Robust Unsupervised StyleGAN Image Restoration

Coforge Data and AI Lab, IIT (BHU) Varanasi

Lakshya	**22095129**
Yash Sachan	**22095127**
Kothapalli Santosh	**22095054**

Date: 12 October, 2025

Introduction:

Problem:

- Images often suffer from degradations such as noise, blur, and compression, reducing their visual quality.Image restoration aims to reconstruct clean images from such corrupted inputs.Modern approaches leverage GANs like StyleGAN, capable of generating realistic, high-fidelity images by learning detailed visual structures.

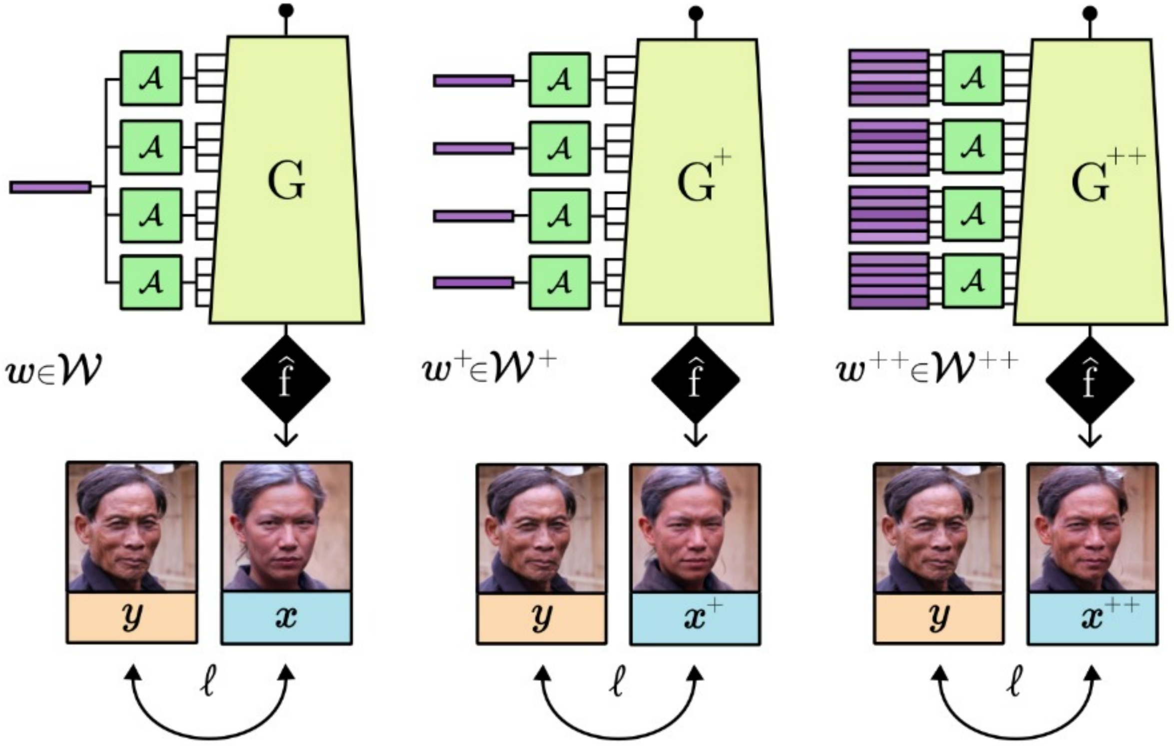
- GAN-based Image Restoration inverts the StyleGAN generative process to recover clean images. It searches for a latent representation whose degraded version matches the corrupted input. This enables restoration tasks such as super-resolution, denoising, inpainting, and deartifacting without paired data.

Motivation/Limitations:

