0.440	18.591	0.710	0.476	17.624	0.674	0.511	16.981	0.641	0.542	16.131	0.607	0.586
IRNeXt	ICML'23	19.952	0.723	0.454	19.766	0.711	0.470	19.288	0.699	0.484	18.672	0.682	0.506	18.002	0.664	0.524	16.466	0.629	0.560
SFHformer	ECCV'24	8.817	0.394	0.807	8.807	0.396	0.811	8.798	0.395	0.812	8.791	0.395	0.813	8.787	0.395	0.813	8.780	0.396	0.813
DarkIR	CVPR'25	9.221	0.175	0.867	9.225	0.175	0.871	9.229	0.175	0.874	9.234	0.175	0.875	9.239	0.175	0.875	9.249	0.175	0.876
PLRL	**Ours**	**27.341**	**0.868**	**0.168**	**27.296**	**0.867**	**0.169**	**27.146**	**0.866**	**0.171**	**26.735**	**0.860**	**0.177**	**25.933**	**0.847**	**0.193**	**23.268**	**0.788**	**0.267**

Camera shake tests whether the screen layout can be recovered when the wall pattern shifts locally. Between 0 and 2 pixels, PLRL changes from 27.341 to 27.146 dB PSNR and from 0.168 to 0.171 LPIPS, showing that small spatial disturbances preserve the broad evidence used to establish the screen. At 6 pixels, PLRL retains 23.268 dB, compared with the strongest alternative at 18.960 dB. Figure 17 gives this result a visible form: the central dialog remains a distinct foreground region within the desktop.

Figure 16 adds the structural comparison: SSIM changes from 0.868 to 0.866 between 0 and 2 pixels. Between 4 and 6 pixels, PSNR changes from 25.933 to 23.268 dB and LPIPS from 0.193 to 0.267. The curve shape places the larger reconstruction change near the upper end of the tested disturbance range.

The leading trajectories in Figure 16 show that PLRL retains its relative advantage as shake affects structural precision. The stronger late change accompanies lower SSIM, linking the pixel and per-

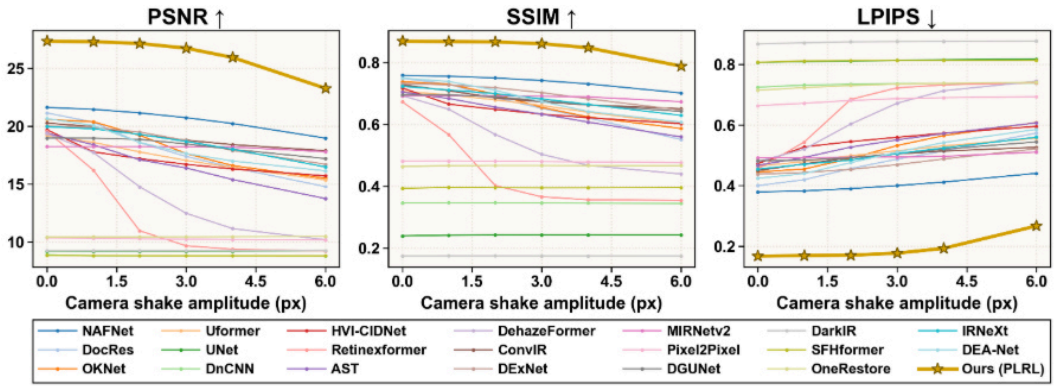


Figure 16: Screen-2000 PSNR, SSIM, and LPIPS curves across camera-shake amplitudes.

ceptual changes to the recovered arrangement of screen content. Bounded updates provide a way to anchor each refinement to the existing estimate, which is particularly relevant when fine observation structure shifts. The curves measure the resulting behavior across the dataset; the next panel shows how a prominent screen element remains represented under the same kind of disturbance.

The central dialog remains a distinct screen element across the shake conditions displayed in Figure 17. At 0, 1, 2, 4, and 6 pixels, PLRL retains the bright dialog against the dark desktop and preserves its position, rectangular extent, and broad internal division more clearly than the compared outputs. These are direct cues to the hidden screen's main content. Their recovery demonstrates how a wall-light image can reveal the presence and placement of a foreground window even as camera shake perturbs the evidence used to reconstruct it.

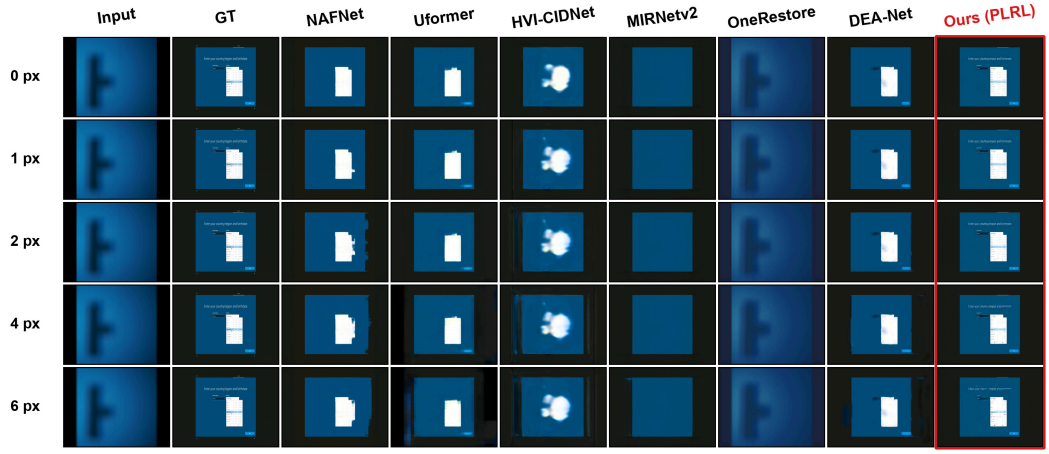


Figure 17: Representative Screen-2000 reconstructions under camera-shake amplitudes of 0, 1, 2, 4, and 6 pixels.

The comparison in Figure 17 shows that increasing shake changes the precision of the dialog more strongly than its broad placement. Several alternatives simplify the window into a solid bright block or an irregular region, while PLRL maintains a structured central window. This visible separation between stable placement and more sensitive local boundaries explains the usefulness of recovering layout before detail. Camera shake changes spatial alignment; motion blur next adds continuous spreading along a direction, placing further pressure on the fine structure already weakened by reflection.

A.8 Motion-Blur Robustness

Motion blur further spreads screen-related variation after diffuse reflection has already weakened fine structure. We apply horizontal motion-blur kernels of 1, 3, 5, 9, 13, and 17 pixels to the recorded Screen-2000 wall-light observations while holding the model and evaluation setting fixed. Increasing the kernel width creates progressively stronger directional smoothing of the wall observation.

This test follows the core reconstruction difficulty directly: as neighboring responses become more mixed, the recovered screen must rely on the broad evidence that remains informative.

PLRL leads all 18 condition-metric comparisons across the motion-blur range in Table 9. At the widest 17-pixel kernel, it achieves 23.018 dB PSNR, 0.779 SSIM, and 0.276 LPIPS. NAFNet provides the strongest competing values of 18.930 dB, 0.698, and 0.446. The advantage extends from the one-pixel reference through the widest tested blur, showing that PLRL's reconstruction remains closer to the hidden-screen target as directional smoothing increases.

Table 9: Screen-2000 results under motion-blur kernels of 1, 3, 5, 9, 13, and 17 pixels.

Method	Source	1 px			3 px			5 px			9 px			13 px			17 px		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																			
OneRestore	ECCV'24	10.409	0.464	0.714	10.421	0.465	0.726	10.426	0.466	0.728	10.439	0.466	0.733	10.455	0.465	0.739	10.470	0.465	0.741
DocRes	CVPR'24	21.138	0.748	0.400	20.092	0.725	0.423	18.527	0.684	0.456	16.547	0.627	0.509	15.113	0.553	0.567	13.734	0.497	0.619
Pixel2Pixel	TPAMI'25	10.325	0.482	0.664	10.300	0.482	0.673	10.261	0.481	0.682	10.194	0.479	0.690	10.158	0.477	0.693	10.135	0.475	0.694
General-Purpose End-to-End Image Restoration Frameworks																			
NAFNet	ECCV'22	**21.614**	**0.758**	**0.379**	**21.495**	**0.754**	**0.384**	**21.200**	**0.750**	**0.390**	**20.561**	**0.736**	**0.406**	**19.656**	**0.717**	**0.426**	**18.930**	**0.698**	**0.446**
DGUNet	CVPR'22	18.961	0.695	0.482	18.939	0.694	0.485	18.825	0.686	0.493	17.765	0.658	0.522	16.840	0.635	0.550	16.678	0.626	0.562
Uformer	CVPR'22	19.201	0.706	0.464	18.452	0.693	0.482	17.631	0.677	0.501	16.475	0.644	0.530	15.678	0.606	0.551	14.785	0.578	0.570
MIRNetv2	TPAMI'22	18.235	0.691	0.492	18.235	0.691	0.493	18.236	0.691	0.493	18.220	0.690	0.495	17.923	0.679	0.505	17.408	0.662	0.523
OKNet	AAAI'24	20.582	0.738	0.445	20.247	0.725	0.460	18.885	0.691	0.495	15.769	0.598	0.589	15.116	0.572	0.620	14.720	0.554	0.646
AST	CVPR'24	19.389	0.704	0.467	17.894	0.676	0.503	16.883	0.652	0.531	15.700	0.612	0.572	13.767	0.565	0.600	12.450	0.526	0.629
ConvIR	TPAMI'24	20.266	0.728	0.449	19.785	0.703	0.476	19.105	0.683	0.496	18.497	0.665	0.512	17.979	0.655	0.526	17.504	0.646	0.535
DnCNN	TIP'17	8.881	0.346	0.724	8.795	0.347	0.731	8.806	0.346	0.733	8.793	0.343	0.737	8.752	0.342	0.738	8.741	0.342	0.738
UNet	MICCAI'15	9.184	0.240	0.806	9.182	0.242	0.809	9.184	0.243	0.809	9.183	0.243	0.811	9.200	0.243	0.815	9.224	0.243	0.819
Models for Related Degradations																			
DExNet	TPAMI'25	20.018	0.731	0.438	19.858	0.726	0.444	19.464	0.718	0.453	18.086	0.688	0.484	16.629	0.642	0.523	16.033	0.620	0.548
HVI-CIDNet	CVPR'25	19.690	0.717	0.472	17.979	0.670	0.520	17.501	0.657	0.535	16.720	0.636	0.557	16.052	0.615	0.579	15.543	0.597	0.601
Retinexformer	ICCV'23	19.563	0.672	0.455	15.059	0.538	0.566	10.041	0.370	0.707	9.325	0.354	0.736	9.183	0.351	0.736	9.129	0.353	0.742
DehazeFormer	TIP'23	19.357	0.692	0.479	17.284	0.640	0.531	14.438	0.562	0.604	10.643	0.456	0.726	9.883	0.434	0.749	9.883	0.432	0.758
DEA-Net	TIP'24	20.645	0.749	0.424	19.955	0.736	0.443	18.131	0.704	0.482	17.156	0.649	0.536	16.325	0.611	0.569	15.783	0.590	0.595
IRNeXt	ICML'23	19.952	0.723	0.454	19.708	0.708	0.473	19.239	0.699	0.487	18.255	0.670	0.521	16.960	0.641	0.555	15.483	0.608	0.582
SFHformer	ECCV'24	8.817	0.394	0.807	8.805	0.396	0.813	8.796	0.394	0.812	8.787	0.395	0.812	8.780	0.395	0.812	8.779	0.397	0.813
DarkIR	CVPR'25	9.221	0.175	0.867	9.225	0.175	0.870	9.229	0.175	0.872	9.238	0.175	0.873	9.246	0.175	0.877	9.255	0.175	0.876
PLRL	**Ours**	**27.341**	**0.868**	**0.168**	**27.281**	**0.867**	**0.169**	**27.180**	**0.866**	**0.170**	**26.509**	**0.856**	**0.182**	**25.065**	**0.829**	**0.216**	**23.018**	**0.779**	**0.276**

Directional blur mixes neighboring wall responses while preserving broader brightness regions. PLRL changes modestly between 1- and 5-pixel kernels, from 27.341 to 27.180 dB PSNR and from 0.168 to 0.170 LPIPS. NAFNet simplifies image restoration through multiplicative operations in place of nonlinear activations, supporting both deblurring and denoising Chen et al. (2022). PLRL establishes regional structure before refining weaker local variation and retains 23.018 dB PSNR under 17-pixel motion blur of the wall observation, compared with 18.930 dB for NAFNet in our Screen-2000 comparison (Table 9).

Figure 18 complements these pixel and perceptual comparisons with SSIM of 0.868 and 0.866 at kernel widths of 1 and 5 pixels. At 13 pixels, PLRL reaches 25.065 dB PSNR and 0.216 LPIPS, retaining substantial agreement before the widest tested blur. These curves connect stronger spatial mixing to the precision of the recovered screen.

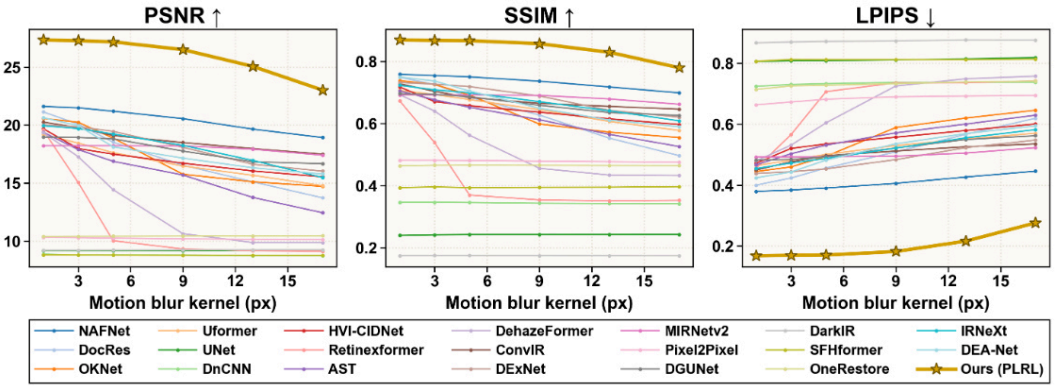


Figure 18: Screen-2000 PSNR, SSIM, and LPIPS curves across motion-blur kernel widths.

The three curves in Figure 18 agree that larger blur widths make screen structure harder to recover precisely. Their joint movement links the loss of pixel accuracy to structural and perceptual changes. This behavior fits the motivation for separating the recovery stages: broad organization can provide useful support while increasingly mixed local evidence makes refinement more demanding. PLRL retains the leading values at every measured width. The qualitative comparison then locates this remaining structural support in the position and extent of a central screen dialog.

The main dialog remains visible in PLRL's outputs across the blur widths shown in Figure 19. At widths of 1, 5, 9, 13, and 17 pixels, the reconstruction preserves a bright central window against

the dark desktop. The panel compares the observation and target with six restoration baselines and PLRL, allowing the dialog's placement and broad divisions to be inspected directly. Table 9 gives a corresponding PSNR separation at the widest kernel: PLRL remains above 23 dB, while the strongest competing value is below 19 dB.

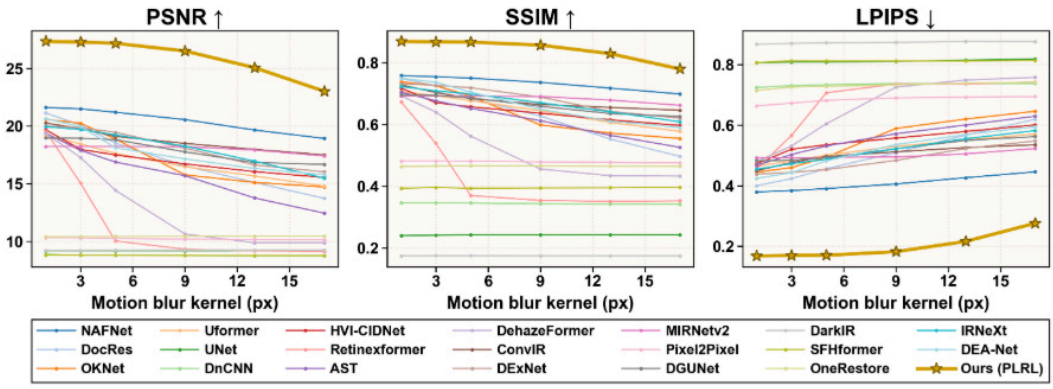


Figure 19: Representative Screen-2000 reconstructions under motion-blur kernel widths of 1, 5, 9, 13, and 17 pixels. The columns compare the wall-light observation, hidden-screen target, six restoration baselines, and PLRL. PLRL retains the central dialog and its broad internal organization across the displayed blur conditions.

The dialog boundaries and broad internal divisions in Figure 19 show the screen structure that remains recoverable after additional directional smoothing. At the 17-pixel width, Table 9 reports 0.779 SSIM and 0.276 LPIPS, linking that visible organization to the structural and perceptual advantage. The progression from layout to calibrated detail is designed for precisely this imbalance between broad and fine evidence. The next experiment changes a different part of the observation: ambient light contributes spatial and color variation that must be interpreted alongside the screen's wall illumination.

A.9 Additional Ambient-Light Robustness

Figure 20 shows acquisition photographs for changes in ambient-light intensity and color temperature. They document how room illumination changes the wall pattern independently of screen content. These photographs illustrate variation in ambient illumination and color temperature; Table 10 evaluates the five directional lighting configurations described below.

Ambient illumination adds its own spatial and color pattern to the wall light that carries the screen content. We evaluate five additional directional lighting configurations on Screen-2000: left warm, right cool, top neutral, bottom warm, and upper-right cool. The model and evaluation setting remain fixed. These configurations test whether PLRL can recover the hidden screen's broad organization when illumination from the surroundings changes the mixture observed by the camera.

PLRL achieves the strongest values in all 15 ambient-light condition-metric comparisons in Table 10. The top-neutral setting gives 23.889 dB PSNR, 0.817 SSIM, and 0.218 LPIPS. Under bottom-warm illumination, PLRL reaches 20.397 dB, 0.727, and 0.321; DEA-Net supplies the strongest competing values of 16.314 dB, 0.659, and 0.528. The leading results span the tested changes in source direction and color, showing that the relative recovery advantage extends across distinct mixtures of screen light and ambient illumination.

Ambient illumination changes the wall pattern while the target screen provides a fixed structural reference. Across the five lighting configurations, PLRL preserves the dominant pale region and its main boundary in Figure 21. Under bottom-warm illumination, it reaches 20.397 dB PSNR, compared with 16.314 dB for the strongest alternative. The comparison therefore connects the metric advantage to the intended outcome: the same screen organization remains recoverable as the surrounding illumination changes.

The dominant pale screen region remains visible across the five lighting configurations in Figure 21. PLRL preserves its main boundary and the placement of visible colored elements, while several compared outputs exhibit stronger dark bands, broad color shifts, or more heavily mixed internal regions. The target provides a fixed reference for comparing the left-warm, right-cool, top-neutral,

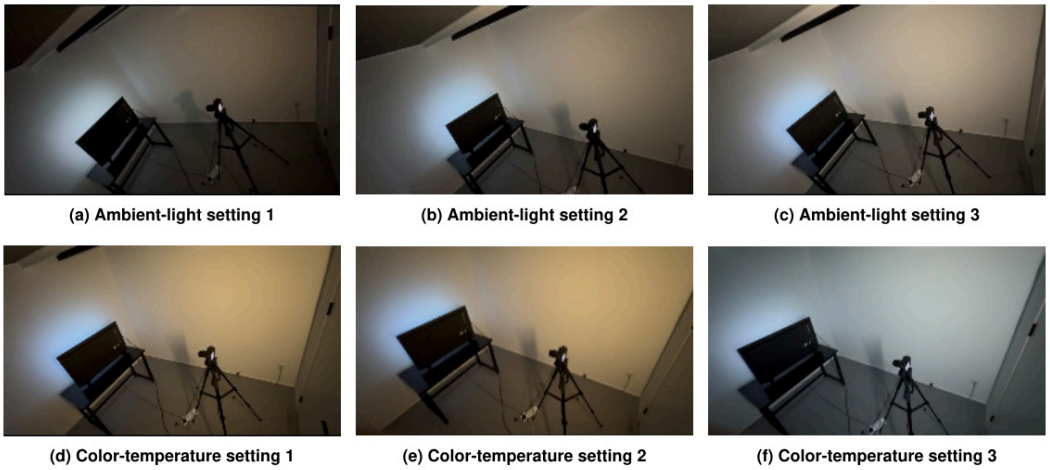


Figure 20: Acquisition photographs illustrating illumination variation. The upper row shows three ambient-light settings, and the lower row shows three color-temperature settings with the screen, wall, and camera present. The images illustrate changes in the background illumination surrounding the screen-induced wall response.

