Motion Reconstruction from a Single Blurry Noisy Image using Deep Image Deblurring and Sequence Modeling

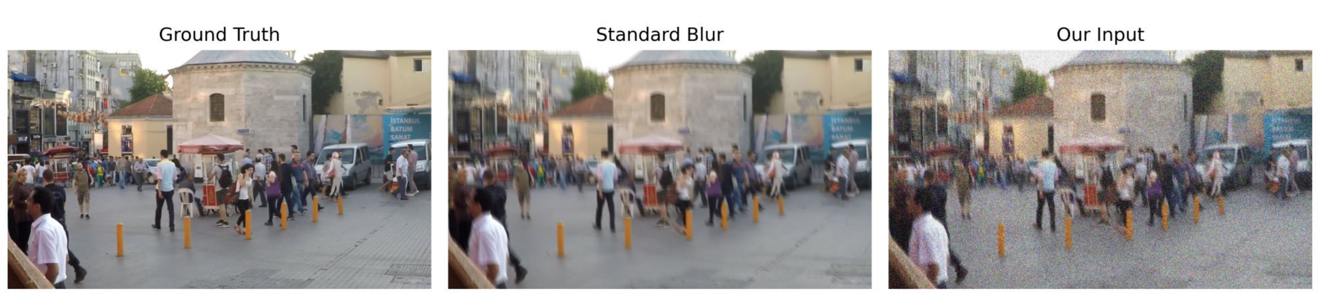
Mahika Annie Verghese, Shreya Shree Srikanth, Tetiana Ufimtseva

Department of Computer Science, University of Toronto

Motivation

- Motion blur is a common challenge in many real-world settings, from low-light surveillance and crowded urban scenes to everyday handheld photography. When only a single blurred frame is available, valuable temporal cues about human movement are lost, making it harder to reconstruct the original motion. This project aims to reconstruct short motion sequences from a single blurred and noisy image using a combination of image processing techniques and sequence modeling.
- Existing motion blur decomposition methods fail when input images contain sensor noise, producing temporally inconsistent and visually degraded frame sequences.
- In real-world surveillance settings, this blur is compounded by **additive noise**. Current methods fail when this noise is present:

$$x_{noisy} = x + \mathcal{N}(0, \sigma^2)$$



Results after the addition of noise:

Work	Mean PSNR (σ = 0)	PSNR (σ = 20)	% drop in quality
[1]	30.0241 dB	21.3247 dB	29.0%
[2]	30.2238 dB	18.4919 dB	38.8%
[3]	19.95987 dB	18.42723 dB	7.7%

State of the art models, designed for clean data, are severely compromised by real-world noise. Our evaluation shows the Animation From Blur [3] performs the most robustly to noise in comparison.

Related Work

We evaluated three baseline works on their native datasets: Learning to Extract a Video Sequence from a Single Motion-Blurred Image [1], Exposure Trajectory Recovery from Motion Blur [2], and Animation from Blur [3] to understand their strengths and limitations. Jin et al. [1] predicts a fixed 7-frame sequence from a single blurred GoPro image, addressing temporal ambiguity via order-invariant losses. Zhang et al. [2] recovers continuous exposure trajectories with a physically interpretable model. Animation from Blur [3] introduces a motion guidance representation and a two-stage network to generate multi-modal sequences.

Across all three works, the core limitations are shared: they struggle with large or complex blur, and none of the original models explicitly handle noise. We address one of these problems by incorporating noise modelling into our approach.

Among these baselines, the Animation from Blur model showed strong generalisation: even without being trained on GoPro dataset its performance degraded only marginally under increasing Gaussian noise (σ = 5, 10, 20), unlike the other models.

References

[1] Jin, Meishvili, Favaro, Learning to Extract a Video Sequence From a Single Motion Blurred Image, CVPR, 2018