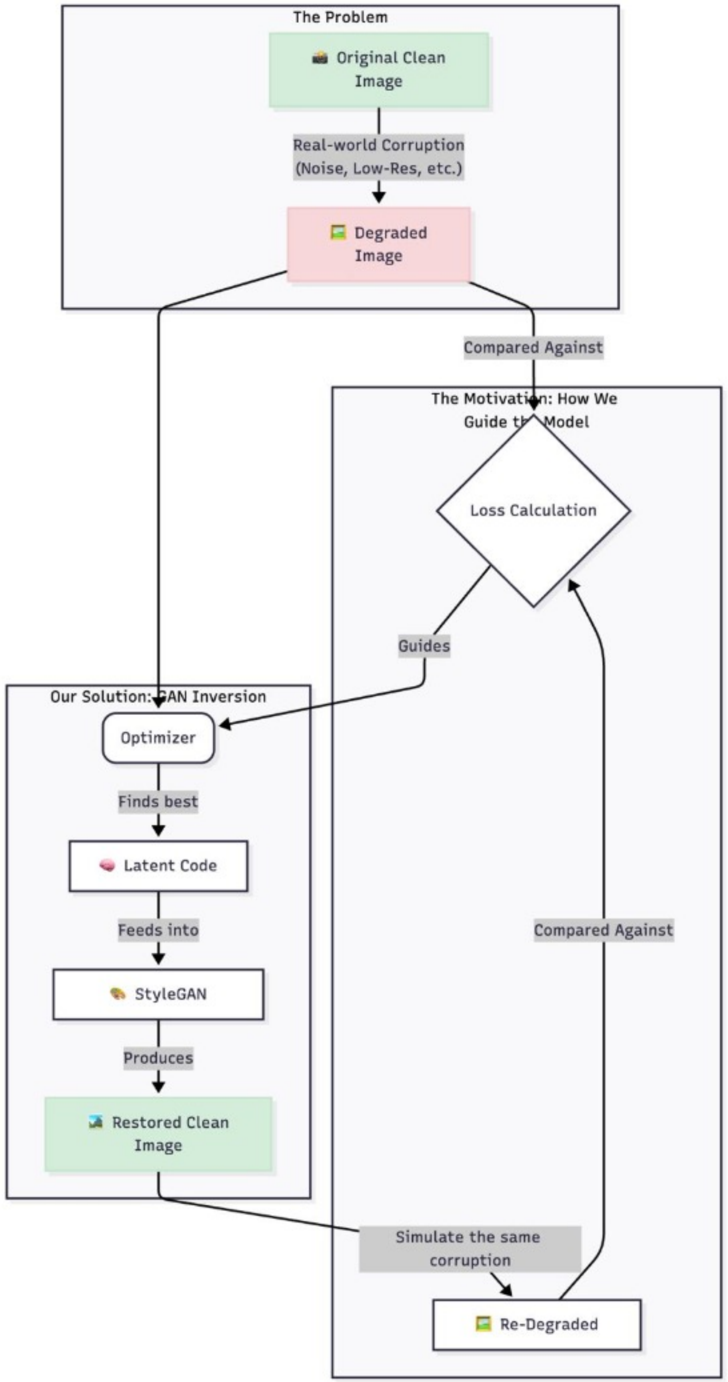
- Existing unsupervised StyleGAN approaches require extensive manual tuning for each degradation level and task.
- Different regularization weights and hyperparameters are needed for every scenario.
- This lack of robustness makes deployment time-consuming and inconsistent across conditions.

Key Idea / Novel Contribution:

- We propose a Robust Unsupervised StyleGAN Restoration framework.
- It introduces a fixed, universal set of hyperparameters that works across all restoration tasks and their combinations.
- By eliminating complex regularization terms, the model achieves stable, high-quality restoration with a simple "set-and-forget" design.



(a) Phase I (b) Phase II (c) Phase III



Related Work

Category	Method	Gap/Limitations
GAN Priors & Regularization	mGANprior (CVPR 2020), GAN Prior + Regularization (ECCV 2023)	Sensitive to noise levels and requires extensive per-task tuning.
Semi-Supervised Inversion	Pivotal Tuning Inversion (PTI) (ICCV 2021)	Requires fine-tuning for each individual image to preserve identity.
Diffusion Models	DDRM (NeurIPS 2022)	Performance is slow and highly dependent on the number of diffusion steps.
Domain Adaptation	StyleGAN-NADA (CVPR 2022)	Fails to generalize to degradation types not seen during adaptation.
Proposed Method (This Work)	**Robust Unsupervised StyleGAN Image Restoration** framework , which restores images by optimizing a latent code through three progressive phases (W, W+, and W++) using StyleGan V2	

Key Limitations:

1. Heavy dependence on manual regularization tuning.

2.Degradation-specific pipelines; fail under mixed noise or unseen artifacts.

3. Poor generalization and high computational overhead during optimization.

Gap Leading to Proposed Method:

There is no unified, robust unsupervised framework that performs consistently across all degradation types.Our work addresses this by introducing a single hyperparameter-free, “set-and-forget” approach, ensuring scalability and reliability.

Categories of Methods:

1. Supervised GAN Restoration – Requires paired clean–corrupted data; limits generalization.

2.Unsupervised GAN Inversion – Learns latent code to reconstruct degraded images without supervision.

3. Regularization-based Optimization – Balances realism and fidelity via custom loss terms, needing per-task tuning.

Proposed Method (Overview):

Our approach enhances the 3-phase (W, W+, W++) latent optimization framework by leveraging the StyleGAN2 model for its superior architectural stability and high-fidelity generative capabilities. We selected StyleGAN2 over models like StyleGAN-NADA because its well-structured latent space is critical for precise, optimization-based inversion.