Table 10: Screen-2000 results under five additional directional ambient-light configurations.

Method	Source	Left warm			Right cool			Top neutral			Bottom warm			Upper-right cool		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.149	0.464	0.714	10.364	0.469	0.708	10.215	0.464	0.714	10.127	0.465	0.708	10.134	0.466	0.711
DocRes	CVPR'24	13.237	0.458	0.667	12.808	0.439	0.684	14.529	0.526	0.597	13.128	0.453	0.669	12.409	0.415	0.705
Pixel2Pixel	TPAMI'25	9.799	0.457	0.676	9.795	0.458	0.671	9.760	0.458	0.676	9.652	0.452	0.674	9.659	0.453	0.678
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	17.425	0.643	0.485	16.827	0.628	**0.500**	17.644	0.658	**0.469**	14.661	0.569	0.562	14.311	0.546	0.587
DGUNet	CVPR'22	14.775	0.570	0.577	16.078	0.600	0.544	16.568	0.607	0.537	13.687	0.547	0.598	14.256	0.555	0.592
Uformer	CVPR'22	17.267	0.657	0.507	14.293	0.628	0.545	16.368	0.663	0.512	14.594	0.632	0.548	13.311	0.606	0.563
MIRNetv2	TPAMI'22	14.148	0.568	0.592	15.436	0.590	0.560	16.622	0.622	0.525	13.749	0.562	0.595	13.398	0.544	0.614
OKNet	AAAI'24	15.288	0.590	0.570	15.694	0.618	0.550	15.758	0.605	0.556	13.048	0.548	0.632	13.561	0.555	0.622
AST	CVPR'24	14.282	0.604	0.562	13.562	0.595	0.571	15.013	0.623	0.546	13.020	0.573	0.601	12.881	0.567	0.594
ConvIR	TPAMI'24	14.789	0.572	0.568	15.059	0.594	0.553	16.029	0.593	0.541	13.458	0.541	0.603	13.723	0.548	0.605
DnCNN	TIP'17	8.593	0.358	0.731	8.684	0.357	0.737	8.511	0.348	0.739	8.202	0.353	0.743	8.277	0.355	0.747
UNet	MICCAI'15	9.143	0.240	0.808	9.385	0.237	0.796	9.231	0.242	0.801	9.269	0.240	0.802	9.306	0.239	0.799
Models for Related Degradations																
DExNet	TPAMI'25	17.408	0.663	0.501	**17.560**	**0.665**	0.501	18.297	0.686	0.479	15.676	0.613	0.545	**16.245**	0.624	**0.538**
HVI-CIDNet	CVPR'25	10.614	0.412	0.751	11.959	0.457	0.712	11.137	0.422	0.736	10.425	0.407	0.756	11.629	0.448	0.727
Retinexformer	ICCV'23	8.831	0.331	0.759	9.317	0.350	0.743	9.278	0.352	0.740	8.856	0.334	0.750	9.172	0.346	0.746
DehazeFormer	TIP'23	15.758	0.606	0.562	16.179	0.614	0.553	16.412	0.623	0.542	14.220	0.577	0.598	14.334	0.574	0.600
DEA-Net	TIP'24	**18.509**	**0.690**	**0.485**	16.881	0.663	0.520	**18.482**	**0.698**	0.476	**16.314**	**0.659**	**0.528**	15.404	**0.627**	0.567
IRNeXt	ICML'23	15.228	0.582	0.567	16.108	0.616	0.531	16.389	0.605	0.537	14.345	0.567	0.589	14.068	0.557	0.600
SFHformer	ECCV'24	8.224	0.365	0.818	8.805	0.396	0.817	8.601	0.382	0.814	8.453	0.381	0.813	8.708	0.390	0.825
DarkIR	CVPR'25	9.168	0.145	0.884	9.247	0.145	0.883	9.176	0.147	0.881	9.156	0.139	0.885	9.185	0.136	0.890
PLRL	**Ours**	**23.064**	**0.790**	**0.249**	**23.347**	**0.796**	**0.246**	**23.889**	**0.817**	**0.218**	**20.397**	**0.727**	**0.321**	**20.559**	**0.729**	**0.318**

bottom-warm, and upper-right-cool outputs. Retaining these large regions links the lighting experiment to the intended task outcome: a coarse reconstruction of the screen's principal appearance.

The fixed screen layout in Figure 21 makes the benefit of interpreting local brightness through wider structure visible. The wall inputs vary with illumination direction, yet PLRL continues to reproduce the large target regions across the displayed conditions. This behavior is consistent with using the reconstruction state and image-wide context when weighting proposed detail. The visual comparison therefore adds a spatial interpretation to Table 10: the retained measurements correspond to a recoverable main boundary and regional arrangement even as the surrounding light changes. Figure 22 then compares the difficulty of the individual lighting configurations.

Different lighting directions produce different reconstruction quality, with PLRL leading throughout the five-condition comparison. Figure 22 gives PSNR values of 23.064, 23.347, and 23.889 dB under left-warm, right-cool, and top-neutral light, followed by 20.397 and 20.559 dB under bottom-warm and upper-right-cool light. SSIM and LPIPS follow the same broad grouping. These trajectories show that the location and color of additional illumination matter to the recoverable screen evidence, beyond the presence of ambient light alone.

The agreement across metrics in Figure 22 ties the illumination-dependent changes to recovered structure and appearance. Source direction changes where additional light overlaps the informative

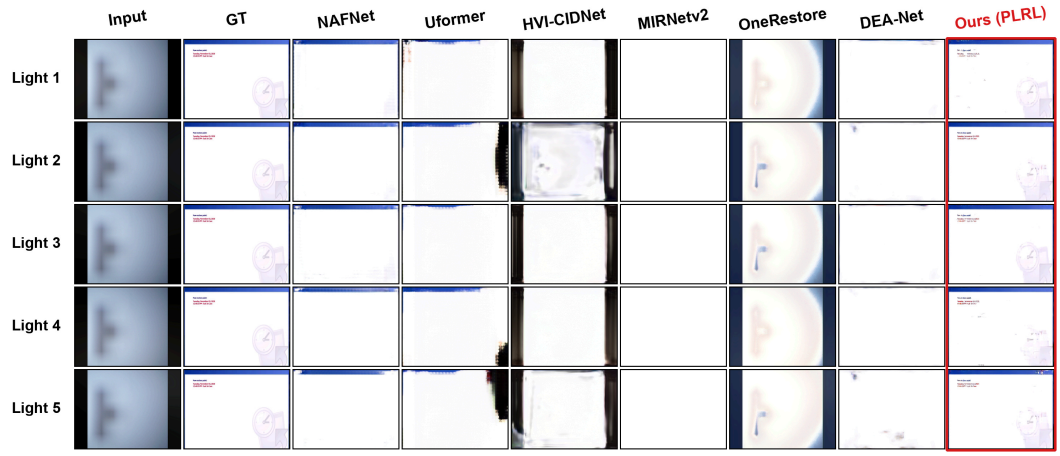


Figure 21: Representative Screen-2000 reconstructions under left-warm, right-cool, top-neutral, bottom-warm, and upper-right-cool ambient illumination. The main screen boundary and large regions remain visible in the PLRL outputs across the five configurations.

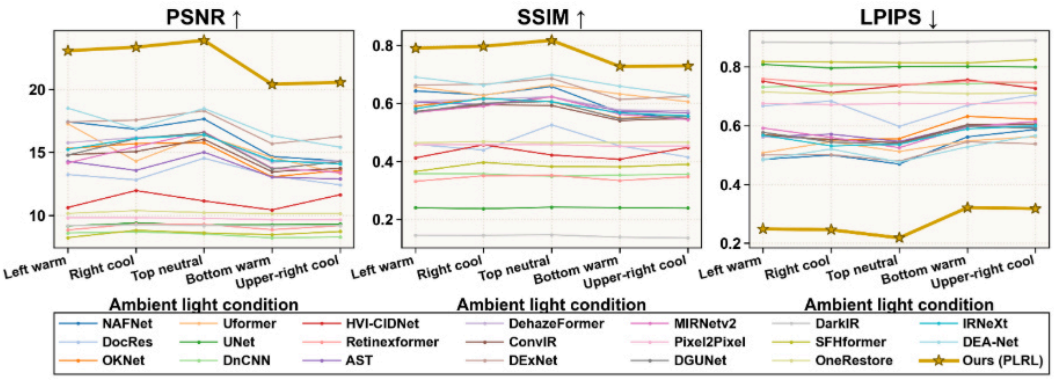


Figure 22: Screen-2000 PSNR, SSIM, and LPIPS curves under left-warm, right-cool, top-neutral, bottom-warm, and upper-right-cool ambient illumination.

wall pattern, and color temperature changes the balance among channels. PLRL retains the best measurements across these configurations, supporting its use of structural and contextual cues under the tested mixtures. The illumination experiments have varied the light reaching the wall. We next examine the surface that reflects that light, beginning with texture that contributes spatial variation of its own.

A.10 Wall-Texture Robustness

Figure 23 documents three wall finishes in the acquisition material: wood, tile, and matte white. The scene views and surface details show why wall reflectance and texture can alter the recorded light pattern. These photographs complement the controlled wall-response perturbations evaluated below; their material names are distinct from the numerical perturbation levels.

Wall texture introduces spatial patterns that compete with the weak structure carried by screen light. We evaluate smooth, fine, medium, coarse, and strong texture conditions on Screen-2000 while keeping the scene, model, and evaluation setting fixed. This experiment addresses a central ambiguity in hidden-screen recovery: an observed local variation can reflect the screen or the surface on which its light falls. The comparison tests how well the reconstructed screen retains its organization as that surface contribution increases.

PLRL leads all 15 condition-metric comparisons across the wall textures in Table 11. Under strong texture, it reaches 19.753 dB PSNR, 0.719 SSIM, and 0.451 LPIPS. DExNet supplies the strongest competing PSNR of 19.709 dB, while OKNet supplies the strongest competing SSIM of 0.684 and



Figure 23: Wall finishes in the robustness acquisition scenes. The upper row shows the display beside wood-finished, tiled, and matte white walls; the lower row shows the corresponding surface details. They illustrate reflectance and surface-texture variation relevant to the wall-texture and wall-roughness studies.

LPIPS of 0.458. The endpoint advantage is clearest in structural similarity, directly connecting the comparison to the task of recovering the hidden display's main regional arrangement.

Table 11: Screen-2000 results under five wall-texture conditions.

Method	Source	Smooth texture			Fine texture			Medium texture			Coarse texture			Strong texture		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.409	0.464	0.714	10.410	0.455	0.721	10.412	0.429	0.774	10.409	0.389	0.818	10.411	0.331	0.826
DocRes	CVPR'24	21.138	0.748	0.400	16.591	0.574	0.547	14.467	0.465	0.640	13.229	0.409	0.699	11.896	0.363	0.762
Pixel2Pixel	TPAMI'25	10.325	0.482	0.664	10.026	0.459	0.700	9.967	0.446	0.717	9.944	0.432	0.730	9.934	0.427	0.719
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	**21.614**	**0.758**	**0.379**	19.879	0.716	**0.423**	17.956	0.649	0.487	17.212	0.608	0.527	16.289	0.534	0.576
DGUNet	CVPR'22	18.961	0.695	0.482	18.890	0.690	0.485	18.855	0.684	0.495	18.818	0.669	0.511	18.775	0.627	0.534
Uformer	CVPR'22	19.201	0.706	0.464	18.556	0.690	0.481	17.062	0.654	0.523	16.316	0.619	0.565	16.457	0.590	0.606
MIRNetv2	TPAMI'22	18.235	0.691	0.492	18.070	0.684	0.497	17.942	0.678	0.501	17.880	0.673	0.506	17.775	0.664	0.508
OKNet	AAAI'24	20.582	0.738	0.445	**20.424**	**0.727**	0.448	14.124	0.368	0.678	13.178	0.294	0.711	14.881	**0.684**	**0.458**
AST	CVPR'24	19.389	0.704	0.467	18.765	0.689	0.483	15.153	0.604	0.557	13.706	0.551	0.607	15.323	0.558	0.618
ConvIR	TPAMI'24	20.266	0.728	0.449	19.970	0.713	0.456	**19.828**	**0.697**	**0.467**	**19.905**	0.674	0.496	19.425	0.604	0.558
DnCNN	TIP'17	8.882	0.346	0.724	8.868	0.339	0.744	8.843	0.320	0.770	8.845	0.286	0.785	8.862	0.246	0.793
UNet	MICCAI'15	9.184	0.240	0.806	9.192	0.237	0.814	9.199	0.229	0.838	9.211	0.215	0.855	9.238	0.200	0.832
Models for Related Degradations																
DExNet	TPAMI'25	20.018	0.731	0.438	18.159	0.688	0.483	17.313	0.663	0.512	19.194	**0.708**	**0.462**	**19.709**	0.386	0.626
HVI-CIDNet	CVPR'25	19.690	0.717	0.472	13.583	0.504	0.663	12.617	0.452	0.705	12.193	0.433	0.710	11.946	0.417	0.719
Retinexformer	ICCV'23	19.563	0.672	0.455	18.438	0.640	0.481	16.769	0.560	0.550	11.343	0.356	0.724	9.523	0.290	0.774
DehazeFormer	TIP'23	19.357	0.692	0.479	17.824	0.655	0.521	16.014	0.591	0.581	15.010	0.560	0.613	15.094	0.565	0.617
DEA-Net	TIP'24	20.645	0.749	0.424	18.471	0.703	0.486	14.631	0.579	0.637	11.060	0.464	0.742	11.938	0.510	0.700
IRNeXt	ICML'23	19.952	0.723	0.454	19.796	0.713	0.458	19.160	0.691	0.483	17.469	0.634	0.542	15.264	0.536	0.629
SFHformer	ECCV'24	8.817	0.394	0.807	8.817	0.391	0.815	8.810	0.380	0.848	8.807	0.361	0.869	8.807	0.330	0.839
DarkIR	CVPR'25	9.221	0.175	0.867	9.205	0.173	0.861	9.190	0.169	0.880	9.171	0.161	0.905	9.143	0.146	0.901
PLRL	**Ours**	**27.341**	**0.868**	**0.168**	**20.901**	**0.732**	**0.353**	**20.317**	**0.725**	**0.451**	**20.158**	**0.717**	**0.455**	**19.753**	**0.719**	**0.451**

PLRL retains the leading reconstruction results across all five wall-texture conditions in Table 11. From fine to strong texture, its SSIM remains between 0.717 and 0.732, showing persistent agreement in the screen's regional structure. At the strong-texture endpoint, PLRL reaches 0.719 SSIM, compared with the strongest alternative at 0.684. Figure 25 locates this structural advantage in the positions and divisions of the main interface panels.

The curves in Figure 24 distinguish the initial introduction of wall texture from changes in its strength. PSNR moves from 27.341 dB on the smooth wall to 20.901 dB under fine texture, and then ranges down to 19.753 dB as texture strengthens. The maintained structural agreement across these later conditions shows that the added wall pattern continues to leave evidence for the main screen organization.

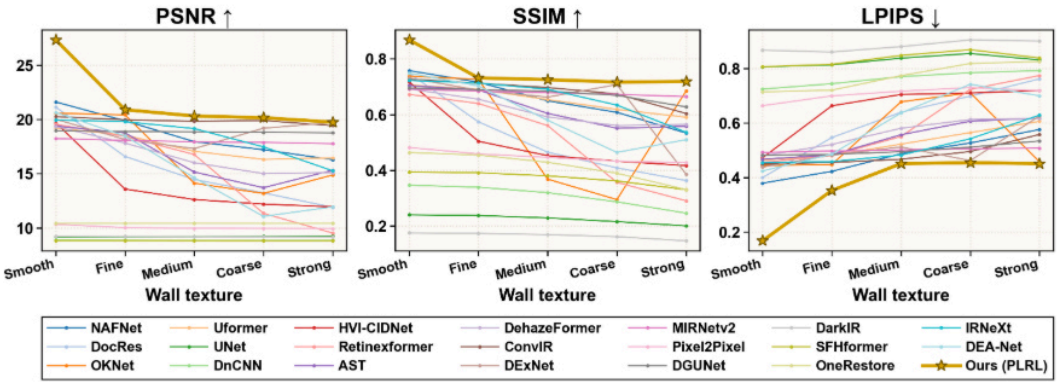


Figure 24: Screen-2000 PSNR, SSIM, and LPIPS curves across wall-texture conditions.

The perceptual curve shows how local appearance can change while structural similarity remains comparatively strong. In Figure 24, LPIPS rises from 0.353 under fine texture to 0.451 under medium texture, then stays within 0.451-0.455 through the strong condition. Together with the SSIM range, this indicates that regional structure and finer appearance respond differently to the surface pattern. The separation is directly relevant to PLRL's design, which establishes the screen structure before calibrating its detail. Figure 25 shows these two aspects of the reconstruction within a dark interface.

The large panel divisions of the dark interface remain evident across the five wall textures in Figure 25. Near the smooth condition, PLRL preserves the visible panel boundaries and colored accents. As texture grows stronger, the broad arrangement persists while additional local variation appears within the output. The panel thus shows both the recovered screen organization and the increasing difficulty of refining its internal appearance from wall light containing surface structure.

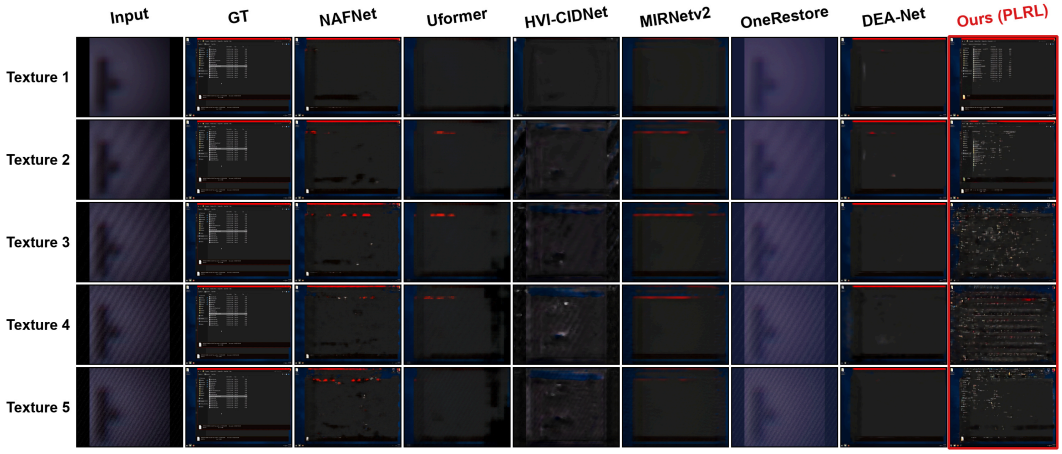


Figure 25: Representative Screen-2000 reconstructions under the five wall-texture conditions.

The retained panel arrangement in Figure 25 explains why screen structure and local detail need distinct treatment. The positions of the main panels and boundaries establish the interface's recognizable organization; its small internal features depend on more weakly preserved evidence. PPRL provides a spatial foundation for those regions, and LPCR regulates the subsequent refinement through compatibility and confidence. This relation connects the visual example to the higher structural similarity in Table 11. The next surface experiment varies roughness, which changes the wall response through which the screen light is reflected.

A.11 Wall-Roughness Robustness

Wall roughness changes the reflected evidence that the camera uses to recover the hidden screen. We vary roughness from 0.1 to 0.5 while fixing the Screen-2000 scene, model, and evaluation setting.

This extends the surface analysis from visible texture to changes in the wall response associated with roughness. The test follows the same question through a different physical variation: whether the light that remains after reflection still supports the screen's main regions and contours.

PLRL retains the leading measurements at all five roughness levels in Table 12. It achieves the best values in all 15 condition-metric comparisons, beginning with 26.728 dB PSNR, 0.858 SSIM, and 0.177 LPIPS at roughness 0.1. At 0.5, it reaches 22.576 dB, 0.762, and 0.286. The strongest competing endpoint values are 17.957 dB PSNR and 0.480 LPIPS from NAFNet and 0.642 SSIM from DEA-Net. The advantage therefore extends through the complete tested change in surface response.

Table 12: Screen-2000 results under wall-roughness levels of 0.1, 0.2, 0.3, 0.4, and 0.5.

Method	Source	0.1			0.2			0.3			0.4			0.5		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.425	0.463	0.720	10.428	0.464	0.724	10.420	0.464	0.727	10.404	0.464	0.728	10.382	0.463	0.729
DocRes	CVPR'24	15.755	0.558	0.551	14.818	0.526	0.596	13.265	0.434	0.690	10.690	0.366	0.760	9.748	0.339	0.787
Pixel2Pixel	TPAMI'25	9.912	0.460	0.680	9.861	0.457	0.686	9.800	0.455	0.690	9.731	0.451	0.694	9.648	0.438	0.699
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	**21.354**	**0.751**	**0.384**	**21.028**	**0.740**	**0.393**	**20.311**	**0.720**	**0.411**	**19.247**	**0.688**	**0.439**	**17.957**	0.636	**0.480**
DGUNet	CVPR'22	18.863	0.671	0.485	18.609	0.641	0.491	18.254	0.621	0.502	17.624	0.603	0.520	16.763	0.582	0.545
Uformer	CVPR'22	18.285	0.692	0.483	17.037	0.672	0.506	16.325	0.653	0.526	15.362	0.621	0.553	13.881	0.565	0.592
MIRNetv2	TPAMI'22	18.181	0.672	0.496	18.022	0.640	0.501	17.481	0.613	0.518	16.167	0.582	0.555	14.377	0.533	0.618
OKNet	AAAI'24	20.274	0.712	0.448	19.803	0.675	0.457	19.208	0.648	0.471	18.511	0.625	0.488	17.658	0.605	0.511
AST	CVPR'24	18.647	0.685	0.486	16.558	0.630	0.539	13.702	0.556	0.605	12.439	0.512	0.640	11.719	0.481	0.666
ConvIR	TPAMI'24	19.809	0.666	0.467	18.249	0.616	0.500	16.832	0.582	0.534	15.772	0.554	0.568	14.874	0.533	0.601
DnCNN	TIP'17	8.728	0.355	0.737	8.631	0.345	0.748	8.483	0.342	0.755	8.353	0.350	0.754	8.242	0.348	0.759
UNet	MICCAI'15	9.249	0.242	0.814	9.297	0.244	0.815	9.328	0.246	0.816	9.345	0.247	0.824	9.353	0.247	0.831
Models for Related Degradations																
DExNet	TPAMI'25	19.609	0.713	0.452	18.503	0.682	0.481	17.301	0.636	0.519	15.984	0.582	0.566	14.577	0.526	0.618
HVI-CIDNet	CVPR'25	11.832	0.432	0.718	11.237	0.408	0.749	10.848	0.388	0.775	10.507	0.376	0.790	10.436	0.372	0.798
Retinexformer	ICCV'23	14.702	0.517	0.576	9.354	0.334	0.739	8.752	0.315	0.765	8.503	0.300	0.776	8.280	0.288	0.782
DehazeFormer	TIP'23	18.773	0.675	0.493	17.731	0.638	0.527	16.242	0.596	0.574	14.555	0.551	0.619	13.053	0.504	0.660
DEA-Net	TIP'24	20.507	0.745	0.431	19.981	0.730	0.447	18.973	0.705	0.475	17.818	0.675	0.510	16.654	**0.642**	0.550
IRNeXt	ICML'23	19.735	0.672	0.466	18.824	0.628	0.491	17.494	0.598	0.521	16.322	0.570	0.555	15.551	0.554	0.581
SFHformer	ECCV'24	8.813	0.397	0.810	8.809	0.400	0.812	8.804	0.403	0.814	8.798	0.406	0.815	8.790	0.409	0.815
DarkIR	CVPR'25	9.188	0.163	0.877	9.153	0.143	0.889	9.116	0.135	0.900	9.078	0.131	0.911	9.038	0.128	0.920
PLRL	**Ours**	**26.728**	**0.858**	**0.177**	**26.055**	**0.846**	**0.190**	**24.985**	**0.825**	**0.211**	**23.789**	**0.798**	**0.245**	**22.576**	**0.762**	**0.286**

The roughness comparison shows that PLRL continues to recover useful screen structure as the reflected wall evidence changes. Such changes make it useful to combine broad spatial support with controlled local refinement. PPRL supplies bounded updates to layout and structure, and LPCR weights detail according to the observation and recovery state. Their combined reconstruction retains higher pixel fidelity and structural similarity together with lower perceptual distance in Table 12. The curves identify how these advantages develop as roughness increases through the intermediate settings.

PLRL leads all three metrics across roughness levels from 0.1 to 0.5. At the highest level, it retains 22.576 dB PSNR, 0.762 SSIM, and 0.286 LPIPS. Figure 27 shows what this retained agreement represents: the central window remains positioned within its background, with its extent and broad internal organization preserved.

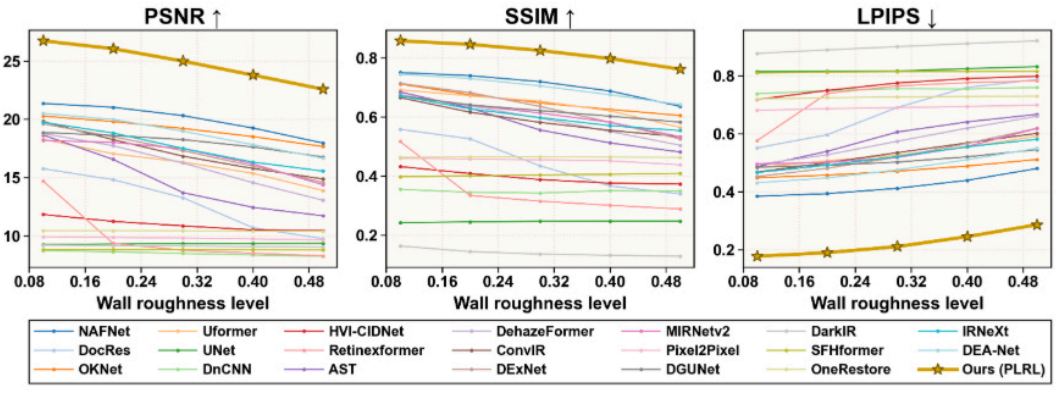


Figure 26: Screen-2000 PSNR, SSIM, and LPIPS curves across wall-roughness levels.

The joint curve movement connects roughness to both broad fidelity and local appearance. The increasing errors are consistent with changes in the wall response that make spatial evidence less precise, while the retained separation in Figure 26 shows that PLRL continues to use the available

information effectively. This behavior matches the motivation for a structural estimate followed by controlled detail: the reconstruction can preserve major screen organization while refinement responds to the quality of the reflected evidence. The next panel makes that organization concrete through a prominent central window.

The central light window remains distinct against its green background across all five roughness levels in Figure 27. PLRL preserves the window's position, rectangular extent, and broad internal organization. Several compared outputs reduce it to a bright block or merge parts of its boundary with the surrounding background as roughness increases. This contrast makes the recovered spatial relationships visible: the reconstruction identifies both the major foreground element and the background region in which it sits.

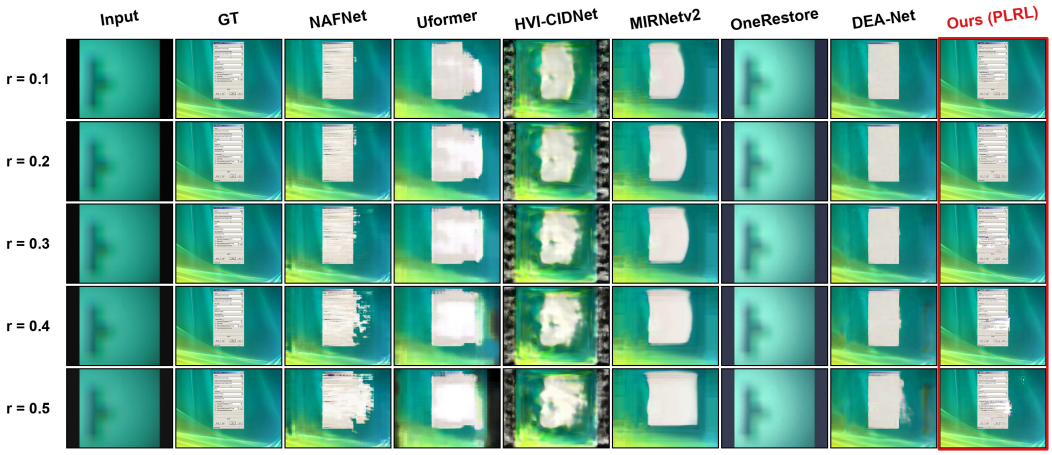


Figure 27: Representative Screen-2000 reconstructions under wall-roughness levels of 0.1, 0.2, 0.3, 0.4, and 0.5.

The persistent window structure in Figure 27 supports the feasibility of recovering principal screen content under the tested surface changes. The observed wall light continues to reveal a major screen element and its placement even as the wall response varies. Establishing structure before detail provides the method's rationale for maintaining that organization. We next add random corruption to the observation itself, testing whether the same recovery priority remains useful when the screen-related signal is mixed with noise across several content domains.

A.12 Gaussian-Noise Robustness

Gaussian noise makes the screen-related wall signal harder to distinguish from random intensity variation. We evaluate OS-Lock, Screen-2000, and Screen-Game at SNRs of 20, 25, 30, 35, and 40 dB using each dataset's fixed model and evaluation setting. Lower SNR corresponds to stronger corruption, so the sweep follows recovery as the observation becomes progressively cleaner. The three domains test whether PLRL's advantage extends across different kinds of screen structure under the same form of added disturbance.

OS-Lock. PLRL leads all 15 Gaussian-noise condition-metric comparisons on OS-Lock in Table 13. At 20 dB SNR, it reaches 16.990 dB PSNR, 0.608 SSIM, and 0.631 LPIPS. At 40 dB SNR, the corresponding values are 17.087 dB, 0.623, and 0.627. Across these endpoints, PSNR changes by 0.097 dB and LPIPS by 0.004. The small changes in these two metrics show that PLRL retains a similar level of pixel fidelity and perceptual appearance throughout the tested noise range.

The OS-Lock results show a stable coarse appearance together with a structural benefit from cleaner observations. In Table 13, the rise in SSIM as SNR increases indicates that local structure still improves as random variation weakens. This combination matches the purpose of PLRL's two recovery components: the structural estimate provides a continuing layout, while the confidence-weighted residual refines it according to the available evidence. The measured stability therefore has a task-specific meaning: the wall pattern continues to support broad screen appearance while the precision of recovered structure responds to observation quality.

The OS-Lock curves locate most of the structural change in the strongest-noise part of the sweep. Figure 28 shows nearly level PSNR and LPIPS trajectories, while SSIM rises from 0.608 to 0.618

Table 13: OS-Lock results under additive Gaussian noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	12.247	0.509	0.727	12.242	0.517	0.721	12.240	0.520	0.720	12.239	0.521	0.721	12.238	0.522	0.721
DocRes	CVPR'24	15.925	0.593	0.686	16.089	0.601	0.673	16.138	0.606	0.666	16.103	0.607	0.664	16.067	0.607	0.663
Pixel2Pixel	TPAMI'25	12.398	0.515	0.737	12.400	0.516	0.734	12.400	0.517	0.732	12.400	0.517	0.731	12.400	0.517	0.731
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	15.217	0.543	0.677	15.232	0.555	0.668	15.238	0.559	0.665	15.240	0.561	0.664	15.241	0.561	0.664
DGUNet	CVPR'22	**16.491**	0.605	0.649	**16.520**	**0.616**	0.649	**16.534**	**0.619**	0.650	**16.540**	**0.621**	**0.651**	**16.543**	**0.621**	**0.651**
Uformer	CVPR'22	15.146	0.554	0.681	15.149	0.561	0.676	15.151	0.563	0.675	15.150	0.564	0.675	15.149	0.564	0.674
MIRNetv2	TPAMI'22	15.264	0.600	0.654	15.249	0.603	0.655	15.235	0.604	0.656	15.225	0.604	0.657	15.220	0.604	0.657
OKNet	AAAI'24	14.644	0.538	0.685	14.652	0.549	0.680	14.658	0.553	0.679	14.660	0.554	0.678	14.662	0.554	0.678
AST	CVPR'24	15.115	0.520	0.701	15.096	0.530	0.695	15.085	0.534	0.693	15.079	0.535	0.692	15.076	0.536	0.692
ConvIR	TPAMI'24	14.941	0.575	0.687	14.930	0.586	0.685	14.922	0.590	0.686	14.917	0.591	0.686	14.914	0.592	0.686
DnCNN	TIP'17	11.741	0.465	0.728	11.732	0.472	0.723	11.740	0.474	0.723	11.759	0.475	0.723	11.777	0.476	0.723
UNet	MICCAI'15	12.109	0.475	0.770	12.123	0.483	0.763	12.146	0.485	0.761	12.168	0.487	0.761	12.177	0.487	0.762
Models for Related Degradations																
DExNet	TPAMI'25	15.404	**0.607**	**0.647**	15.456	0.608	**0.649**	15.492	0.608	**0.650**	15.508	0.608	0.651	15.513	0.608	0.652
HVI-CIDNet	CVPR'25	13.716	0.394	0.846	13.898	0.431	0.812	14.055	0.460	0.789	14.217	0.479	0.773	14.341	0.493	0.760
Retinexformer	ICCV'23	14.594	0.523	0.677	14.616	0.546	0.662	14.626	0.555	0.655	14.627	0.558	0.653	14.622	0.559	0.652
DehazeFormer	TIP'23	14.851	0.524	0.677	14.866	0.533	0.671	14.871	0.536	0.669	14.872	0.537	0.669	14.873	0.537	0.669
DEA-Net	TIP'24	15.942	0.605	0.673	15.958	0.606	0.674	15.959	0.606	0.674	15.961	0.606	0.674	15.961	0.606	0.674
IRNeXt	ICML'23	16.135	0.596	0.676	16.133	0.603	0.676	16.134	0.606	0.677	16.134	0.607	0.678	16.134	0.607	0.678
SFHformer	ECCV'24	12.455	0.504	0.739	12.458	0.507	0.738	12.459	0.508	0.739	12.460	0.509	0.740	12.460	0.509	0.740
DarkIR	CVPR'25	8.790	0.096	0.811	8.788	0.095	0.807	8.788	0.095	0.805	8.788	0.095	0.804	8.787	0.095	0.804
PLRL	**Ours**	**16.990**	**0.608**	**0.631**	**17.057**	**0.618**	**0.627**	**17.073**	**0.621**	**0.627**	**17.082**	**0.622**	**0.627**	**17.087**	**0.623**	**0.627**

between 20 and 25 dB SNR. Its subsequent increase to 0.623 at 40 dB is smaller. This shape shows that reducing strong additive noise initially benefits structure most, after which the remaining change in the recovered image becomes modest.

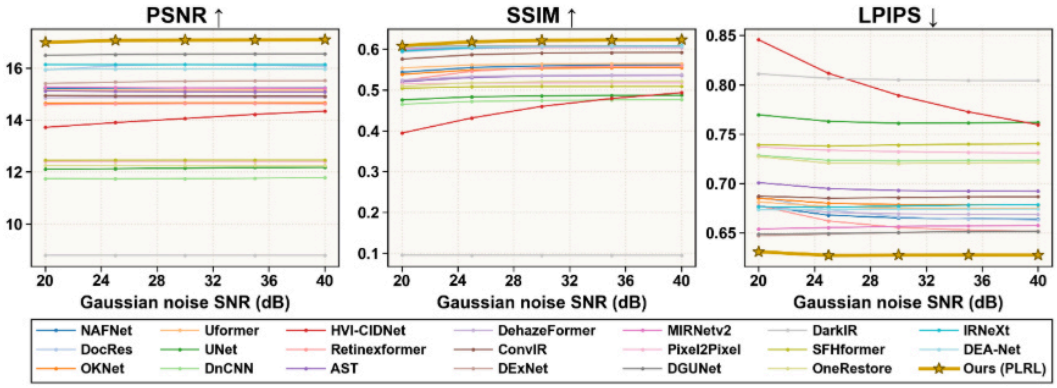


Figure 28: OS-Lock PSNR, SSIM, and LPIPS curves across Gaussian-noise SNRs.

PLRL's lowest LPIPS throughout Figure 28 connects noise tolerance to the appearance of the reconstructed display. The perceptual separation persists even where several methods show similar PSNR trends, indicating that comparable pixel-level trends can accompany different agreement in screen appearance. For the present task, that appearance is carried chiefly by the large regions and major contours preserved after diffuse reflection. Figure 29 gives a direct view of these coarse cues in a landscape lock screen and shows the level of visual organization represented by the stable aggregate measurements.

The landscape example in Figure 29 retains broad color regions across the Gaussian-noise levels. Starting from a dark wall-light observation, PLRL produces a blue upper region and warmer lower regions across the displayed 20-to-40-dB rows, recovering the target's broad color organization. Fine landscape structure remains diffuse, so the visible evidence concerns regional appearance.

The repeated regional pattern in Figure 29 connects the small endpoint changes in Table 13 to a visible screen-level outcome: the wall observation continues to reveal the display's main color and brightness organization. This persistent regional structure gives subsequent detail refinement an established layout to build on. The next domain examines how the same recovery process responds to Gaussian noise in more varied screen content.

Screen-2000. PLRL leads all 15 Gaussian-noise comparisons on Screen-2000 in Table 14. At 20 dB SNR, it obtains 23.476 dB PSNR, 0.747 SSIM, and 0.266 LPIPS; DExNet supplies the strongest competing values of 19.946 dB, 0.724, and 0.441. At 40 dB SNR, PLRL reaches 27.267 dB, 0.867,

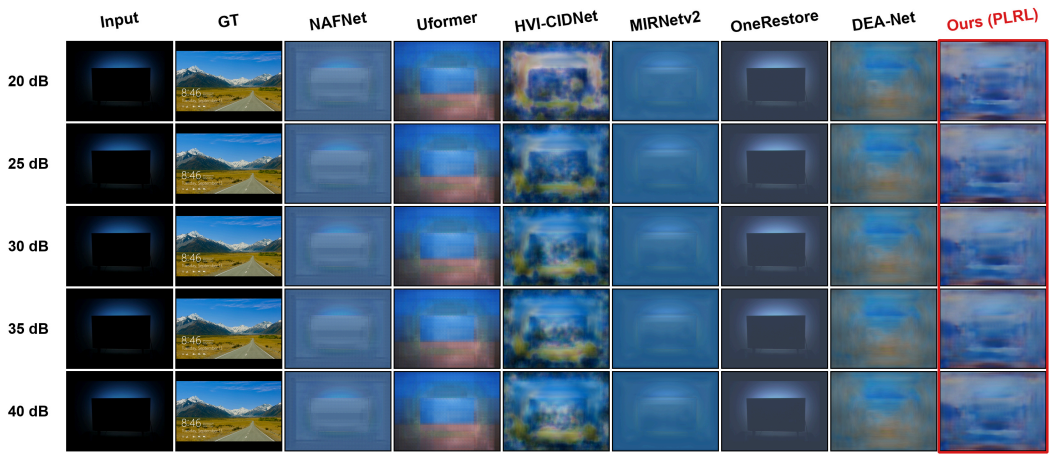


Figure 29: Representative OS-Lock reconstructions under Gaussian-noise SNRs of 20, 25, 30, 35, and 40 dB.

and 0.169. Its advantage at the strongest tested noise level shows that useful screen recovery already persists before the observation becomes cleaner.