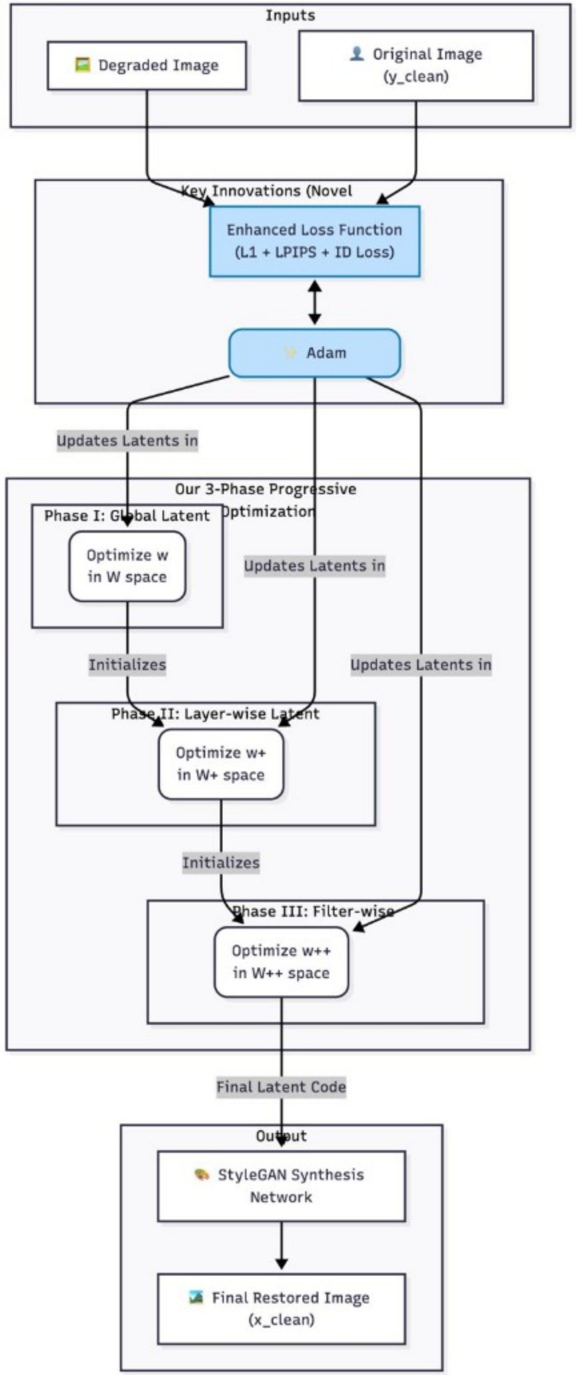
It addressed 4 core restorations tasks:

1) **Upsampling:** To increase the resolution of an image, effectively turning a low-resolution input into a high-resolution, detailed output.

2) **Denoising:** To remove unwanted noise and grain from an image. The framework was specifically tested against difficult, **non-linear noise models** like Poisson and Bernoulli noise

3) **Inpainting:** Reconstruct missing or occluded portions of an image. The experiments simulated this degradation by filling in areas with **random strokes.**

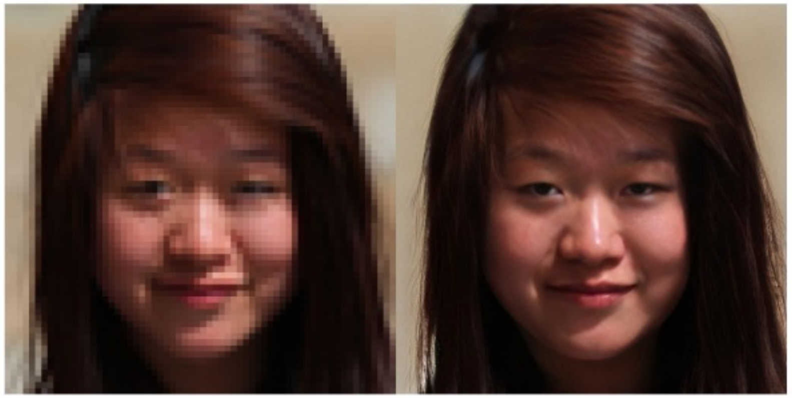
4) **Deartifacting:** To eliminate visual distortions and artifacts caused by lossy compression algorithms. The primary focus of the evaluation was on correcting artifacts from JPEG compression.



Deartifacting

Denoising

Inpainting



Upsampling

Key Enhancement Made to StyleGan V2

```
Algorithm: Robust Image Restoration with Progressive Latent Expansion

Input: Degraded image y, Original image y_clean, Generator G
Output: Restored image x++

// Phase I: Global Latent Optimization
1. w = G.mapping.w_avg
2. optimizer = Adam([w], lr=lr_1)
3. For step = 1 to N_1 do:
4.     x = G.synthesis(w)
5.     loss, _ = loss_fn(x, y, y_clean) // Uses L1+LPIPS+ID Loss
6.     loss.backward()
7.     optimizer.step()
8.     // Check early stopping condition

// Phase II: Layer-Wise Latent Optimization
9. w+ = expand(w) // Repeat w for each layer
10. optimizer = Adam([w+], lr=lr_2)
11. For step = 1 to N_2 do:
12.     x+ = G.synthesis(w+)
13.     loss, _ = loss_fn(x+, y, y_clean)
14.     loss.backward()
15.     optimizer.step()
16.     // Check early stopping condition

// Phase III: Filter-Wise Latent Optimization
17. w++ = expand(w+) // Repeat w+ for each filter
18. optimizer = Adam([w++], lr=lr_3)
19. For step = 1 to N_3 do:
20.     x++ = G.synthesis(w++)
21.     loss, _ = loss_fn(x++, y, y_clean)
22.     loss.backward()
23.     optimizer.step()
24.     // Check early stopping condition

25. return x++
```

Advanced Identity-Preserving Loss Function:

- ➢ Integrated a new **Identity Loss** into the original loss function to maintain facial identity
- ➢ Utilizes **pre-trained ArcFace model** to compare facial embeddings of restored and ground truth images
- ➢ Minimizes the **identity gap** between restored and original images
- ➢ Final loss = weighted combination of **Perceptual (LPIPS), Pixel-wise (L1),** and **Identity (ID)** losses

Improved Optimization and Efficiency:

The original Stochastic Gradient Descent (SGD) optimizer was replaced with the Adam optimizer for all three phases to achieve faster and more stable convergence.

```
// Initialization (at the start of each phase)
1. latent_vector.requires_grad = True
2. optimizer = Adam([latent_vector], lr=learning_rate)

// Update Step (inside the optimization loop)
3. loss.backward()       // Calculate gradients
4. optimizer.step()      // Update latent_vector based on gradients
5. optimizer.zero_grad()  // Reset gradients for the next step
```

Rigorous and Automated Evaluation:

A dedicated metrics module was created to formalize performance evaluation.After each task, we automatically calculate three key metrics based on the original paper's definitions:

- Accuracy (LPIPS): Perceptual similarity between the restored image and the ground truth.
- Realism (FID): Measures the photo-realism of the restored images using Fréchet Inception Distance.
- Fidelity (LPIPS): Compares a re-degraded restored image against the original degraded input to measure faithfulness.

```
// After processing all images for a task
1. DEFINE ground_truth_dir, restored_dir, degraded_dir

2. // Calculate metrics based on the paper's definitions
3. accuracy_score = calculate_accuracy_lpips(restored_dir, ground_truth_dir)
4. realism_score = calculate_realism_fid(restored_dir, ground_truth_dir)
5. fidelity_score = calculate_fidelity_lpips(restored_dir, degraded_dir)

6. // Display the final report
7. PRINT "----------------------------------------"
8. PRINT "PERFORMANCE REPORT FOR TASK: [Task Name]"
9. PRINT "- Accuracy (LPIPS ↓):", accuracy_score
10. PRINT "- Realism (FID ↓):", realism_score
11. PRINT "- Fidelity (LPIPS ↓):", fidelity_score
12. PRINT "----------------------------------------"
```

```
// Initialization (before the loop for each phase)
1. best_loss = infinity
2. patience_counter = 0
3. PATIENCE_LIMIT = 15

// Inside the optimization loop (after loss calculation)
4. IF current_loss < best_loss THEN
5.     best_loss = current_loss
6.     patience_counter = 0  // Reset counter on improvement
7. ELSE
8.     patience_counter += 1 // Increment counter if no improvement
9. END IF

10. IF patience_counter >= PATIENCE_LIMIT THEN
11.     print("Stopping early...")
12.     BREAK LOOP
13. END IF
```

Early Stopping for time optimization:

An early stopping mechanism was introduced to automatically halt the optimization process if the loss stops improving, significantly reducing computation time without sacrificing quality.

Problem Formulation:

1. Formal Task Definition

The primary objective is to recover a **latent code** $\tilde{\mathbf{w}}$. This code, when passed through the **StyleGAN generator** $G(\tilde{\mathbf{w}})$, produces a clean, high-quality image. The success of the restoration is measured by minimizing the difference between a degraded version of our generated image, $f(G(\tilde{\mathbf{w}}))$, and the original corrupted **target image**, $\mathbf{y}$.

2. Notations

- $\mathbf{y}$: The observed, degraded target image we want to restore.
- $G(\cdot)$: The pre-trained StyleGAN **Generator** network.
- $f(\cdot)$: A differentiable function that approximates the real-world **degradation** (e.g., blur, noise).
- $\mathbf{w}$: The latent code in a specific latent space ($\mathcal{W}$, $\mathcal{W}^+$, etc.) that we are optimizing.
- L_{total}: The **total loss function** that quantifies the error.

Problem Formulation(contd.)