Table 14: Screen-2000 results under additive Gaussian noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.438	0.344	0.801	10.427	0.412	0.750	10.420	0.445	0.723	10.415	0.457	0.716	10.413	0.462	0.715
DocRes	CVPR'24	15.543	0.526	0.588	16.626	0.588	0.530	17.124	0.611	0.505	17.312	0.618	0.497	17.308	0.618	0.498
Pixel2Pixel	TPAMI'25	9.687	0.436	0.682	9.718	0.447	0.683	9.880	0.453	0.684	10.075	0.465	0.674	10.190	0.473	0.667
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	19.383	0.604	0.453	**20.822**	0.702	**0.401**	**21.291**	0.738	**0.386**	**21.467**	**0.752**	**0.381**	**21.588**	**0.756**	**0.380**
DGUNet	CVPR'22	18.760	0.582	0.511	18.900	0.652	0.488	18.939	0.681	0.482	18.956	0.690	0.481	18.960	0.694	0.482
Uformer	CVPR'22	18.820	0.586	0.526	19.083	0.662	0.482	19.143	0.692	0.468	19.186	0.702	0.465	19.205	0.705	0.464
MIRNetv2	TPAMI'22	18.039	0.675	0.493	18.124	0.684	0.492	18.183	0.687	0.493	18.222	0.689	0.492	18.232	0.690	0.492
OKNet	AAAI'24	18.749	0.598	0.481	20.116	0.688	0.449	20.422	0.721	0.445	20.527	0.732	0.445	20.570	0.736	0.445
AST	CVPR'24	18.891	0.572	0.522	19.271	0.653	0.481	19.358	0.687	0.470	19.384	0.698	0.467	19.390	0.702	0.467
ConvIR	TPAMI'24	19.154	0.557	0.515	20.088	0.667	0.457	20.235	0.708	0.448	20.263	0.721	0.448	20.269	0.725	0.449
DnCNN	TIP'17	8.815	0.263	0.769	8.881	0.312	0.739	8.887	0.334	0.726	8.887	0.342	0.725	8.885	0.345	0.725
UNet	MICCAI'15	9.292	0.214	0.826	9.218	0.228	0.810	9.190	0.235	0.805	9.183	0.238	0.805	9.181	0.239	0.806
Models for Related Degradations																
DExNet	TPAMI'25	**19.946**	**0.724**	**0.441**	19.995	0.729	0.439	20.014	0.731	0.438	20.017	0.731	0.438	20.018	0.731	0.438
HVI-CIDNet	CVPR'25	12.032	0.348	0.759	12.112	0.403	0.719	12.158	0.430	0.705	12.247	0.443	0.700	12.298	0.452	0.695
Retinexformer	ICCV'23	15.514	0.505	0.570	18.414	0.617	0.489	19.019	0.646	0.469	19.346	0.660	0.460	19.485	0.667	0.455
DehazeFormer	TIP'23	18.803	0.664	0.496	19.248	0.684	0.483	19.316	0.690	0.480	19.351	0.692	0.480	19.351	0.692	0.480
DEA-Net	TIP'24	19.170	0.720	0.464	20.179	**0.740**	0.436	20.525	**0.747**	0.427	20.606	0.749	0.424	20.639	0.749	0.424
IRNeXt	ICML'23	16.559	0.510	0.581	18.752	0.648	0.483	19.835	0.702	0.455	19.941	0.716	0.453	19.954	0.721	0.454
SFHformer	ECCV'24	8.801	0.344	0.843	8.809	0.376	0.823	8.813	0.388	0.813	8.814	0.392	0.809	8.816	0.393	0.807
DarkIR	CVPR'25	9.216	0.154	0.922	9.220	0.168	0.892	9.221	0.173	0.876	9.221	0.174	0.870	9.221	0.175	0.868
PLRL	**Ours**	**23.476**	**0.747**	**0.266**	**25.337**	**0.820**	**0.207**	**26.529**	**0.853**	**0.181**	**27.063**	**0.863**	**0.172**	**27.267**	**0.867**	**0.169**

Cleaner wall evidence produces a larger reconstruction gain on Screen-2000 than on OS-Lock. Across the endpoints in Table 14, PLRL gains 3.791 dB PSNR and 0.120 SSIM. The same type of perturbation therefore interacts differently with the content and observation distributions of the two domains. The joint gain in fidelity and structure is consistent with using improved evidence to strengthen the progressive estimate and its refinement. Figure 30 shows where along the SNR sweep this improvement is concentrated.

Most of the Screen-2000 improvement occurs as the strongest Gaussian corruption is reduced. In Figure 30, PLRL rises from 23.476 to 25.337 dB PSNR between 20 and 25 dB SNR, then reaches 26.529 dB at 30 dB. Between 35 and 40 dB SNR, the PSNR gain is 0.204 dB and LPIPS improves from 0.172 to 0.169. The decreasing increments show that added noise has a progressively smaller effect as the observation becomes cleaner.

The flattening trajectories in Figure 30 connect the available evidence to the amount of additional recovery. PLRL's scale-based reconstruction and confidence weighting operate across the full sweep, with cleaner observations producing closer structural and perceptual agreement. At higher SNR, the smaller metric changes indicate that further reductions in the added noise yield smaller gains. The selected natural-image screen in Figure 31 shows how this recovery appears for textured content with overlapping desktop elements.

The natural-image screen in Figure 31 shows coarse color variation emerging from a diffuse wall observation. Across the displayed SNRs, PLRL recovers warm and cool regions related to the target,

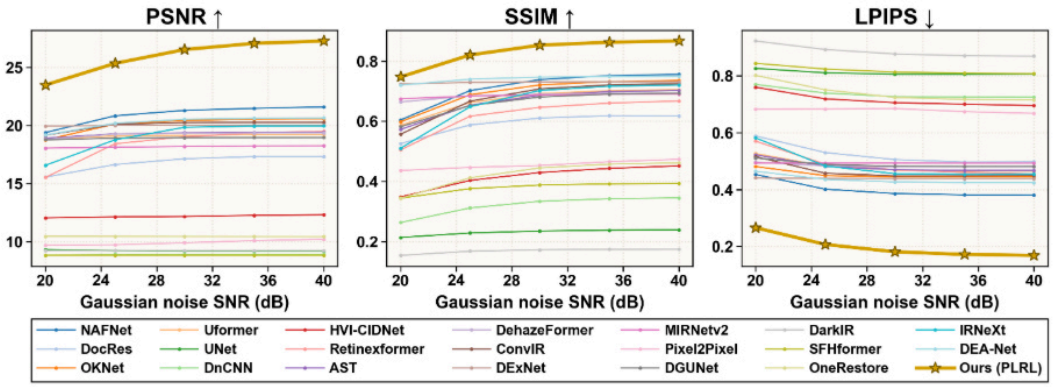


Figure 30: Screen-2000 PSNR, SSIM, and LPIPS curves across Gaussian-noise SNRs.

which combines a textured image with desktop elements. This example shows recovery of broad color and brightness organization from wall light produced by richly textured screen content.

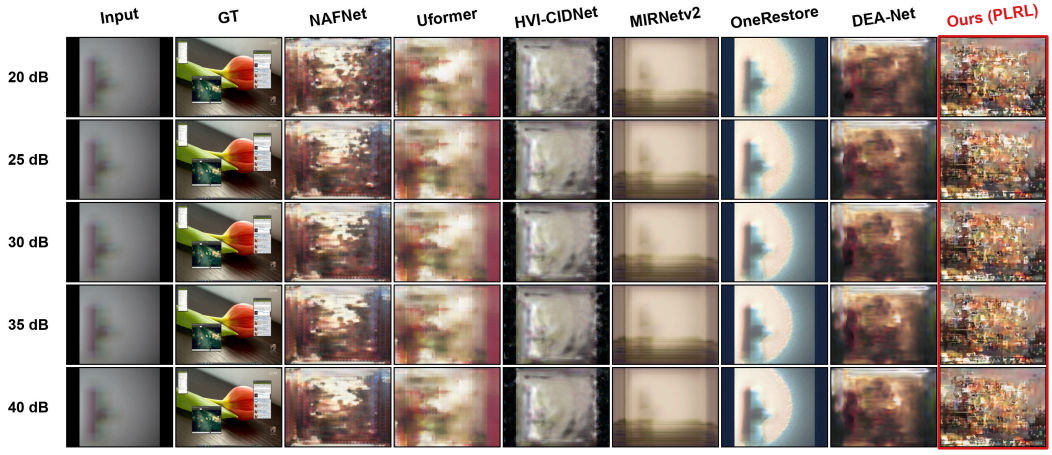


Figure 31: Representative Screen-2000 reconstructions under Gaussian-noise SNRs of 20, 25, 30, 35, and 40 dB.

Broad spatial appearance provides the clearest task-relevant evidence in Figure 31. The reconstructed color and brightness distribution conveys partial information about the display, while small object boundaries remain difficult to resolve. Table 14 establishes PLRL's relative advantage across the full dataset, and the selected panel shows the form that coarse recovery takes in this demanding example. Together, they connect aggregate improvement with the paper's intended level of reconstruction. The Screen-Game comparison next examines the same noise sweep in a domain with a more stable perceptual response.

Screen-Game. PLRL achieves the strongest values in all 15 Gaussian-noise comparisons on Screen-Game in Table 15. At 20 dB SNR, it reaches 18.326 dB PSNR, 0.652 SSIM, and 0.474 LPIPS. The strongest competing values are 18.093 dB PSNR and 0.571 LPIPS from IRNeXt and 0.649 SSIM from MIRNetv2. At 40 dB SNR, PLRL reaches 18.374 dB, 0.674, and 0.474. Its advantage therefore extends across the full tested noise range in this third content domain.

The most persistent Screen-Game advantage concerns perceptual appearance. Table 15 places PLRL's LPIPS between 0.473 and 0.474 throughout the sweep, compared with strongest competing values between 0.571 and 0.574. PSNR changes little, while SSIM improves as the observation becomes cleaner. This combination shows that a similar broad appearance can be retained while local structural agreement benefits from weaker noise. It also matches the method's separation between an established screen estimate and detail whose contribution depends on local confidence.

The Screen-Game curves show structural improvement alongside an almost unchanged perceptual distance. In Figure 32, PLRL's SSIM rises from 0.652 at 20 dB SNR to 0.671 at 30 dB and 0.674

Table 15: Screen-Game results under additive Gaussian noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.248	0.469	0.752	10.286	0.492	0.746	10.303	0.501	0.745	10.309	0.504	0.746	10.312	0.505	0.747
DocRes	CVPR'24	8.065	0.455	0.849	8.193	0.449	0.823	8.525	0.442	0.786	9.201	0.439	0.748	10.398	0.454	0.713
Pixel2Pixel	TPAMI'25	11.316	0.319	0.779	9.946	0.336	0.758	9.700	0.360	0.748	10.233	0.373	0.749	10.762	0.398	0.737
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	17.577	0.609	0.597	17.982	0.641	0.582	18.117	0.653	0.578	**18.160**	0.657	0.576	**18.156**	0.657	0.576
DGUNet	CVPR'22	17.852	0.645	0.594	17.871	0.654	0.594	17.880	0.657	0.595	17.886	0.658	0.595	17.886	0.659	0.595
Uformer	CVPR'22	17.520	0.601	0.630	17.538	0.630	0.617	17.543	0.641	0.614	17.547	0.645	0.614	17.547	0.646	0.614
MIRNetv2	TPAMI'22	17.607	**0.649**	0.616	17.662	0.655	0.614	17.671	0.655	0.613	17.673	0.655	0.613	17.677	0.655	0.612
OKNet	AAAI'24	17.857	0.645	0.574	17.898	0.656	0.574	17.901	0.660	0.575	17.893	0.661	0.575	17.893	0.662	0.576
AST	CVPR'24	17.416	0.599	0.632	17.447	0.628	0.618	17.450	0.639	0.616	17.453	0.643	0.616	17.456	0.644	0.616
ConvIR	TPAMI'24	17.597	0.608	0.603	17.626	0.635	0.600	17.638	0.646	0.600	17.640	0.649	0.601	17.640	0.650	0.601
DnCNN	TIP'17	7.347	0.317	0.769	8.329	0.354	0.760	8.753	0.416	0.728	8.745	0.425	0.726	8.741	0.427	0.726
UNet	MICCAI'15	10.060	0.442	0.772	10.303	0.462	0.760	10.306	0.469	0.757	10.301	0.471	0.757	10.296	0.471	0.758
Models for Related Degradations																
DExNet	TPAMI'25	16.871	0.631	0.619	17.510	0.651	0.600	17.620	0.656	0.596	17.633	0.657	0.596	17.633	0.658	0.595
HVI-CIDNet	CVPR'25	12.635	0.427	0.754	12.617	0.466	0.743	12.692	0.483	0.745	12.620	0.488	0.748	12.551	0.490	0.748
Retinexformer	ICCV'23	15.391	0.578	0.648	17.239	0.622	0.616	17.897	0.649	0.598	17.924	0.653	0.594	17.925	0.654	0.593
DehazeFormer	TIP'23	17.748	0.643	0.608	17.764	0.647	0.607	17.773	0.648	0.607	17.772	0.648	0.607	17.773	0.649	0.607
DEA-Net	TIP'24	16.962	0.638	0.629	17.628	0.655	0.599	17.683	0.656	0.591	17.714	0.657	0.589	17.722	0.657	0.588
IRNeXt	ICML'23	**18.093**	0.645	**0.571**	**18.129**	**0.661**	**0.571**	**18.141**	**0.665**	**0.573**	18.140	**0.667**	**0.573**	18.134	**0.667**	**0.574**
SFHformer	ECCV'24	9.994	0.447	0.779	10.004	0.457	0.763	10.008	0.460	0.759	10.009	0.461	0.759	10.010	0.461	0.759
DarkIR	CVPR'25	4.471	0.040	0.957	4.467	0.041	0.943	4.465	0.041	0.939	4.465	0.041	0.937	4.465	0.041	0.936
PLRL	**Ours**	**18.326**	**0.652**	**0.474**	**18.342**	**0.665**	**0.473**	**18.376**	**0.671**	**0.474**	**18.379**	**0.673**	**0.473**	**18.374**	**0.674**	**0.474**

at 40 dB, while LPIPS remains nearly level. Most of the structural response is concentrated in the lower-SNR conditions. As in the other domains, the curve shape makes the influence of observation quality explicit while retaining PLRL's relative reconstruction advantage.

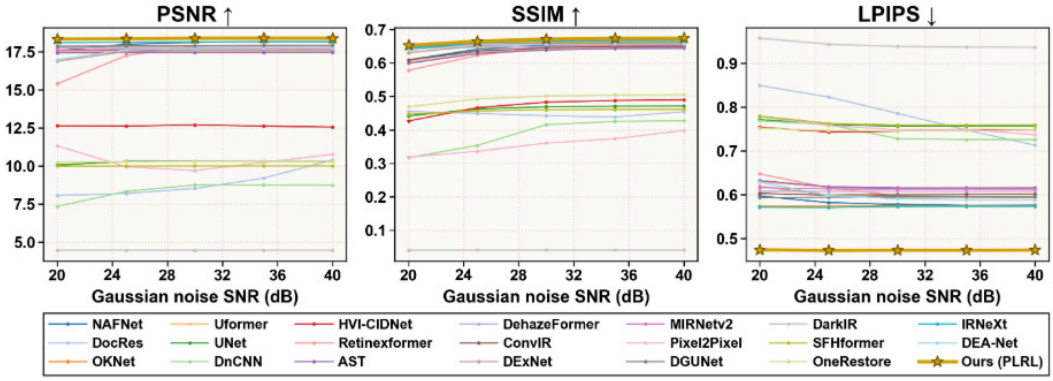


Figure 32: Screen-Game PSNR, SSIM, and LPIPS curves across Gaussian-noise SNRs.

The different sensitivities of SSIM and LPIPS in Figure 32 help explain what noise robustness means for this task. PLRL maintains a similar perceptual appearance across the tested noise levels, while cleaner evidence improves agreement in local structure. These metrics therefore connect two levels of recovery that the model handles separately: large-scale organization and its refinement. The interface in Figure 33 makes the large spatial relationships visible, allowing the reader to inspect the screen regions that remain stable across the SNR sweep.

The interface in Figure 33 retains its principal spatial divisions across all five Gaussian-noise levels. PLRL preserves the pale workspace, the narrow panel on the right, and the lower colored strip in consistent positions. Small items inside the workspace remain simplified. The workspace, right-side panel, and lower strip retain their relative positions, revealing the interface's main layout and supplying a concrete visual counterpart to the stable LPIPS and improving SSIM in Table 15 and Figure 32.

The comparative columns in Figure 33 show the benefit of preserving regional relationships during reconstruction. Several methods produce nearly uniform central regions or strong local artifacts, while PLRL maintains a cleaner distinction between the workspace and right-side controls. This example connects the three Gaussian-noise suites to the same central story: useful coarse screen content survives in the wall observation, and detail receives a separate confidence-dependent contribution. We next test abrupt, isolated intensity disturbances through salt-and-pepper corruption, which challenges the interpretation of local variation in a different way.

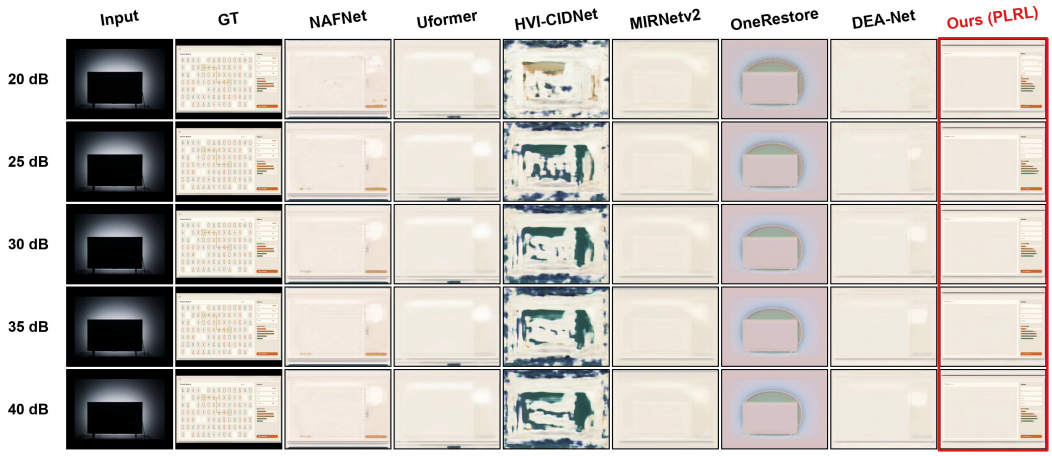


Figure 33: Representative Screen-Game reconstructions under Gaussian-noise SNRs of 20, 25, 30, 35, and 40 dB.

A.13 Salt-and-Pepper-Noise Robustness

Salt-and-pepper corruption introduces abrupt local intensity changes that can compete with weak screen detail in the wall image. We evaluate OS-Lock, Screen-2000, and Screen-Game at SNRs of 20, 25, 30, 35, and 40 dB, retaining each dataset's model and settings. This complements Gaussian noise by testing a different form of disturbance within the same recovery problem. The key question remains whether broad screen structure provides enough support for a useful reconstruction while local corrections are calibrated from corrupted evidence.

OS-Lock. PLRL leads PSNR and LPIPS at every tested SNR in Table 16 and achieves the highest displayed SSIM or a tie at each condition. At 20 dB SNR, it obtains 17.007 dB PSNR, 0.619 SSIM, and 0.630 LPIPS. At 40 dB SNR, the values are 17.088 dB, 0.623, and 0.627. The strongest competing PSNR and LPIPS at 40 dB are 16.547 dB from DGUNet and 0.647 from DocRes. The measured advantage therefore persists as the impulsive perturbation weakens.

Table 16: OS-Lock results under additive salt-and-pepper noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	12.238	0.521	0.721	12.237	0.522	0.721	12.237	0.522	0.721	12.237	0.522	0.721	12.237	0.522	0.721
DocRes	CVPR'24	16.339	0.607	**0.648**	16.337	0.608	**0.647**	16.357	0.609	**0.647**	16.355	0.609	**0.647**	16.357	0.609	**0.647**
Pixel2Pixel	TPAMI'25	12.406	0.516	0.730	12.404	0.517	0.730	12.405	0.517	0.730	12.404	0.517	0.730	12.405	0.517	0.730
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	15.234	0.560	0.664	15.240	0.561	0.664	15.241	0.561	0.664	15.241	0.561	0.664	15.241	0.561	0.664
DGUNet	CVPR'22	**16.416**	**0.618**	0.654	**16.491**	**0.621**	0.652	**16.544**	**0.622**	0.652	**16.533**	**0.622**	0.652	**16.547**	**0.622**	0.652
Uformer	CVPR'22	15.146	0.563	0.675	15.149	0.564	0.674	15.151	0.564	0.675	15.145	0.564	0.675	15.149	0.565	0.674
MIRNetv2	TPAMI'22	15.209	0.603	0.658	15.210	0.603	0.657	15.211	0.603	0.658	15.211	0.603	0.658	15.211	0.604	0.658
OKNet	AAAI'24	14.664	0.554	0.679	14.662	0.554	0.678	14.663	0.555	0.678	14.664	0.555	0.678	14.663	0.555	0.678
AST	CVPR'24	15.058	0.535	0.692	15.069	0.536	0.692	15.070	0.536	0.692	15.072	0.536	0.692	15.071	0.536	0.692
ConvIR	TPAMI'24	14.884	0.590	0.687	14.903	0.591	0.687	14.912	0.592	0.687	14.901	0.591	0.687	14.910	0.592	0.687
DnCNN	TIP'17	11.820	0.477	0.722	11.818	0.478	0.722	11.819	0.478	0.722	11.819	0.478	0.722	11.819	0.478	0.722
UNet	MICCAI'15	12.205	0.485	0.762	12.184	0.486	0.762	12.181	0.487	0.762	12.177	0.486	0.762	12.179	0.487	0.762
Models for Related Degradations																
DExNet	TPAMI'25	15.499	0.606	0.653	15.515	0.607	0.652	15.517	0.607	0.652	15.516	0.607	0.652	15.518	0.608	0.652
HVI-CIDNet	CVPR'25	15.591	0.573	0.672	15.598	0.575	0.672	15.604	0.575	0.672	15.600	0.575	0.672	15.605	0.575	0.671
Retinexformer	ICCV'23	14.594	0.557	0.651	14.598	0.558	0.651	14.600	0.558	0.651	14.600	0.558	0.651	14.600	0.558	0.651
DehazeFormer	TIP'23	14.872	0.536	0.670	14.871	0.536	0.669	14.873	0.537	0.669	14.866	0.536	0.669	14.873	0.537	0.669
DEA-Net	TIP'24	15.790	0.601	0.678	15.862	0.603	0.676	15.961	0.606	0.674	15.942	0.605	0.675	15.962	0.606	0.674
IRNeXt	ICML'23	16.027	0.604	0.681	16.095	0.606	0.679	16.130	0.607	0.679	16.108	0.607	0.679	16.134	0.607	0.679
SFHformer	ECCV'24	12.459	0.508	0.740	12.460	0.509	0.740	12.461	0.509	0.740	12.460	0.509	0.741	12.461	0.509	0.740
DarkIR	CVPR'25	8.787	0.095	0.803	8.787	0.095	0.803	8.787	0.095	0.804	8.787	0.095	0.804	8.787	0.095	0.804
PLRL	**Ours**	**17.007**	**0.619**	**0.630**	**17.078**	**0.621**	**0.627**	**17.098**	**0.623**	**0.627**	**17.076**	**0.622**	**0.628**	**17.088**	**0.623**	**0.627**

The small changes in Table 16 show that coarse OS-Lock recovery remains stable over the tested corruption range. PPRL's scale organization provides a broad structural reference, while LPCR assesses detail through the relation between the observation and candidate. This design addresses the local ambiguity created by isolated intensity changes: each proposed refinement is considered together with the larger recovery state. The table records the resulting advantage in screen reconstruction, and Figure 34 shows how closely the individual PLRL measurements remain grouped across SNRs.

The OS-Lock curves vary within a narrow range under salt-and-pepper corruption. In Figure 34, PLRL's PSNR spans 17.007-17.098 dB, SSIM spans 0.619-0.623, and LPIPS spans 0.627-0.630. PSNR reaches its maximum at 30 dB SNR, so the measurements describe approximate stability across the sweep with small condition-specific changes. This pattern connects the table's endpoint comparison to the behavior of all five tested levels.

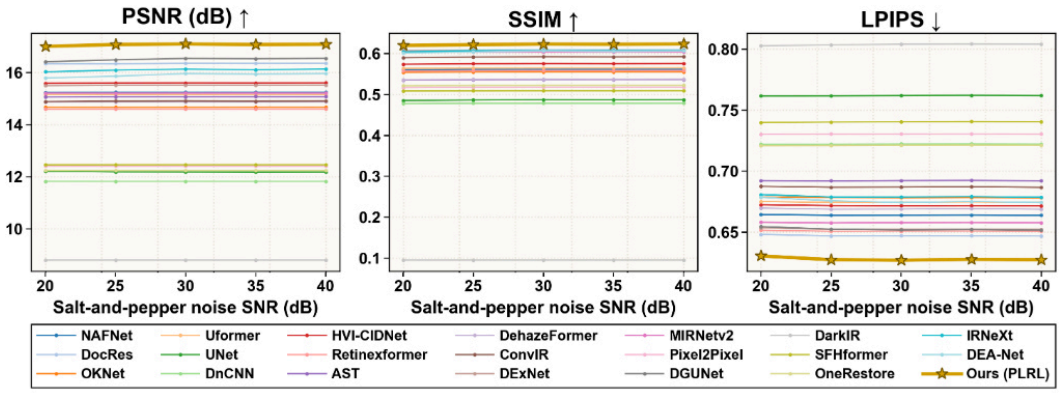


Figure 34: OS-Lock PSNR, SSIM, and LPIPS curves across salt-and-pepper-noise SNRs.

The maintained curve separation in Figure 34 shows that stable recovery also carries a comparative advantage. PLRL retains the lowest perceptual distance and remains at the top of the pixel and displayed structural comparisons. The agreement among these measures supports the persistence of useful coarse appearance under the evaluated impulsive disturbances. Figure 35 expands the visual inspection to several selected lock-screen contents, making the recovered color and brightness regions visible across different examples.

The selected lock screens in Figure 35 show broad color and brightness organization under the displayed corruption conditions. Across the examples, PLRL retains large blue, warm, or green regions related to parts of the corresponding targets, while local scene details remain diffuse. The rows contain different screen contents, so each reconstruction is read against its adjacent target. Across these selected contents, the reconstructed blue, warm, and green regions provide coarse visual information about the corresponding hidden displays.



Figure 35: Representative OS-Lock reconstructions under salt-and-pepper-noise SNRs of 20, 25, 30, 35, and 40 dB. Rows present selected screen contents at the indicated SNRs; each reconstruction is paired with the target shown in the same row.

Regional correspondence is the shared visual evidence across the different targets in Figure 35. The relative placement of large bright and dark areas and the overall color distribution provide direct cues to the hidden display's appearance. The selected rows show partial color correspondence, while Table 16 supplies the aggregate comparison across noise levels. This combination keeps the feasibility result concrete: wall observations can support coarse visual information across the

displayed contents. Screen-2000 next examines a setting where the strength of impulsive corruption produces a much larger change in recovery.

Screen-2000. PLRL leads all 15 salt-and-pepper condition-metric comparisons on Screen-2000 in Table 17. At 20 dB SNR, it reaches 20.605 dB PSNR, 0.708 SSIM, and 0.375 LPIPS. DExNet gives the strongest competing values of 19.451 dB, 0.673, and 0.459. At 40 dB SNR, PLRL reaches 27.278 dB, 0.868, and 0.169. Thus, even the strongest tested impulsive corruption leaves useful evidence for PLRL to reconstruct a closer screen estimate than the evaluated alternatives.

Table 17: Screen-2000 results under additive salt-and-pepper noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.391	0.452	0.712	10.396	0.459	0.713	10.407	0.462	0.714	10.409	0.463	0.714	10.409	0.464	0.714
DocRes	CVPR'24	16.746	0.640	0.490	19.368	0.712	**0.429**	20.534	**0.739**	0.408	20.985	0.746	0.402	21.094	0.748	0.401
Pixel2Pixel	TPAMI'25	10.315	0.466	0.669	10.322	0.476	0.665	10.323	0.480	0.664	10.325	0.481	0.664	10.325	0.481	0.664
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	16.957	0.577	0.534	19.609	0.696	0.432	**20.864**	0.737	**0.399**	**21.474**	**0.753**	**0.384**	**21.589**	**0.758**	**0.380**
DGUNet	CVPR'22	16.467	0.612	0.551	17.811	0.661	0.509	18.600	0.687	0.489	18.842	0.692	0.484	18.929	0.694	0.483
Uformer	CVPR'22	17.244	0.654	0.509	18.527	0.690	0.479	18.915	0.700	0.470	19.153	0.705	0.465	19.197	0.706	0.464
MIRNetv2	TPAMI'22	16.684	0.633	0.542	17.450	0.667	0.511	17.974	0.682	0.499	18.170	0.690	0.493	18.222	0.691	0.493
OKNet	AAAI'24	17.288	0.624	0.544	19.550	0.706	0.470	19.951	0.722	0.460	20.539	0.736	0.446	20.578	0.737	0.445
AST	CVPR'24	15.163	0.608	0.541	17.809	0.672	0.492	18.770	0.693	0.477	19.202	0.701	0.469	19.371	0.704	0.467
ConvIR	TPAMI'24	17.602	0.640	0.529	19.361	0.702	0.471	19.759	0.716	0.462	20.207	0.726	0.451	20.266	0.727	0.450
DnCNN	TIP'17	8.701	0.323	0.730	8.820	0.338	0.726	8.867	0.345	0.724	8.873	0.346	0.725	8.882	0.346	0.725
UNet	MICCAI'15	9.174	0.230	0.798	9.177	0.237	0.803	9.182	0.239	0.805	9.184	0.240	0.806	9.184	0.240	0.806
Models for Related Degradations																
DExNet	TPAMI'25	**19.451**	**0.673**	**0.459**	**19.839**	**0.724**	0.444	19.935	0.729	0.441	20.012	0.731	0.439	20.012	0.731	0.438
HVI-CIDNet	CVPR'25	17.933	0.653	0.520	19.107	0.696	0.487	19.478	0.711	0.478	19.636	0.716	0.474	19.683	0.717	0.473
Retinexformer	ICCV'23	15.157	0.534	0.556	17.754	0.624	0.491	19.114	0.663	0.463	19.425	0.669	0.458	19.542	0.672	0.456
DehazeFormer	TIP'23	16.156	0.597	0.563	18.058	0.655	0.509	18.869	0.679	0.491	19.227	0.689	0.482	19.353	0.692	0.480
DEA-Net	TIP'24	15.774	0.635	0.545	18.724	0.710	0.466	19.880	0.737	0.439	20.480	0.747	0.427	20.624	0.748	0.425
IRNeXt	ICML'23	16.611	0.615	0.553	18.806	0.690	0.482	19.381	0.709	0.468	19.833	0.720	0.457	19.955	0.722	0.454
SFHformer	ECCV'24	8.806	0.384	0.798	8.811	0.390	0.802	8.814	0.393	0.806	8.817	0.393	0.806	8.817	0.394	0.807
DarkIR	CVPR'25	9.203	0.173	0.835	9.213	0.174	0.856	9.219	0.175	0.865	9.220	0.175	0.866	9.220	0.175	0.867
PLRL	**Ours**	**20.605**	**0.708**	**0.375**	**24.367**	**0.808**	**0.235**	**26.231**	**0.852**	**0.187**	**27.026**	**0.861**	**0.175**	**27.278**	**0.868**	**0.169**

The Screen-2000 result shows that reducing impulsive corruption improves both screen fidelity and recovered organization. Between 20 and 40 dB SNR, PLRL gains 6.673 dB PSNR and 0.160 SSIM and lowers LPIPS by 0.206. These changes indicate that local corruption affects the structural content available for recovery as well as pixel accuracy. The persistent lead in Table 17 supports PLRL's ability to build an estimate from the remaining wall evidence. Figure 36 identifies where the largest benefit from cleaner observations appears.

The largest reconstruction gains occur as the strongest salt-and-pepper corruption is reduced. Figure 36 shows PLRL improving by 3.762 dB PSNR between 20 and 25 dB SNR, while LPIPS falls from 0.375 to 0.235. The changes become smaller toward 40 dB, where the curves approach the corresponding nominal Screen-2000 measurements. The sweep therefore traces a clear progression from heavily disturbed evidence to an observation whose added corruption has a smaller effect on screen recovery.

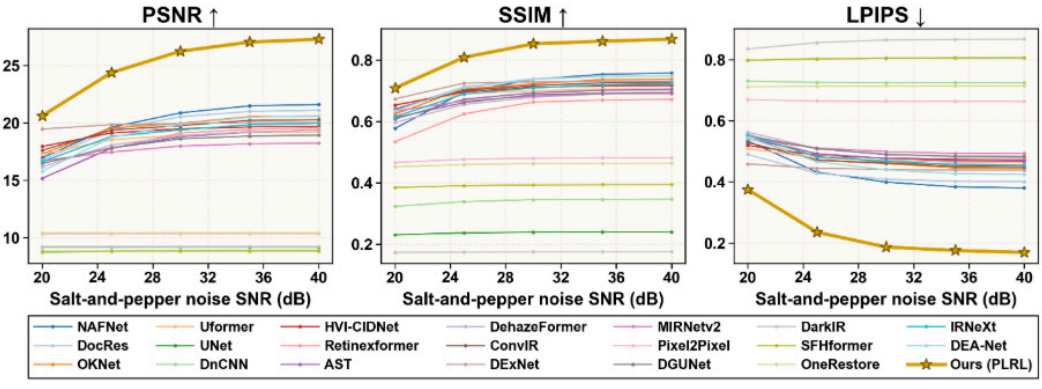


Figure 36: Screen-2000 PSNR, SSIM, and LPIPS curves across salt-and-pepper-noise SNRs.

The structural and perceptual curves in Figure 36 improve together as the wall evidence becomes cleaner. Their agreement connects the gain in reconstruction accuracy to better recovered

screen organization. This is consistent with a progressive estimate gaining a clearer foundation for confidence-calibrated detail as impulsive disturbances weaken. PLRL retains the leading result throughout the full range. Figure 37 then shows how the surviving evidence takes different visible forms in natural-image, desktop, and document screens.

The examples in Figure 37 reveal distinct forms of coarse recovery across the displayed screen contents. In the higher-SNR document rows, PLRL recovers page boundaries and a vertical arrangement of text-like structures. The more heavily corrupted natural-image example preserves a coarser distribution of color and brightness, and the panel also includes desktop content. These examples connect the numerical noise sweep to concrete output structures while showing how the type of displayed content shapes the visible result.

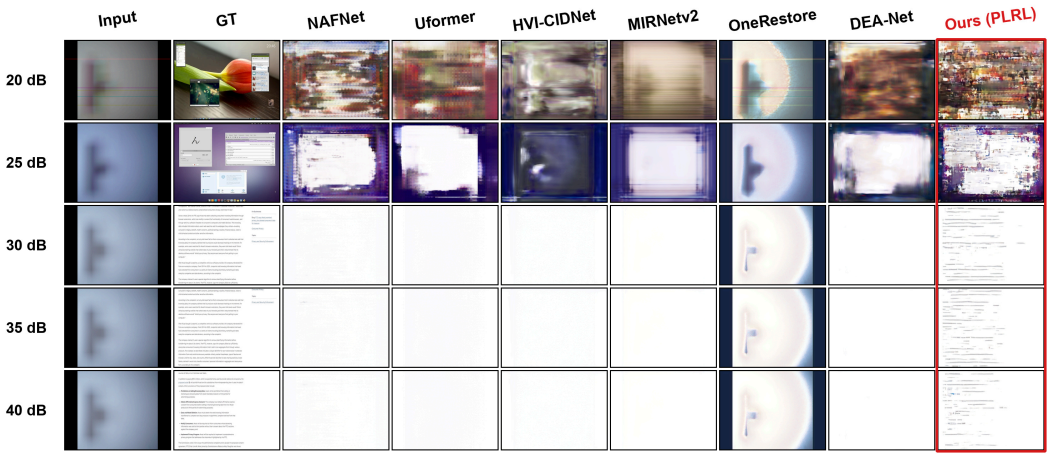


Figure 37: Representative Screen-2000 reconstructions under salt-and-pepper-noise SNRs of 20, 25, 30, 35, and 40 dB. Rows present selected screen contents at the indicated SNRs; each reconstruction is paired with the target shown in the same row. The fine horizontal line patterns in the 20-dB row are present in the saved observation and model outputs.

The different targets in Figure 37 show why the screen's regional organization is the common basis for evaluating feasibility. Large document areas and repeated line arrangements provide clearer organizational cues than the fine variations of a natural image. PLRL expresses the available cues at a coarse level, including page extent, line placement, and broad color distribution. These outputs support the same task objective across the selected contents: recovering the main displayed organization from reflected wall light. The final noise comparison extends this examination to Screen-Game.

Screen-Game. PLRL leads all 15 salt-and-pepper-noise comparisons on Screen-Game in Table 18. At 20 dB SNR, it obtains 18.227 dB PSNR, 0.665 SSIM, and 0.481 LPIPS. NAFNet provides the strongest competing PSNR of 17.867 dB and LPIPS of 0.581, while HVI-CIDNet provides the strongest competing SSIM of 0.652. At 40 dB SNR, PLRL reaches 18.377 dB, 0.674, and 0.474. Its advantage therefore extends to the full noise range in this third domain.

The modest variation in Table 18 shows that Screen-Game recovery retains a similar level of quality across these impulsive perturbations. At the strongest tested corruption, PLRL's LPIPS is 0.100 lower than the best competing value, giving a clear advantage in recovered appearance. The persistence of this gap complements the stronger changes measured on Screen-2000. Together, the domain comparisons support the usefulness of progressive recovery and confidence calibration across wall observations with different content distributions and local disturbances.

PLRL's Screen-Game measurements approach a stable level by 30 dB SNR in Figure 38. Its PSNR is 18.377, 18.376, and 18.377 dB at 30, 35, and 40 dB SNR, respectively; SSIM remains 0.674 and LPIPS remains 0.474 across those conditions. The larger changes occur between the lower-SNR settings. This trajectory identifies the part of the sweep where reducing the added corruption yields most of the measured reconstruction improvement.

The nearly level upper-SNR curves in Figure 38 show that the remaining impulsive perturbation has a small additional effect in that range. PLRL's continuing separation in LPIPS ties this stability to closer recovered appearance than the compared methods. The curve therefore complements the table by showing both the sensitivity at stronger corruption and the stable behavior as the observation

Table 18: Screen-Game results under additive salt-and-pepper noise at SNRs of 20, 25, 30, 35, and 40 dB.