3. Objective (Composed Degradations)

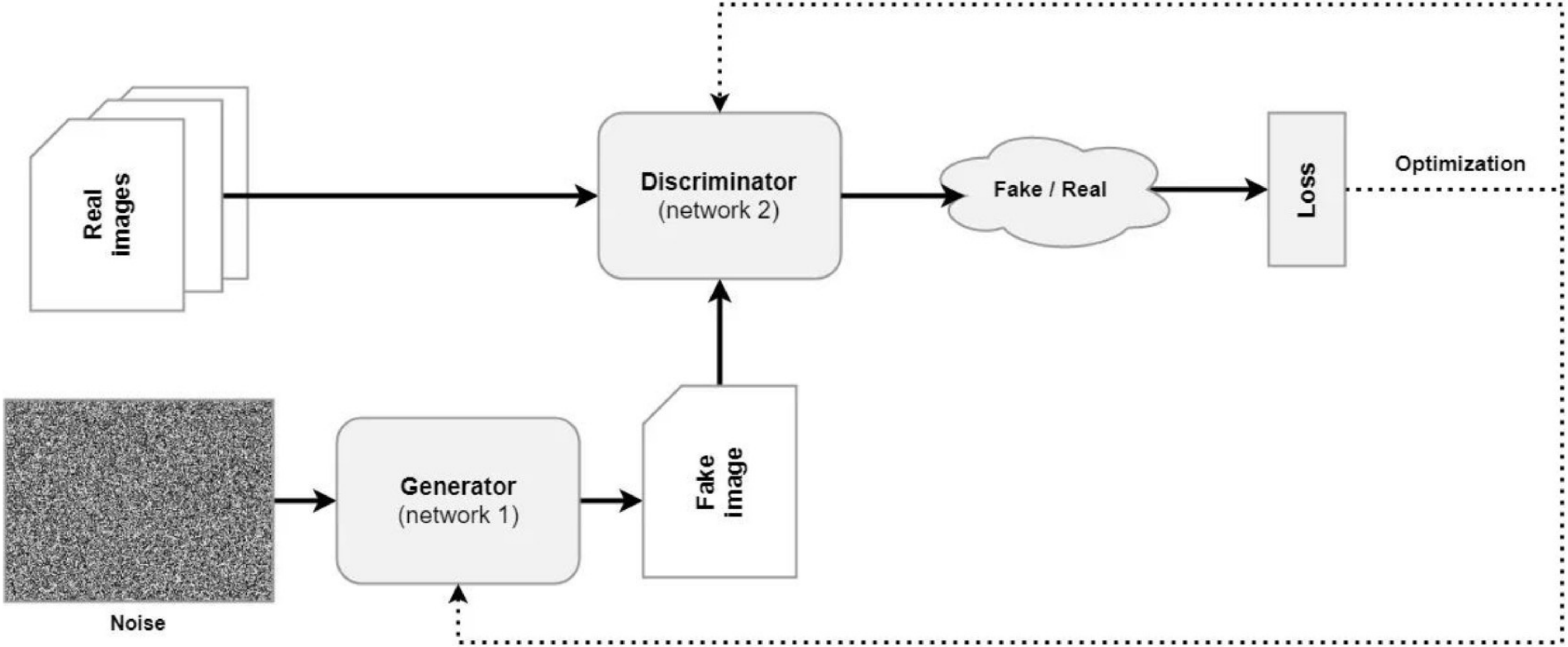
We aim to find the latent code $\tilde{\mathbf{w}}$ that minimizes our total loss function:

$$\tilde{\mathbf{w}} = \underset{\mathbf{w}}{\operatorname{argmin}}\, L_{\text{total}}(f(G(\mathbf{w})), \mathbf{y})$$

A key strength of this method is its ability to handle **composed degradations**. For a sequence of k multiple corruptions, $f_1, \ldots, f_k$, the objective is simply to minimize the distance over the entire chain of degradation approximations:

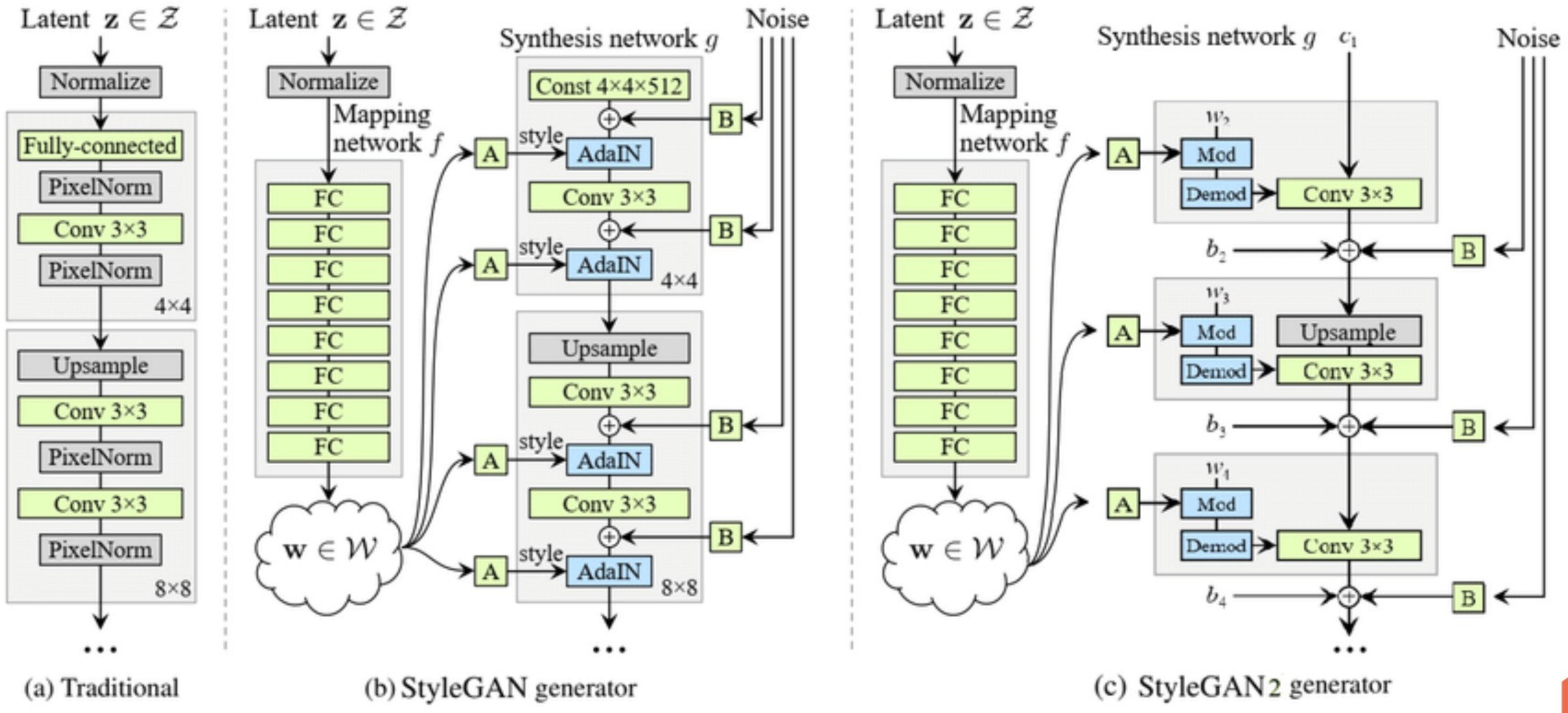
$$\tilde{\mathbf{w}} = \underset{\mathbf{w}}{\operatorname{argmin}}\, L_{\text{total}}((f_1 \circ \cdots \circ f_k)(G(\mathbf{w})), \mathbf{y})$$