Method	Source	20			25			30			35			40		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones																
OneRestore	ECCV'24	10.301	0.502	0.747	10.310	0.505	0.747	10.313	0.505	0.747	10.312	0.505	0.747	10.313	0.506	0.747
DocRes	CVPR'24	17.442	0.641	0.582	17.499	0.645	0.580	17.517	0.646	0.580	17.521	0.646	0.580	17.526	0.646	0.580
Pixel2Pixel	TPAMI'25	11.054	0.444	0.695	11.108	0.448	0.693	11.127	0.449	0.692	11.125	0.449	0.692	11.140	0.450	0.692
General-Purpose End-to-End Image Restoration Frameworks																
NAFNet	ECCV'22	**17.867**	0.646	**0.581**	**18.119**	0.656	0.577	18.109	0.657	0.577	**18.140**	0.657	0.576	**18.158**	0.658	0.576
DGUNet	CVPR'22	17.270	0.635	0.610	17.760	0.654	0.598	17.825	0.656	0.597	17.857	0.658	0.596	17.875	0.658	0.595
Uformer	CVPR'22	17.514	0.642	0.613	17.533	0.645	0.614	17.541	0.646	0.614	17.543	0.646	0.614	17.546	0.646	0.614
MIRNetv2	TPAMI'22	17.145	0.638	0.624	17.583	0.652	0.614	17.622	0.654	0.613	17.660	0.655	0.612	17.670	0.655	0.612
OKNet	AAAI'24	17.408	0.643	0.587	17.856	0.659	0.577	17.881	0.661	0.576	17.890	0.662	0.576	17.865	0.660	0.577
AST	CVPR'24	17.197	0.637	0.617	17.398	0.643	0.617	17.370	0.643	0.616	17.453	0.644	0.616	17.441	0.644	0.616
ConvIR	TPAMI'24	17.366	0.640	0.608	17.586	0.649	0.602	17.631	0.650	0.601	17.637	0.650	0.601	17.626	0.650	0.602
DnCNN	TIP'17	8.707	0.421	0.728	8.737	0.426	0.726	8.736	0.427	0.726	8.738	0.427	0.726	8.738	0.427	0.726
UNet	MICCAI'15	10.254	0.464	0.759	10.277	0.469	0.758	10.281	0.470	0.758	10.285	0.471	0.758	10.283	0.471	0.758
Models for Related Degradations																
DExNet	TPAMI'25	17.446	0.650	0.598	17.601	0.656	0.596	17.605	0.657	0.596	17.619	0.657	0.595	17.627	0.658	0.596
HVI-CIDNet	CVPR'25	17.743	**0.652**	0.597	17.779	0.655	0.598	17.764	0.654	0.598	17.771	0.655	0.598	17.774	0.655	0.598
Retinexformer	ICCV'23	17.487	0.639	0.600	17.831	0.652	0.594	17.906	0.654	0.593	17.903	0.654	0.593	17.896	0.653	0.593
DehazeFormer	TIP'23	17.410	0.631	0.617	17.693	0.645	0.609	17.738	0.648	0.608	17.764	0.648	0.608	17.756	0.648	0.608
DEA-Net	TIP'24	16.346	0.625	0.625	17.391	0.649	0.596	17.541	0.653	0.593	17.609	0.655	0.590	17.707	0.657	0.589
IRNeXt	ICML'23	17.644	0.650	0.585	18.104	**0.666**	**0.574**	**18.124**	**0.667**	**0.574**	18.133	**0.667**	**0.574**	18.113	**0.666**	**0.574**
SFHformer	ECCV'24	10.000	0.458	0.759	10.010	0.461	0.759	10.009	0.461	0.760	10.009	0.461	0.760	10.010	0.461	0.760
DarkIR	CVPR'25	4.465	0.041	0.930	4.465	0.041	0.935	4.465	0.041	0.936	4.465	0.041	0.936	4.465	0.041	0.936
PLRL	**Ours**	**18.227**	**0.665**	**0.481**	**18.327**	**0.671**	**0.476**	**18.377**	**0.674**	**0.474**	**18.376**	**0.674**	**0.474**	**18.377**	**0.674**	**0.474**

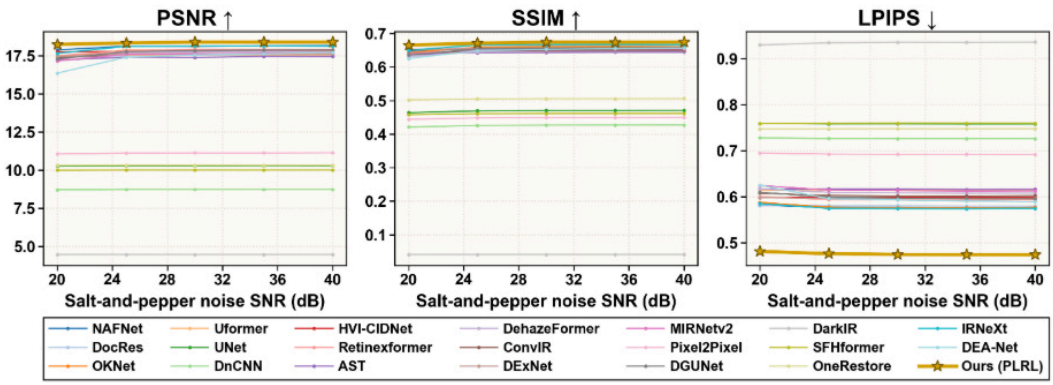


Figure 38: Screen-Game PSNR, SSIM, and LPIPS curves across salt-and-pepper-noise SNRs.

improves. Figure 39 adds the spatial interpretation through selected pale and dark interfaces with different surrounding panels and content regions.

The interfaces in Figure 39 retain distinct major regions in the PLRL reconstructions. The pale and dark examples show large central areas separated from surrounding panels, and some spatially distinct colored items remain visible in the darker screens. These elements supply coarse cues to the content and organization of the hidden display. The panel extends the quantitative comparison by showing the regional relationships that the wall observations support under the displayed salt-and-pepper conditions.

The recovered regional separation in Figure 39 follows the same priority as PLRL's reconstruction sequence. Large regions establish the screen organization, and smaller items contribute where their evidence supports recovery. Across the Gaussian and salt-and-pepper suites, the repeated metric advantages and these coarse outputs support initial feasibility under the tested noise conditions. The experiments so far have changed the light or disturbed the observation. We next change the camera position, which alters how the wall evidence is sampled.

A.14 Camera-Distance Robustness

Figure 40 collects the acquisition material for camera distance, angle, and height. Each row pairs a scene overview with two example camera placements. These images locate the camera relative to the display and wall and also support the viewing-direction discussions in Appendices A.15 and A.16. The distance factors and angular offsets used for quantitative evaluation are specified separately in the corresponding tables and experiment descriptions.

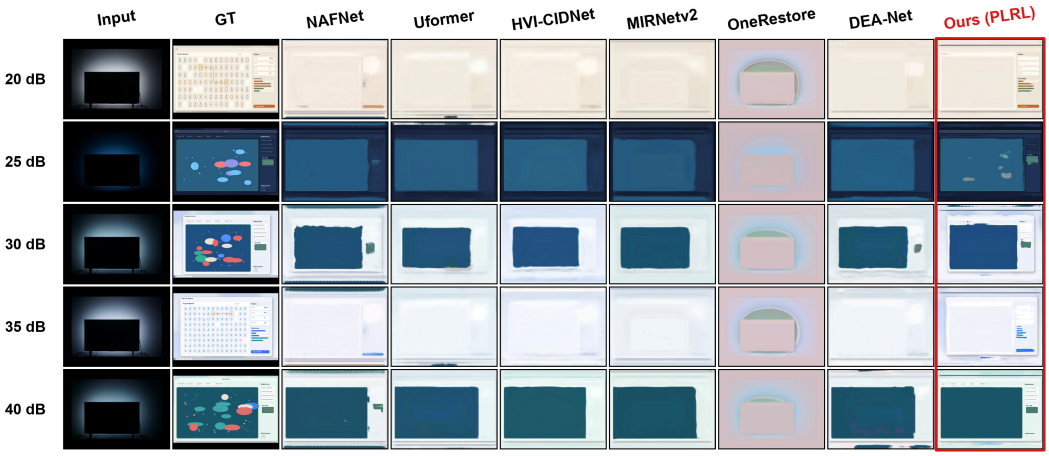


Figure 39: Representative Screen-Game reconstructions under salt-and-pepper-noise SNRs of 20, 25, 30, 35, and 40 dB. Rows present selected screen contents at the indicated SNRs; each reconstruction is paired with the target shown in the same row.

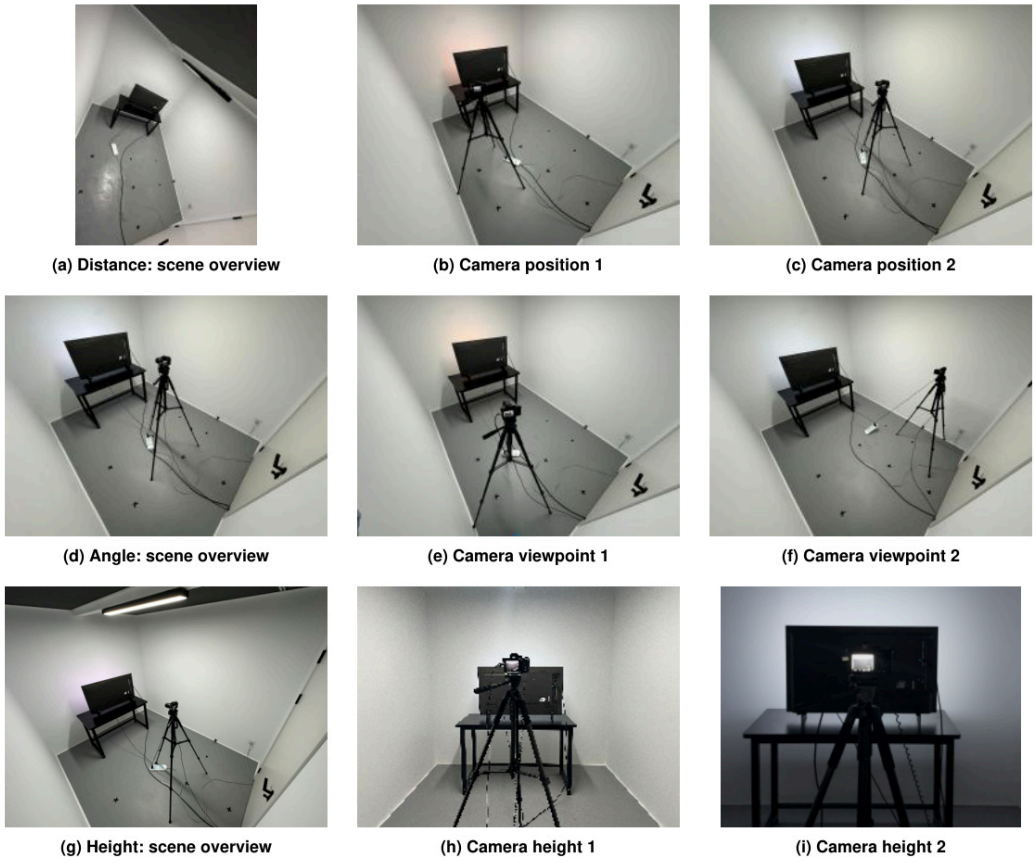


Figure 40: Camera placements in the robustness acquisition scenes. Rows show the distance, angle, and height acquisition arrangements, respectively; columns show a scene overview and two example camera placements. The display and wall form the reference scene while the observation position changes. The photographs show how camera distance, viewing direction, and height are varied relative to the display and wall.

Moving the camera changes the spatial sampling of the wall pattern from which the screen must be recovered. The quantitative Screen-2000 comparison evaluates relative camera-distance factors of 1.5, 2, and 3 with the model and evaluation setting fixed. The table and image panels examine recovery as the wall-light region occupies a smaller part of the camera view. This test asks whether the broad evidence used to establish screen layout remains informative across the evaluated observation distances.

PLRL leads all nine distance condition-metric comparisons in Table 19. At factors of 1.5 and 2, its PSNR is 27.274 and 27.275 dB, respectively, with LPIPS of 0.170 in both conditions. At factor 3, PLRL retains 26.864 dB PSNR, 0.861 SSIM, and 0.175 LPIPS. NAFNet supplies the strongest competing values at the farthest position: 21.023 dB, 0.742, and 0.392. The retained separation shows that PLRL continues to recover closer screen content as the camera moves back.

Table 19: Screen-2000 results at relative camera-distance factors of 1.5, 2, and 3, with the screen and wall fixed.

Method	Source	1.5			2			3		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones										
OneRestore	ECCV'24	10.420	0.465	0.723	10.420	0.464	0.720	10.433	0.459	0.729
DocRes	CVPR'24	16.611	0.592	0.525	16.820	0.600	0.517	15.489	0.561	0.554
Pixel2Pixel	TPAMI'25	10.264	0.478	0.671	10.241	0.475	0.670	10.118	0.466	0.692
General-Purpose End-to-End Image Restoration Frameworks										
NAFNet	ECCV'22	**21.548**	**0.755**	**0.382**	**21.461**	**0.754**	**0.383**	**21.023**	**0.742**	**0.392**
DGUNet	CVPR'22	18.948	0.694	0.484	18.950	0.694	0.483	18.794	0.680	0.493
Uformer	CVPR'22	18.567	0.695	0.479	18.786	0.697	0.475	17.453	0.670	0.504
MIRNetv2	TPAMI'22	18.266	0.691	0.492	18.319	0.691	0.492	18.297	0.690	0.494
OKNet	AAAI'24	20.340	0.728	0.456	20.483	0.732	0.451	19.098	0.694	0.486
AST	CVPR'24	18.543	0.685	0.491	19.106	0.694	0.478	16.822	0.645	0.533
ConvIR	TPAMI'24	19.967	0.710	0.469	20.096	0.713	0.467	19.070	0.678	0.495
DnCNN	TIP'17	8.797	0.346	0.730	8.806	0.345	0.729	8.812	0.341	0.732
UNet	MICCAI'15	9.182	0.241	0.808	9.181	0.240	0.808	9.188	0.240	0.807
Models for Related Degradations										
DExNet	TPAMI'25	19.925	0.728	0.442	19.994	0.730	0.440	19.504	0.719	0.452
HVI-CIDNet	CVPR'25	12.237	0.456	0.698	12.197	0.449	0.699	12.054	0.440	0.708
Retinexformer	ICCV'23	17.123	0.593	0.523	18.307	0.632	0.489	9.989	0.365	0.708
DehazeFormer	TIP'23	18.387	0.669	0.502	19.104	0.685	0.487	14.322	0.558	0.607
DEA-Net	TIP'24	20.386	0.744	0.432	20.490	0.744	0.430	17.314	0.684	0.501
IRNeXt	ICML'23	19.795	0.711	0.469	19.858	0.712	0.467	19.247	0.693	0.486
SFHformer	ECCV'24	8.807	0.396	0.811	8.808	0.395	0.809	8.793	0.393	0.814
DarkIR	CVPR'25	9.222	0.175	0.869	9.220	0.175	0.868	9.226	0.174	0.878
PLRL	**Ours**	**27.274**	**0.867**	**0.170**	**27.275**	**0.866**	**0.170**	**26.864**	**0.861**	**0.175**

The small changes in Table 19 indicate that the tested distances preserve enough sampled wall evidence for coarse recovery. PLRL's PSNR changes by approximately 0.41 dB from the closer conditions to factor 3, while its structural and perceptual advantages remain. This supports feasibility across the evaluated camera positions and links the result to the geometry of acquisition. The curves and examples below show the measured and visible consequences of this change.

The distance experiment changes the camera's position while preserving the screen-to-wall light path. As the camera moves farther away, the wall-light region occupies fewer camera pixels, reducing the spatial sampling available for reconstruction. This makes the distinction between broad and fine evidence especially relevant: the large-scale pattern used by PPRL can still provide a layout reference as local wall variations receive coarser representation. Figure 41 reports the reconstruction response across the three quantitatively evaluated distances.

The distance curves show almost identical PLRL measurements at factors 1.5 and 2, followed by a modest change at factor 3. In Figure 41, SSIM moves from 0.867 to 0.866 and then 0.861, while LPIPS changes from 0.170 at the closer positions to 0.175 at the farthest one. PLRL's trajectories remain separated from the alternatives throughout. This close grouping shows that the evaluated change in sampling has a limited effect on its aggregate coarse reconstruction quality.

The agreement across the curves in Figure 41 shows that distance tolerance extends across pixel fidelity, structure, and appearance. The broad wall pattern remains sufficiently informative to support the leading reconstruction at each measured position. This result fits the motivation for using scale-dependent evidence to establish screen organization before refining detail. The supplementary Screen-Game example in Figure 42 extends the visual inspection to the intermediate positions displayed in its rows, allowing the same interface regions to be compared as camera distance increases.

The interface in Figure 42 preserves its main divisions across relative camera-distance factors of 1.0, 1.5, 2.0, 2.5, and 3.0. PLRL retains the large pale workspace, the panel on the right, and the lower colored strip throughout the displayed positions. These regional relationships remain visible as the captured wall pattern changes with distance. The panel therefore gives a spatial counterpart

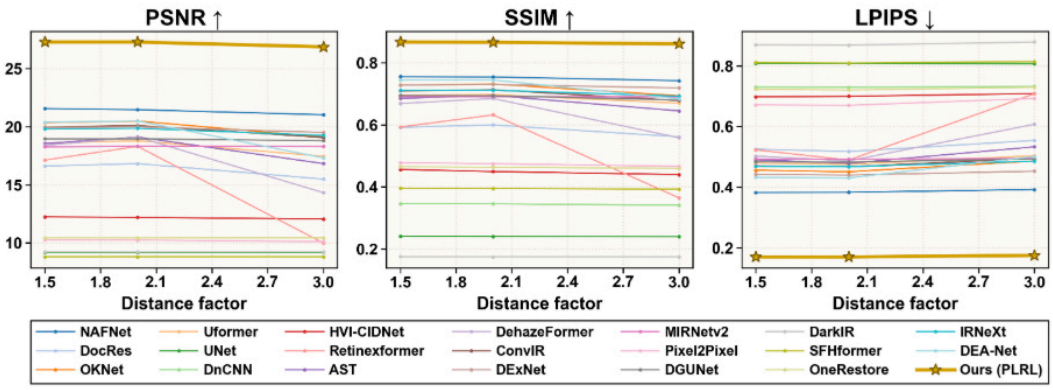


Figure 41: Screen-2000 PSNR, SSIM, and LPIPS curves across relative camera-distance factors of 1.5, 2, and 3.

to the stable measurements: the recovered screen continues to express the arrangement of its major content regions.

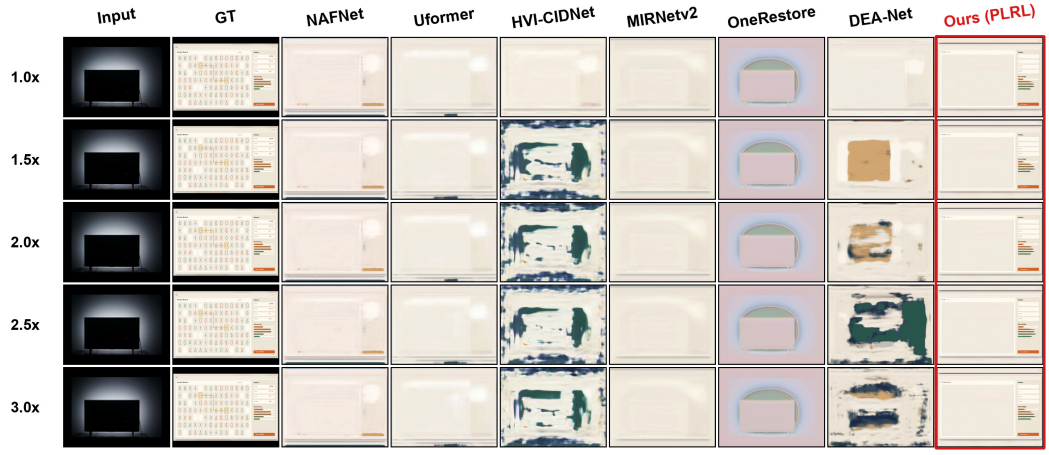


Figure 42: Supplementary Screen-Game reconstructions at relative camera-distance factors of 1.0, 1.5, 2.0, 2.5, and 3.0. The pale workspace, right-side panel, and lower colored strip remain visible across the displayed camera positions.

The retained placement of the right-side controls and lower boundary strengthens the correspondence with the target in Figure 42. PLRL preserves more of these relationships than the compared outputs, connecting distance tolerance to the actual screen content being recovered. This shows why the coarse layout is a meaningful outcome for the task: the display's main areas remain distinguishable across the tested camera positions. We next rotate the camera view, which changes the wall projection across the image and introduces a different spatial challenge.

A.15 Left-to-Right Camera-View Robustness

Changing the horizontal viewing direction alters the wall projection used to reconstruct the hidden screen. We evaluate views at 10 degrees left, centered, and 10 degrees right while keeping the Screen-2000 scene, model, and evaluation setting fixed. These offsets complement the distance experiment by changing viewing direction. The comparison asks whether PLRL retains useful screen organization when the camera observes the wall evidence under a different horizontal projection.

PLRL leads all nine horizontal-view condition-metric comparisons in Table 20. At the left view, it reaches 19.056 dB PSNR, 0.713 SSIM, and 0.464 LPIPS; at the right view, the values are 17.946 dB, 0.700, and 0.477. The strongest competing LPIPS is 0.474 from OKNet on the left and 0.493 from Retinexformer on the right. PLRL also achieves the highest PSNR and SSIM at both offsets, retaining its advantage in screen fidelity, structure, and appearance.

Table 20: Screen-2000 results for 10-degree left, centered, and 10-degree right horizontal camera views.

Method	Source	Left 10 deg			Center 0 deg			Right 10 deg		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones										
OneRestore	ECCV'24	10.274	0.460	0.719	10.409	0.464	0.714	10.410	0.464	0.713
DocRes	CVPR'24	17.942	0.687	0.552	21.138	0.748	0.400	15.473	0.631	0.500
Pixel2Pixel	TPAMI'25	10.206	0.477	0.671	10.325	0.482	0.664	10.263	0.480	0.666
General-Purpose End-to-End Image Restoration Frameworks										
NAFNet	ECCV'22	17.004	0.654	0.490	**21.614**	**0.758**	**0.379**	17.024	0.619	0.515
DGUNet	CVPR'22	17.443	0.671	0.510	18.961	0.695	0.482	17.139	0.644	0.534
Uformer	CVPR'22	12.886	0.545	0.594	19.201	0.706	0.464	11.929	0.508	0.623
MIRNetv2	TPAMI'22	17.064	0.674	0.512	18.235	0.691	0.492	17.117	**0.677**	0.509
OKNet	AAAI'24	15.974	0.515	**0.474**	20.582	0.738	0.445	17.104	0.644	0.536
AST	CVPR'24	11.600	0.507	0.634	19.389	0.704	0.467	12.276	0.525	0.619
ConvIR	TPAMI'24	18.061	0.663	0.508	20.266	0.728	0.449	17.596	0.649	0.524
DnCNN	TIP'17	8.769	0.343	0.727	8.882	0.346	0.724	8.761	0.345	0.728
UNet	MICCAI'15	9.115	0.243	0.804	9.184	0.240	0.806	9.149	0.242	0.802
Models for Related Degradations										
DExNet	TPAMI'25	16.650	0.662	0.508	20.018	0.731	0.438	16.547	0.640	0.522
HVI-CIDNet	CVPR'25	17.776	0.683	0.510	19.690	0.717	0.472	17.516	0.668	0.518
Retinexformer	ICCV'23	15.358	0.553	0.554	19.563	0.672	0.455	**17.630**	0.633	**0.493**
DehazeFormer	TIP'23	9.150	0.414	0.776	19.357	0.692	0.479	10.068	0.423	0.761
DEA-Net	TIP'24	17.310	0.672	0.507	20.645	0.749	0.424	15.450	0.474	0.582
IRNeXt	ICML'23	**18.187**	**0.690**	0.494	19.952	0.723	0.454	16.939	0.657	0.527
SFHformer	ECCV'24	8.569	0.373	0.822	8.817	0.394	0.807	8.964	0.397	0.839
DarkIR	CVPR'25	9.128	0.171	0.871	9.221	0.175	0.867	9.183	0.173	0.860
PLRL	**Ours**	**19.056**	**0.713**	**0.464**	**27.341**	**0.868**	**0.168**	**17.946**	**0.700**	**0.477**

Table 20 gives centered PLRL values of 27.341 dB PSNR, 0.868 SSIM, and 0.168 LPIPS, placing both offset views farther from the target than the centered condition. The continued lead at the offsets supports the usefulness of recovering broad organization and calibrating locally uncertain detail when projection changes. Figure 43 then separates the response to the left and right directions, showing how scene-dependent evidence affects the two views differently.

The centered view gives the strongest recovery, and the two horizontal offsets have different effects. Figure 43 shows a central maximum in PSNR and SSIM and a central minimum in LPIPS. Left-view PSNR exceeds right-view PSNR by 1.110 dB. This directional dependence is consistent with the scene-dependent wall projection and illumination represented in the forward model. The same screen can therefore provide wall evidence of different usefulness when observed from opposite horizontal offsets.

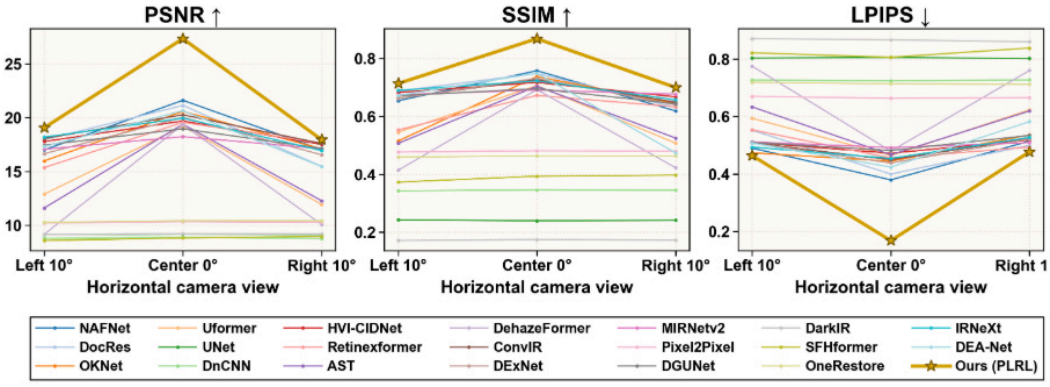


Figure 43: Screen-2000 PSNR, SSIM, and LPIPS curves across left-to-right camera-view offsets.

PLRL retains useful structural agreement at both offset views in Figure 43. SSIM is 0.713 on the left and 0.700 on the right, while the centered peak shows the advantage of the nominal observation. The curves connect viewing direction to reconstruction quality in direct terms: the shifted wall projection changes the precision of the recovered screen, and PLRL retains the leading measurements across the tested directions. The supplementary Screen-Game examples in Figure 44 show which screen regions remain visible in pale and dark interfaces under these horizontal view changes.

The examples in Figure 44 retain coarse cues to two different forms of screen content. In the pale interface, PLRL preserves the broad layout of the main workspace and surrounding regions. In the darker example, it recovers some separated colored regions within the principal content area.

The panel displays a centered pale interface together with left and right views of the pale and dark examples, making their recovered organization visible under the camera-direction changes.

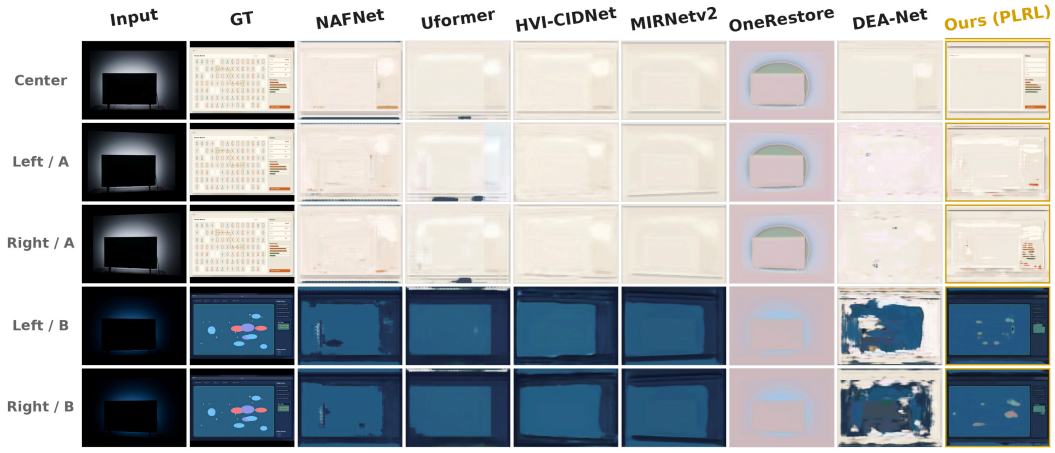


Figure 44: Supplementary Screen-Game reconstructions under horizontal camera-view offsets. A and B denote the pale and dark targets, respectively; the centered row uses A. Left and right correspond to $+10^\circ$ and -10° yaw in the evaluation convention. The examples show the main content regions and their placement under the three viewing directions.

The contrast between the two screens in Figure 44 shows how content shapes the visual evidence available for recovery. The pale interface provides large panel boundaries, while the darker screen relies more on separated items against a relatively uniform background. PLRL expresses both forms of coarse organization through the same progression from layout to structure and confidence-calibrated detail. This extends the horizontal-view result from aggregate metrics to recognizable regional relationships. The vertical-view experiment next changes the other projection axis.

A.16 Top-to-Bottom Camera-View Robustness

Vertical view changes provide a complementary test of recovery from a shifted wall projection. We evaluate 10 degrees upward, the centered view, and 10 degrees downward with the Screen-2000 scene, model, and evaluation setting fixed. As in the horizontal experiment, the camera observes the wall-light pattern from a changed direction. The test asks whether the remaining evidence supports the screen's broad visual arrangement when the projection changes along the vertical axis.

PLRL achieves the strongest values in all nine vertical-view condition-metric comparisons in Table 21. At the upward view, it reaches 20.232 dB PSNR, 0.720 SSIM, and 0.348 LPIPS. At the downward view, it reaches 19.588 dB, 0.704, and 0.365. NAFNet gives the strongest competing upward-view LPIPS of 0.461, and DExNet gives the strongest competing downward-view LPIPS of 0.472. These measurements show a persistent advantage in recovered screen appearance at both offsets.

PLRL retains the leading reconstruction results at both vertical offsets, extending the evidence across horizontal and vertical viewing changes. Its LPIPS reaches 0.348 in the upward view and 0.365 in the downward view. Figure 46 connects this advantage to the recovery of major desktop regions and dialog placement.

The vertical-view curves again place the strongest measurements at the centered observation. In Figure 45, the centered PLRL result is 27.341 dB PSNR and 0.168 LPIPS; both upward and downward offsets produce lower pixel fidelity and higher perceptual distance. Their SSIM values remain relatively close, at 0.720 and 0.704. The trajectories therefore show a shared effect from changing the vertical projection together with a difference between the two directions.

PLRL's perceptual advantage remains clear at both vertical offsets in Figure 45. Its lower LPIPS indicates closer overall appearance as the wall projection changes, complementing the pixel and structural comparisons. The physical model accounts for direction-dependent observation evidence, while the method uses progressive structure and confidence-calibrated detail to recover from that evidence. Figure 46 then shows the corresponding screen-level outcome through large desktop regions, a prominent dialog, and broad line arrangements.

Table 21: Screen-2000 results for 10-degree top, centered, and 10-degree bottom vertical camera views.

Method	Source	Top 10 deg			Center 0 deg			Bottom 10 deg		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Screen-Oriented Restoration Backbones										
OneRestore	ECCV'24	10.407	0.463	0.711	10.409	0.464	0.714	10.396	0.463	0.712
DocRes	CVPR'24	17.164	0.650	0.496	21.138	0.748	0.400	16.766	0.638	0.506
Pixel2Pixel	TPAMI'25	10.310	0.481	0.666	10.325	0.482	0.664	10.310	0.481	0.668
General-Purpose End-to-End Image Restoration Frameworks										
NAFNet	ECCV'22	18.343	0.682	**0.461**	**21.614**	**0.758**	**0.379**	17.506	0.663	0.484
DGUNet	CVPR'22	17.898	0.669	0.513	18.961	0.695	0.482	17.916	0.670	0.514
Uformer	CVPR'22	14.092	0.585	0.572	19.201	0.706	0.464	13.990	0.582	0.576
MIRNetv2	TPAMI'22	17.689	0.682	0.501	18.235	0.691	0.492	17.690	0.682	0.501
OKNet	AAAI'24	**18.906**	0.693	0.497	20.582	0.738	0.445	17.658	0.661	0.532
AST	CVPR'24	13.868	0.573	0.586	19.389	0.704	0.467	13.044	0.564	0.598
ConvIR	TPAMI'24	18.771	0.680	0.494	20.266	0.728	0.449	18.680	0.679	0.496
DnCNN	TIP'17	8.823	0.345	0.728	8.882	0.346	0.724	8.862	0.346	0.728
UNet	MICCAI'15	9.164	0.241	0.801	9.184	0.240	0.806	9.166	0.241	0.801
Models for Related Degradations										
DExNet	TPAMI'25	18.367	**0.699**	0.471	20.018	0.731	0.438	18.442	**0.700**	**0.472**
HVI-CIDNet	CVPR'25	17.916	0.670	0.521	19.690	0.717	0.472	17.782	0.668	0.524
Retinexformer	ICCV'23	15.879	0.569	0.547	19.563	0.672	0.455	15.429	0.555	0.563
DehazeFormer	TIP'23	9.612	0.420	0.770	19.357	0.692	0.479	9.726	0.421	0.767
DEA-Net	TIP'24	17.977	0.690	0.502	20.645	0.749	0.424	17.426	0.678	0.514
IRNeXt	ICML'23	18.699	0.696	0.488	19.952	0.723	0.454	**18.693**	0.696	0.489
SFHformer	ECCV'24	8.806	0.389	0.817	8.817	0.394	0.807	8.749	0.385	0.816
DarkIR	CVPR'25	9.209	0.174	0.866	9.221	0.175	0.867	9.212	0.173	0.866
PLRL	**Ours**	**20.232**	**0.720**	**0.348**	**27.341**	**0.868**	**0.168**	**19.588**	**0.704**	**0.365**

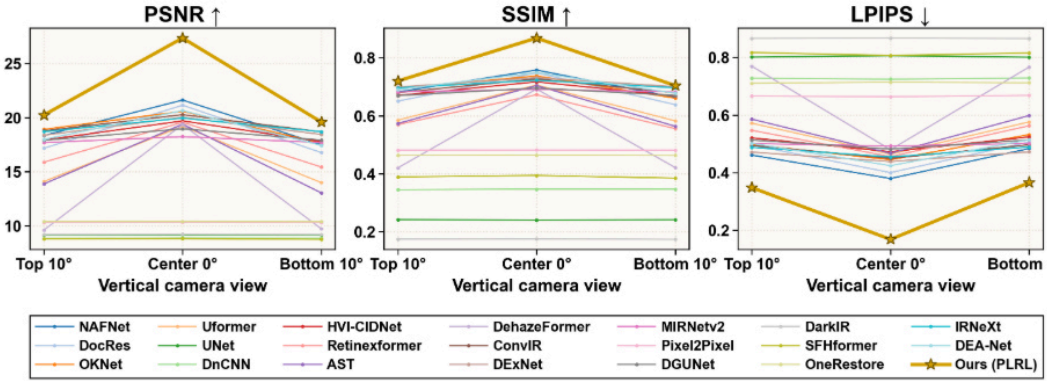


Figure 45: Screen-2000 PSNR, SSIM, and LPIPS curves across top-to-bottom camera-view offsets.

The main desktop regions and dialog placement remain visible under vertical view changes in Figure 46. In the green-background example, PLRL retains the large background area and the light dialog near the bottom. The blue-screen examples preserve their principal color field and portions of the line structure. These outputs identify major screen elements and their locations, connecting the view-robustness measurements to the kind of coarse content that an ordinary camera can recover from wall light.

The prominent dialog in Figure 46 reveals the hidden display's main foreground element and its position within the desktop. Fine text and local boundaries remain simplified, while these coarse spatial relationships persist across the displayed views. Having examined changes in the camera's view of the full wall pattern, we next test how recovery responds when individual wall regions are occluded.

A.17 Nine-Grid Local-Occlusion Analysis

Masking parts of the wall observation tests whether the remaining light still supports the hidden screen's main layout. In Figure 47, we occlude each cell of a fixed 3×3 partition in turn on a representative Screen-Game example. The top row shows the wall observation, hidden-screen target, full-input PLRL output, and masked-input PSNR map; the lower panels show the nine masked observations, indexed 0-8, and their reconstructions. The broad pale workspace remains visible across all nine masks, showing that coarse recovery persists from each partially observed input.

PLRL preserves the broad workspace under all nine local masks in Figure 47, while reconstruction precision varies with the occluded position. PSNR ranges from 18.8 to 19.7 dB: several upper and

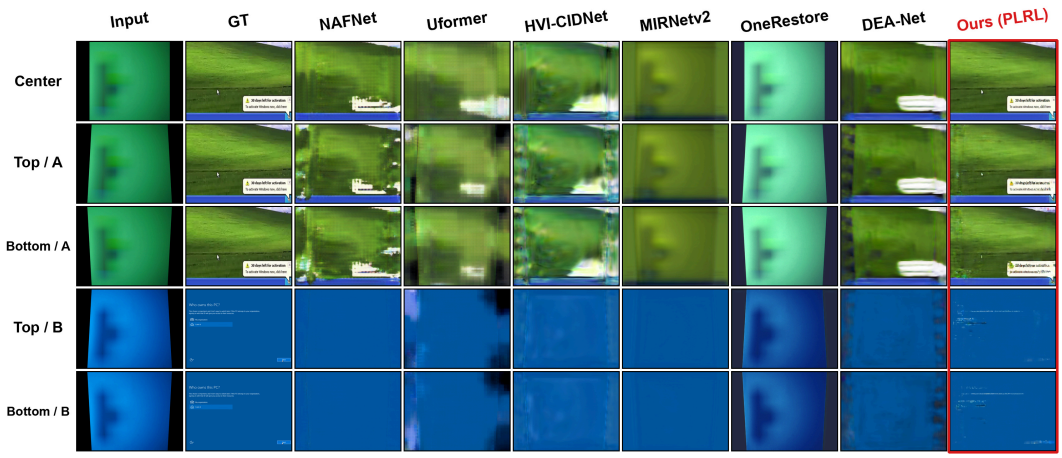


Figure 46: Representative Screen-2000 reconstructions under vertical camera-view offsets. A and B denote the green-background and blue-screen targets, respectively; the centered row uses A. Top and bottom correspond to $+10^\circ$ and -10° pitch in the evaluation convention.