Model Architecture:



- **Optimizer: Adam** Optimizer is used for all the three phases.

StyleGAN v2 Model:



Loss Functions:

Motivation:

- **Perceptual Loss (L_{LPIPS}):**
  Uses the LPIPS metric based on deep VGG features to measure perceptual similarity between the restored and target images.
  Motivation: Pixel metrics (like L2) often fail to reflect human visual judgment, whereas LPIPS aligns better with perceived image quality — making it the primary driver of realism.

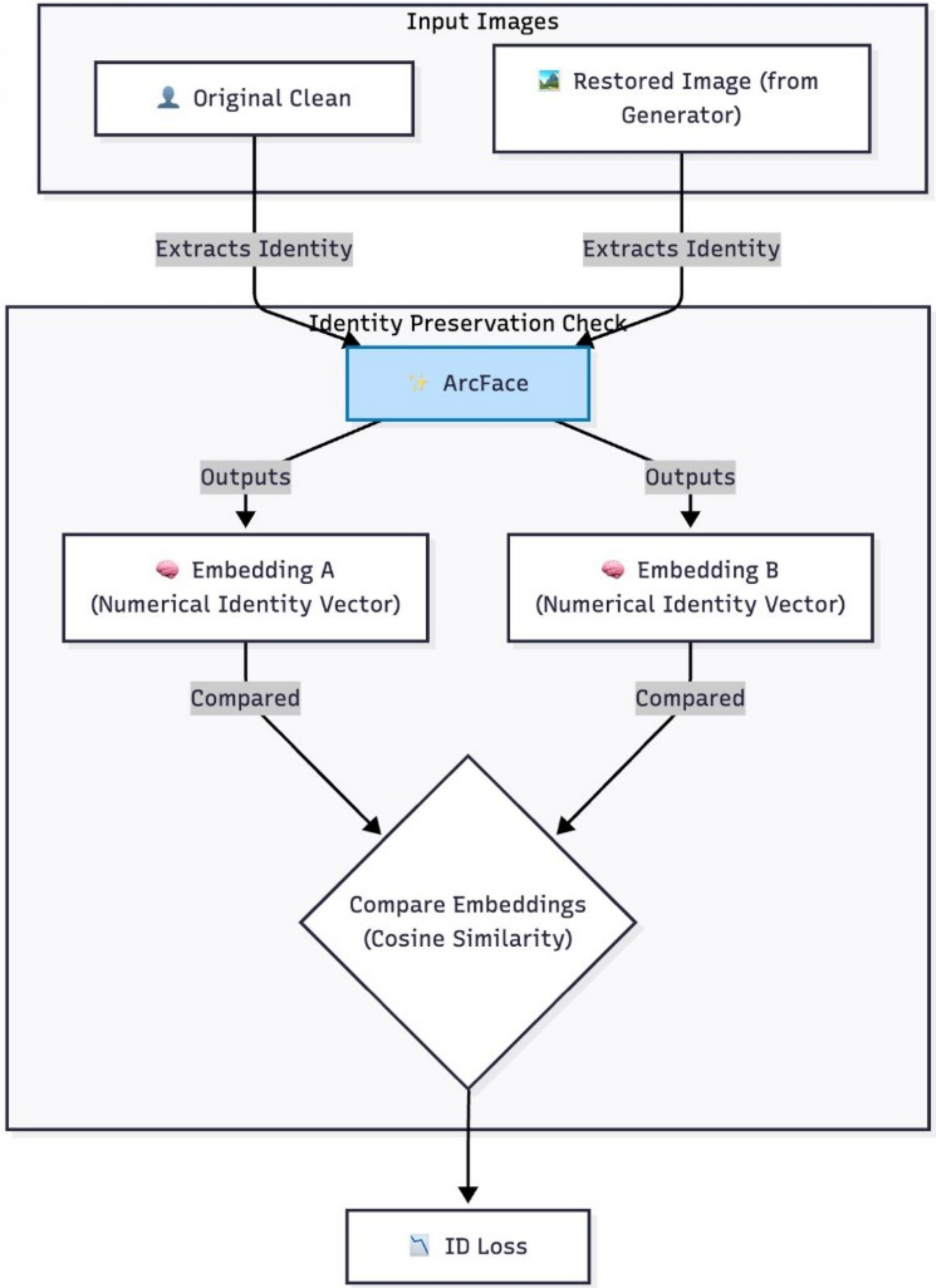
- **Pixel-wise Loss (L_{L1}):**
  Calculates the mean absolute pixel difference to ensure structural and color accuracy. (λ_{L1}=0.1)
  Motivation: Complements LPIPS by preserving fine-grained details and correcting local pixel errors that perceptual loss might overlook.