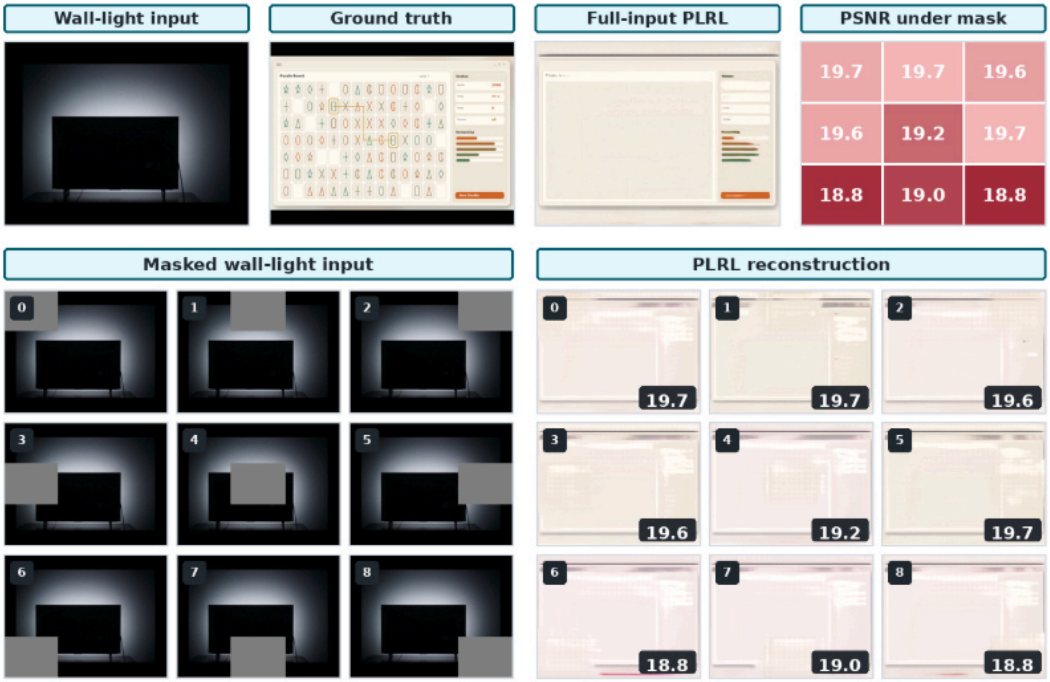


Figure 47: Nine-grid local-occlusion analysis on a Screen-Game example. The top row shows the wall-light observation, hidden-screen target, full-input PLRL reconstruction, and PSNR under each mask position. In the lower panels, each cell of a fixed 3×3 partition is occluded in turn (0–8), followed by its PLRL reconstruction. The masked-input PSNR values range from 18.8 to 19.7 dB, showing position-dependent sensitivity alongside retained broad screen organization.

middle mask positions give 19.6-19.7 dB, the central mask gives 19.2 dB, and the lower-row masks give 18.8, 19.0, and 18.8 dB. This spatial variation connects the recovered screen to the location of the available evidence. Such position-dependent sensitivity motivates using a spatial confidence field to assess detail within the recovery state.

The combination of retained layout and position-dependent sensitivity supports using wider context together with local restoration evidence. The following controls examine how the recovery strategy learns this organization under shared training resources and model architecture.

A.18 Fixed-Training-Duration Control

The training-duration control tests whether the advantage of progressive recovery persists as optimization continues. Table 22 compares Direct L1 ERM, Direct L2 ERM, Deep Supervision, Curriculum Learning, Self-Paced Learning, and PPRL after 10, 15, and 20 epochs on Screen-Figure and Screen-Game, using the same U-Net backbone. PPRL leads all 18 dataset-epoch-metric comparisons. Its consistent ranking across the three durations connects the improvement to a learning strategy that remains useful as the shared reconstruction network receives more optimization.

Table 22: Learning-strategy comparison on Screen-Figure and Screen-Game at 10, 15, and 20 training epochs under the same U-Net backbone and training setting.

Screen-Figure

Method	Epochs: 10			Epochs: 15			Epochs: 20		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Direct L1 ERM	15.314	0.5436	0.6192	15.328	0.5567	0.6288	15.336	0.5669	0.6120
Direct L2 ERM	15.391	0.5748	0.6224	15.404	0.5779	0.6191	15.411	0.5790	0.6097
Deep Supervision	15.304	0.5734	0.6190	15.318	0.5838	0.6174	15.325	0.5869	0.6131
Curriculum Learning	15.321	0.5563	0.6203	15.328	0.5467	0.6169	15.344	0.5829	0.6140
Self-Paced Learning	15.318	0.5436	0.6192	15.327	0.5597	0.6118	15.335	0.5849	0.6012
PPRL (Ours)	16.315	0.5966	0.6092	16.529	0.6267	0.6089	16.616	0.6269	0.6008

Screen-Game

Method	Epochs: 10			Epochs: 15			Epochs: 20		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Direct L1 ERM	13.307	0.5174	0.6846	13.317	0.5181	0.6836	13.326	0.5186	0.6830
Direct L2 ERM	13.332	0.5211	0.6892	13.340	0.5218	0.6892	13.345	0.5221	0.6878
Deep Supervision	13.313	0.5176	0.6843	13.338	0.5190	0.6835	13.354	0.5205	0.6831
Curriculum Learning	13.323	0.5197	0.6864	13.321	0.5189	0.6844	13.326	0.5191	0.6838
Self-Paced Learning	13.307	0.5174	0.6846	13.317	0.5181	0.6837	13.337	0.5216	0.6831
PPRL (Ours)	13.338	0.5475	0.6648	13.358	0.5675	0.6638	13.368	0.5485	0.6548

The Screen-Figure results show a sustained structural benefit from organizing recovery by scale. In Table 22, PPRL's PSNR rises from 16.315 dB at 10 epochs to 16.616 dB at 20 epochs, and SSIM rises from 0.5966 to 0.6269. At 20 epochs, the strongest competing PSNR is 15.411 dB from Direct L2 ERM and the strongest competing SSIM is 0.5869 from Deep Supervision. These comparisons support the value of assigning layout and structure explicit roles during learning, allowing the backbone to recover closer screen organization across the tested training durations.

The Screen-Game control places the clearest gain in structural and perceptual agreement. At 20 epochs, PPRL reaches 0.5485 SSIM and 0.6548 LPIPS, compared with the strongest competing values of 0.5221 from Direct L2 ERM and 0.6830 from Direct L1 ERM. PPRL reaches 0.5675 SSIM at 15 epochs, so its persistent ranking accompanies metric-specific changes as training proceeds. Scale-specific supervision supplies an order for learning the wall-to-screen mapping, beginning with layout and structure before finer refinement. Lee et al. (2015) introduced auxiliary objectives at hidden layers to strengthen intermediate feature learning. With scale-specific supervision on the shared backbone, PPRL reaches 0.5966 SSIM on Screen-Figure and 0.5475 on Screen-Game after 10 epochs, exceeding all five alternatives after 20 epochs, whose best values are 0.5869 and 0.5221, respectively (Table 22). The next control asks how this learning advantage changes when fewer paired examples are available.

A.19 Limited-Training-Data Control

The sample-size control examines whether progressive supervision remains useful when paired wall-screen examples are limited. Table 23 evaluates training sets of 100, 300, and 1,000 pairs on Screen-Figure and Screen-Game using the same U-Net backbone, optimization setting, and compared learning strategies. PPRL leads all 18 dataset-sample-metric comparisons. The advantage across these sample sizes shows that organizing recovery by scale benefits learning from both the smaller and larger evaluated sets of paired observations.

Table 23: Learning-strategy comparison on Screen-Figure and Screen-Game with 100, 300, and 1,000 training samples under the same U-Net backbone and training setting.

Screen-Figure

Method	Samples: 100			Samples: 300			Samples: 1000		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Direct L1 ERM	15.245	0.5856	**0.6170**	15.270	0.5863	0.6162	**15.302**	0.5867	0.6153
Direct L2 ERM	**15.284**	0.5858	0.6177	**15.315**	0.5868	0.6168	**15.376**	0.5880	**0.6100**
Deep Supervision	15.265	**0.5862**	0.6175	15.278	0.5869	**0.6161**	15.299	0.5875	0.6152
Curriculum Learning	15.243	0.5855	0.6172	15.268	0.5862	0.6163	15.301	0.5867	0.6143
Self-Paced Learning	15.245	0.5856	0.6173	15.271	**0.5872**	0.6163	**15.302**	**0.5897**	0.6144
PPRL (Ours)	**15.345**	**0.5956**	**0.6072**	**15.385**	**0.5966**	**0.6022**	**15.415**	**0.6166**	**0.5922**

Screen-Game

Method	Samples: 100			Samples: 300			Samples: 1000		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Direct L1 ERM	13.246	0.5144	0.6868	13.285	0.5151	0.6858	13.303	0.5166	0.6846
Direct L2 ERM	**13.269**	**0.5154**	0.6880	**13.297**	**0.5171**	**0.6827**	**13.323**	**0.5194**	**0.6564**
Deep Supervision	13.248	0.5151	0.6874	13.289	0.5160	0.6867	13.302	0.5166	0.6843
Curriculum Learning	13.245	0.5145	0.6868	13.287	0.5153	0.6859	13.306	0.5170	0.6849
Self-Paced Learning	13.246	0.5144	**0.6867**	13.285	0.5151	0.6858	13.303	0.5166	0.6847
PPRL (Ours)	**13.295**	**0.5344**	**0.6668**	**13.305**	**0.5444**	**0.6568**	**13.325**	**0.5574**	**0.6478**

The Screen-Figure comparison shows a structural advantage even at the smallest training size. With 100 pairs in Table 23, PPRL reaches 0.5956 SSIM and 0.6072 LPIPS, compared with the strongest competing values of 0.5862 from Deep Supervision and 0.6170 from Direct L1 ERM. With 1,000 pairs, PPRL reaches 0.6166 SSIM and 0.5922 LPIPS. More paired evidence improves its reconstruction, while the persistent comparative advantage supports giving each progressive stage a defined scale of screen structure to learn.

The Screen-Game results extend that learning advantage across the same sample-size range. PPRL reaches 0.5344 SSIM and 0.6668 LPIPS with 100 pairs and 0.5574 and 0.6478 with 1,000 pairs. At 1,000 pairs, the strongest competing values are 0.5194 SSIM and 0.6564 LPIPS from Direct L2 ERM. These results support progressive supervision as an effective way to learn screen organization from the available wall-screen pairs. Curriculum learning, self-paced learning, and deep supervision provide established ways to organize learning and intermediate training signals Bengio et al. (2009); Kumar et al. (2010); Lee et al. (2015). With 100 training pairs, PPRL reaches 0.5956 SSIM on Screen-Figure and 0.5344 on Screen-Game, exceeding all five alternatives trained with 1,000 pairs, whose best SSIM values are 0.5897 and 0.5194, respectively (Table 23). We next examine the learning strategy and confidence representation separately within the shared backbone.

A.20 Same-Backbone Controlled Comparisons

The shared-backbone comparisons connect reconstruction gains to how the model learns and represents wall evidence. All alternatives use the same U-Net backbone and training conditions. Direct L1 ERM, Direct L2 ERM, and Pixel-only ERM denote the evaluated empirical-risk-minimization objectives; Curriculum Learning (Bengio et al., 2009), Self-Paced Learning (Kumar et al., 2010), and Deep Supervision (Lee et al., 2015) provide representative learning families. Jigsaw (Noroozi & Favaro, 2016), RotNet (Gidaris et al., 2018), Autoencoder (Hinton & Salakhutdinov, 2006), and Denoising Autoencoder (Vincent et al., 2008) provide the compared representation families. Tables 24 and 25 report the point estimates obtained in these settings, examining the two complementary choices behind PLRL.

PPRL leads all six learning-strategy comparisons across the two domains in Table 24. On Screen-Figure, it reaches 15.376 dB PSNR, 0.5880 SSIM, and 0.6172 LPIPS. Self-Paced Learning gives the strongest competing PSNR of 15.322 dB and SSIM of 0.5867, while Direct L2 ERM gives the strongest competing LPIPS of 0.6200. On Screen-Game, PPRL reaches 13.353 dB, 0.5494, and 0.6843, compared with the strongest competing values of 13.336 dB from Curriculum Learning and 0.5366 SSIM and 0.6846 LPIPS from Direct L1 ERM. The gain appears across all three metrics in this shared-architecture setting.

The highest SSIM on both domains in Table 24 supports the structural purpose of progressive learning. PPRL assigns intermediate estimates distinct recovery roles, aligning supervision with the different strengths of information in the wall observation. The resulting advantage indicates that this organization helps the shared backbone reconstruct the screen's broad arrangement and structure. Once that structural estimate is established, detail still varies in reliability across the image. Table 25 therefore examines the complementary representation choice used to assess the evidence for local refinement.

Table 24: Learning-strategy comparison on Screen-Figure and Screen-Game with the same U-Net backbone and fixed training conditions.

	Screen-Figure			Screen-Game		
Method	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Direct L2 ERM	15.302	0.5757	0.6200	13.314	0.5167	0.6888
Direct L1 ERM	**15.313**	**0.5837**	**0.6893**	**13.323**	**0.5366**	**0.6846**
Curriculum Learning	15.314	0.5767	0.6865	13.336	0.5170	0.6849
Self-Paced Learning	15.322	0.5867	0.6596	13.303	0.5156	0.6847
Deep Supervision	15.299	0.5735	0.6254	13.302	0.5126	0.6848
PPRL (Ours)	**15.376**	**0.5880**	**0.6172**	**13.353**	**0.5494**	**0.6843**

LPCR leads all six representation-strategy comparisons in Table 25. On Screen-Figure, it obtains 15.884 dB PSNR, 0.5947 SSIM, and 0.5978 LPIPS; on Screen-Game, it obtains 14.100 dB, 0.5732, and 0.6513. The strongest competing Screen-Game values are 13.312 dB PSNR from Pixel-only ERM, 0.5366 SSIM from Jigsaw, and 0.6568 LPIPS from Autoencoder. This comparison shows improvements in fidelity, structure, and perceptual appearance when the shared backbone uses the evaluated light-propagation confidence representation.

Table 25: Representation-strategy comparison on Screen-Figure and Screen-Game with the same U-Net backbone and fixed training conditions.

	Screen-Figure			Screen-Game		
Method	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Jigsaw	15.290	0.5566	0.6176	13.302	0.5366	0.6965
Pixel-only ERM	**15.286**	**0.5868**	**0.6174**	**13.312**	**0.5276**	**0.6835**
RotNet	15.194	0.5565	0.6076	13.301	0.5146	0.6864
Autoencoder	15.078	0.5745	0.6875	13.300	0.5198	0.6568
Denoising Autoencoder	15.248	0.5865	0.6177	13.279	0.5036	0.6637
LPCR (Ours)	**15.884**	**0.5947**	**0.5978**	**14.100**	**0.5732**	**0.6513**

Jigsaw and RotNet learn visual representations by predicting patch order and image rotation, respectively Noroozi & Favaro (2016); Gidaris et al. (2018). LPCR learns to assess proposed screen detail through the wall observation and current reconstruction, reaching 0.5947 SSIM on Screen-Figure versus 0.5566 and 0.5565 for the Jigsaw- and RotNet-based alternatives in our shared-backbone comparison (Table 25). Scale compatibility relates the candidate to the observation, edge state records structural changes introduced during recovery, and image-wide context places local variation within the complete pattern. LPCR combines these cues to weight detail, tying the representation directly to the uncertainty in recovering screen structure from mixed wall light. The final ablation returns to the preceding step and examines the physical guidance and recovery order that establish the structure being refined.

A.21 Supplementary Ablation

The internal PPRL ablation examines the measured differences associated with physical guidance and progressive order in hidden-screen recovery. Table 26 compares the complete model with physical-guidance and progressive-learning ablations, together with Retinex, gradient, chromaticity, reverse-frequency, parallel-multiband, and alternating-band alternatives. Under the shared full-training setting on OS-Lock, OS-Home, Screen-Figure, and Screen-Game, PLRL leads all 12 dataset-metric comparisons. The physical-guidance alternatives draw on Retinexformer (Cai et al., 2023), DExNet (Huang et al., 2025), and HVI-CIDNet (Yan et al., 2025). UHD-Processor provides a related progressive-frequency-learning reference (Liu et al., 2025). The three frequency-learning alternatives are internal PLRL controls: they share the PLRL architecture and change only the frequency-band loss-weight schedule. Reverse-frequency learning shifts the largest weight from high to middle to low frequencies over training; parallel multiband learning assigns equal weights to all three bands; alternating-band learning cycles the active band across epochs.

The physical-guidance comparison supports organizing recovery around the way screen information reaches the wall. In Table 26, adding the full physical guidance raises OS-Lock PSNR from 15.038 to 17.088 dB. Retinexformer guides restoration with estimated illumination, while HVI-CIDNet decouples color and intensity for low-light enhancement Cai et al. (2023); Yan et al. (2025). PLRL guides reconstruction by the screen structure preserved at each spatial scale, reaching 14.950 dB PSNR on OS-Home compared with 13.811 and 13.908 dB for the Retinexformer- and HVI-CIDNet-based guidance alternatives in our replacement experiment (Table 26). The advantage connects the method to its physical motivation: spatial scales retain different strengths of screen information after transport, and the reconstruction benefits from giving those scales defined structural roles.

Table 26: Internal PPRL ablation on OS-Lock, OS-Home, Screen-Figure, and Screen-Game under the shared full-training setting. The three PLRL variants use the same architecture with different frequency-band loss-weight schedules.

Method	Model	OS-Lock			OS-Home			Screen-Figure			Screen-Game		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Physical Guidance													
Physical-guidance ablation	—	15.038	0.5553	0.6760	13.267	0.5034	0.7534	18.429	0.6932	0.6349	**16.734**	**0.6058**	0.6393
Retinex guidance	Retinexformer	14.666	**0.5635**	0.6705	13.811	0.5231	0.7475	18.003	**0.7078**	0.6350	16.574	0.6028	**0.6325**
Gradient prior	DExNet	**15.042**	0.5483	0.6727	13.752	0.5242	0.7484	18.173	0.7057	0.6599	16.587	0.5923	0.6356
Chromaticity prior	HVI-CIDNet	14.760	0.5545	0.6718	13.908	0.5212	0.7477	**18.789**	0.6999	**0.6178**	16.730	0.5978	0.6345
Progressive Learning													
Progressive-learning ablation	—	14.853	0.5558	0.6715	**13.951**	**0.5353**	0.7379	14.600	0.5919	0.6665	16.445	0.5959	0.6361
Reverse-frequency learning	UHD-Processor	15.002	0.5625	**0.6688**	13.922	0.5311	**0.7349**	18.690	0.6977	0.6212	16.702	0.5830	0.6354
Parallel multiband learning	DEA-Net	14.869	0.5513	0.6749	13.352	0.5068	0.7547	18.694	0.6959	0.6223	16.573	0.5849	0.6394
Alternating-band learning	DGUNet	14.961	0.5581	0.6711	13.351	0.5086	0.7577	18.651	0.6893	0.6263	16.175	0.5761	0.6520
Ours	**PLRL**	**17.088**	**0.6217**	**0.6273**	**14.950**	**0.5493**	**0.7002**	**19.336**	**0.7136**	**0.4712**	**18.378**	**0.6743**	**0.4736**

The frequency-learning controls show how changing the frequency-band supervision schedule affects the recovered screen. On Screen-Figure, PLRL achieves 0.4712 LPIPS, compared with 0.6212 for reverse-frequency learning, 0.6223 for parallel-multiband learning, and 0.6263 for alternating-band learning. On Screen-Game, PLRL reaches 0.6743 SSIM, compared with 0.5830, 0.5849, and 0.5761 for those variants. These results favor the full PLRL training setting over the three alternative frequency-band supervision schedules.

The supplementary evidence supports the initial feasibility of recovering principal hidden-screen content from an ordinary camera's view of wall light. Across the four main screen-content domains, the external comparisons show closer coarse screen appearance with PLRL than with the evaluated restoration methods. The surface, illumination, noise, and camera-position experiments extend that result across the tested conditions. The shared-backbone comparisons and ablations connect the improvement to progressive structural recovery and light-propagation confidence. Together, these findings support the paper's central outcome: the diffuse wall pattern contains usable information from which PLRL recovers the hidden display's brightness distribution, regional layout, and major contours.