New loss introduced: <u>Identity Loss (L_{ID})</u>

- **Implementation:** It utilizes the **ArcFace model ($f_{ID}(\cdot)$)**, a pre-trained face recognition network, to generate robust facial feature embeddings.
- **Mechanism:** Minimizing L_{ID} maximizes the cosine similarity between the generated and input embeddings, guaranteeing the restored face has the **same identity** as the degraded input.

L_{total}, is a weighted sum of three distinct metrics:

$$\mathcal{L}_{total} = \mathcal{L}_{LPIPS} + \lambda_{L1}\mathcal{L}_{L1} + \lambda_{ID}\mathcal{L}_{ID}$$



$$L_{LPIPS}(x,y) = \sum_{l} \frac{1}{H_l W_l} \sum_{h,w} ||w_l \odot (f^l_{hw}(x) - f^l_{hw}(y))||_2^2$$

- Perceptual Loss (LLPIPS)

$$L_1(x,y) = \frac{1}{N} \sum_{i=1}^{N} |x_i - y_i|$$

- Pixel-wise Loss (LL1)

$$L_{ID}(x, y_{clean}) = 1 - \frac{E(x) \cdot E(y_{clean})}{||E(x)|| \cdot ||E(y_{clean})||}$$

- Identity Loss (LID)

Experimental Setup:

Evaluation Metrics and Baselines

Dataset & Degradations:

- **Dataset:** FFHQ-X (100 new test images aligned with StyleGAN's domain).
- **Tasks:** Upsampling, Denoising (non-linear Poisson/Bernoulli), Deartifacting (JPEG), Inpainting (random strokes).
- **Levels:** Tested across 5 levels of severity (XS to XL) and their combinations.

Key Experimental Detail (Robustness Test):

- **Ours:** Uses a **single, fixed set of hyperparameters (LRs/Steps)** for *all* 20 individual tasks and all composed tasks.
- **Baselines:** PULSE and L-BRGM were **individually optimized** (tuned) for each (task, level) pair for best performance, creating a strong handicap for our method.

Metrics (Lower is Better for all):

- **Accuracy (LPIPS):** Prediction vs. Ground Truth (overall quality).
- **Realism (pFID):** PatchFID (pFID) measures image naturalness.
- **Fidelity (LPIPS):** Degraded Prediction vs. Degraded Target (match to input).

Results (I): Quantitative Comparison

Task (Level)	Metric	PULSE (Tuned)	L-BRGM (Tuned)	OURS (Fixed HP)
Upsampling (L)	Accuracy (LPIPS)	0.490	0.487	**0.501**
Denoising (L)	Accuracy (LPIPS)	0.457	0.481	**0.501**
Deartifacting (M)	Realism (pFID)	33.2	24.1	**15.4**
Composed (3)	Realism (pFID)	28.5	22.1	**20.1**

Key Findings:

- **Individual Tasks:** Our method, despite using a fixed hyperparameter set, achieves superior accuracy in the majority of individual tasks against individually tuned baselines.
- **Realism:** Our method consistently achieves significantly better realism (lower pFID) across nearly all scenarios.
- **Compositions:** We maintain high performance when tackling combined degradations.

Results (II): Qualitative Comparison

“Target”

“Ours”

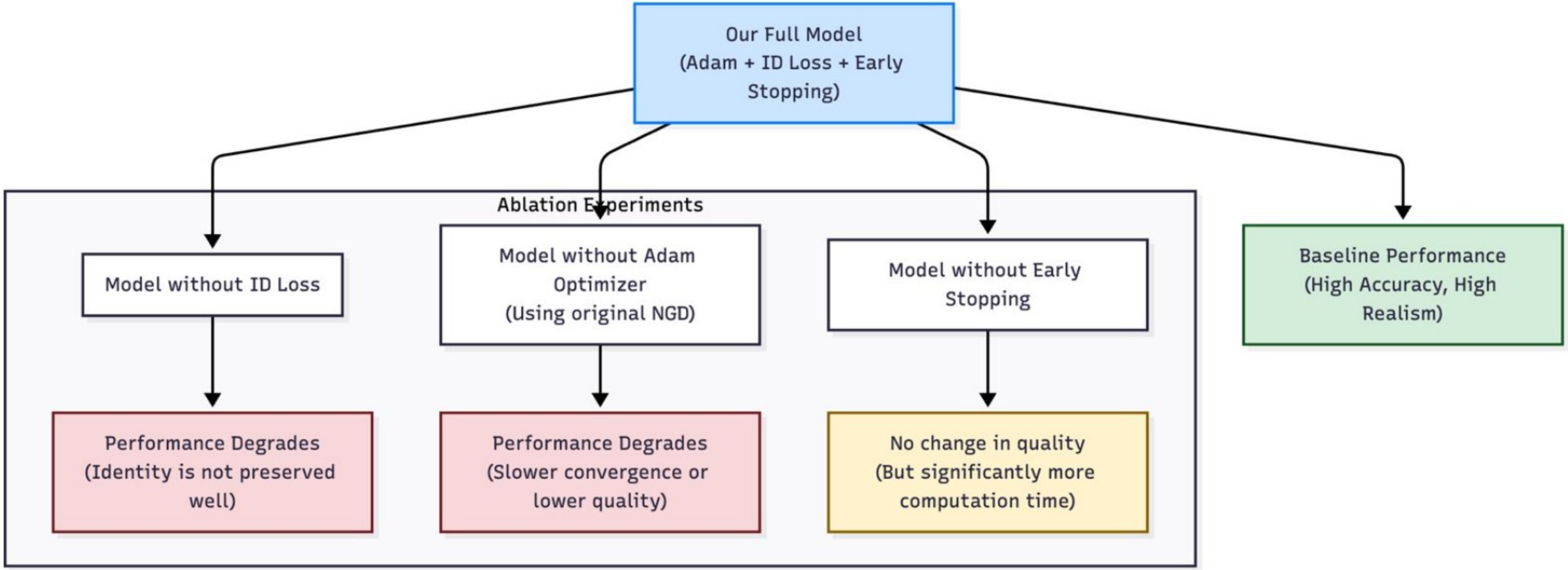
Ablation & Discussion:

Why Does it Work? (Avoiding Regularization)

- **Ablation/Discussion:**
- The **Progressive Extension** from , initialized at the mean, provides a strong **realism prior** (Phase I) that is only slowly perturbed by the higher fidelity focus of later phases.
- The **Conservative NGD Optimizer** acts as an implicit regularizer, preventing the solution from moving too far from the realistic manifold established in Phase I.

This combination achieves:

- **Strong Realism:** When degradations are high (relying on the prior).
- **High Fidelity:** When degradations are low (exploiting the space).
- **Generalization:**
- Tested successfully on other domains (LSUN, AFHQ-v2, BreCaHAD) at resolution, proving it is not face-specific.


Our Full Model
(Adam + ID Loss + Early Stopping)
Ablation Experiments
Model without ID Loss
Model without Adam Optimizer
(Using original NGD)
Model without Early Stopping
Baseline Performance
(High Accuracy, High Realism)
Performance Degrades
(Identity is not preserved well)
Performance Degrades
(Slower convergence or lower quality)
No change in quality
(But significantly more computation time)

Conclusion:

Contributions:

- Introduced a highly robust and simple 3-phase inversion framework for StyleGAN.
- Successfully replaced complex, tuned regularization losses with implicit regularization via progressive latent extension and conservative optimization (NGD).

Key Findings:

- Achieved competitive and SOTA results on a variety of individual and composed restoration tasks.
- Demonstrated performance using a single, fixed set of hyperparameters, validating the method's inherent robustness.
- Showed flexibility by tackling non-linear inverse problems more effectively than diffusion models.

Limitations:

- Restoration is inherently constrained by the domain on which the underlying StyleGAN model was trained.

Future Work:

- **Exploring Different Generative Priors:** While our model uses a StyleGAN2-ADA generator, this framework could be adapted to leverage other powerful generative models. Investigating newer architectures like StyleGAN-XL could allow for restoration on more diverse and unstructured datasets beyond human faces.
- **Scaling to Higher Resolutions and Compositions:** We can test the model's performance on higher-resolution images and more complex combinations of degradations. This would involve stress-testing the limits of the Adam optimizer and the enhanced ID loss to maintain both fidelity and realism under extreme conditions.

References:

Key Citations

- Poirier-Ginter, Y., & Lalonde, J.-F. (2023). **Robust Unsupervised StyleGAN Image Restoration**. *CVPR*.
- Karras, T., et al. (2020). **Analyzing and improving the image quality of StyleGAN**. *CVPR*.
- Conmy, A., et al. (2022). **Stylegan-induced data-driven regularization for inverse problems**. *Int. Conf. Acous. Sp. Sig. Proc.*.
- Kawar, B., et al. (2022). **Denoising diffusion restoration models**. *Adv. Neural Inform. Process. Syst.*.
- Zhang, R., et al. (2018). **The unreasonable effectiveness of deep features as a perceptual metric**. *CVPR*.