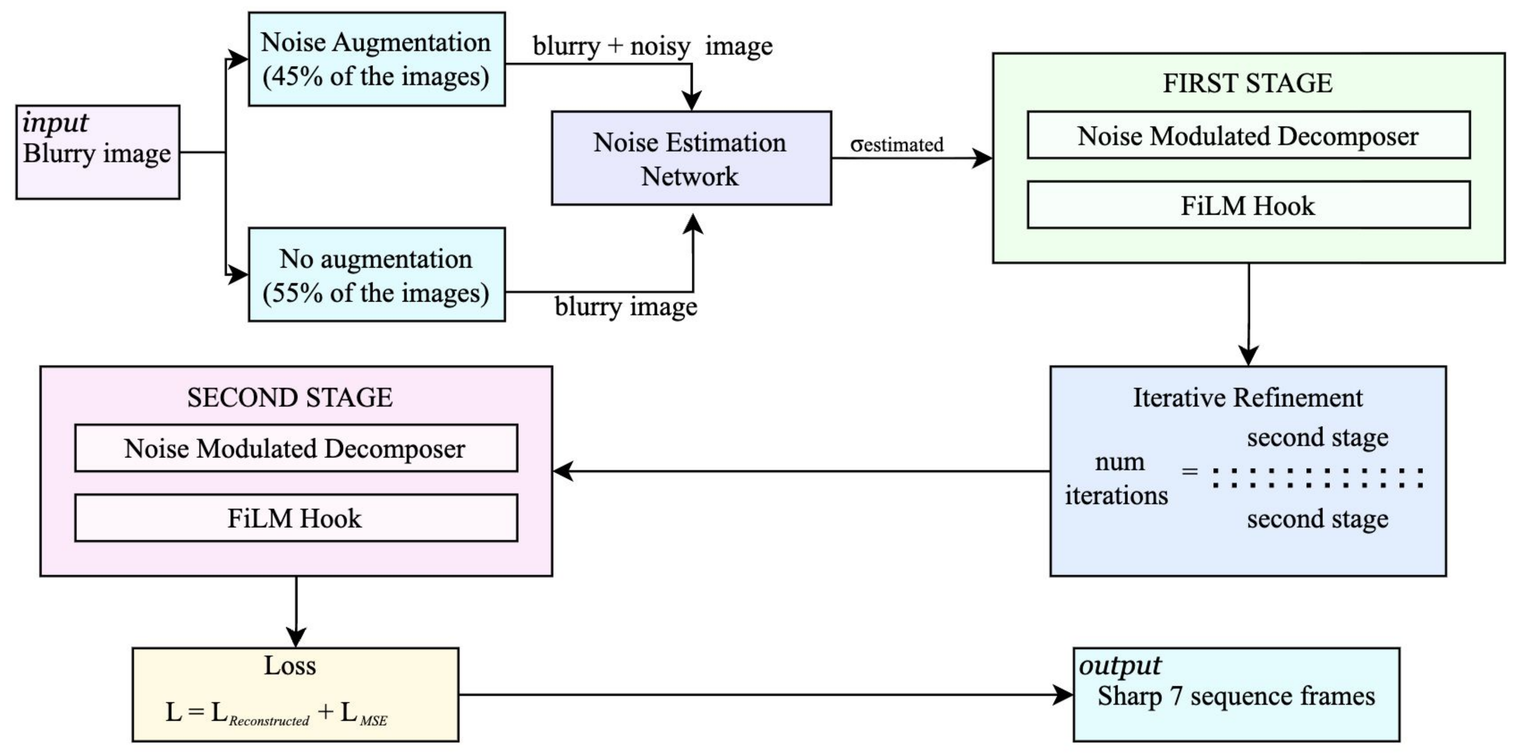
[2] Zhang, Wang, Maybank, Tao, Exposure trajectory recovery from motion blur, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021

[3] Zhong, Sun, Wu, Zheng, Lin, Sato, Animation from blur: Multi modal blur decomposition with motion guidance, European Conference on Computer Vision, 2022

New Technique

We propose three independent enhancements to improve motion blur decomposition. The first algorithm aims to improve the visual quality and smoothness of the reconstructed animation, while the second 2 focus on the model performance improvements under noisy conditions and the stabilisation of the reconstructed frame sequences.

1. The original code included a perceptual loss module that was inactive; we enabled it and added a temporal consistency loss to reduce jitter and enforce smoother frame evolution.
2. Next, we modified the model inference algorithm to include a Restormer-based pre-denoising stage before decomposition of the noisy blurry inputs, addressing the limitation that existing methods assume clean, noise-free inputs.
3. As the next step, we designed a **Noise-Aware Motion Blur Decomposition** (NAMBD) model that is based on the noise-adaptive feature modulation mechanism with Feature-wise Linear Modulation (FiLM), where a noise estimator predicts the noise level and conditions scale-shift parameters that dynamically modulate CNN features, allowing the network to adapt to both clean and noisy inputs.
4. To further strengthen our noise robust solution, we combined the second and third approaches so that in the inference mode the model first uses Restomer to denoise the input and then apply the Noise-Aware Motion Blur Decomposition.



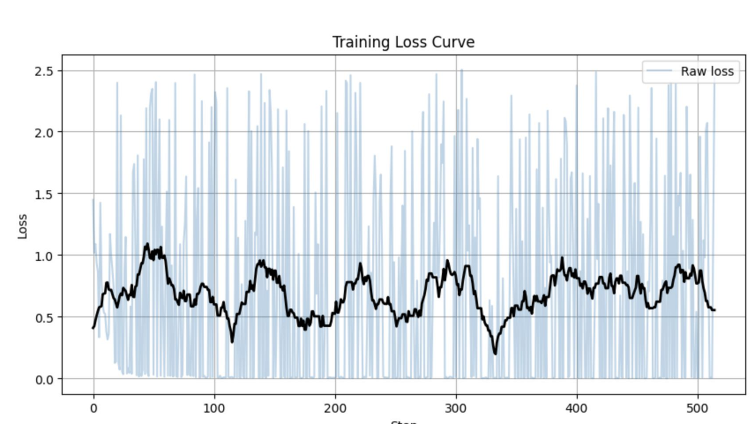
Experimental Results

- NAMBD achieves the best performance on clean data (σ=0: PSNR 24.75 dB), showing the blur decomposition network works optimally without noise interference. However, performance degrades rapidly as noise increases, dropping 3.6 dB by σ=50, so it is not inherently robust to noisy inputs.
- The two-stage approach outperforms standalone NAMBD at moderate-to-high noise levels (σ ≥ 20). At σ = 30, the Restormer + NAMBD model achieves 23.65 dB PSNR versus 23.25 dB for NAMBD alone. This crossover point suggests denoising preprocessing is beneficial when input noise exceeds ~20σ.
- At σ = 20, both models achieved their best LPIPS (NAMBD: 0.24; Restormer + NAMBD: 0.3) and then started to degrade as noise increased. Overall, the two-stage pipeline maintains more consistent perceptual quality on high noise levels and NAMBD performs better on low noise.

  Both of the suggested architectures outperform the original Animation from Blur model even on clean inputs, confirming that our modifications strengthen the baseline.

Noise level (σ)	NAMBD			Restormer + NAMBD		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
0	24.75	0.764	0.268	22.06	0.64	0.43
5	24.71	0.762	0.266	22.74	0.69	0.39
10	24.59	0.754	0.259	24.25	0.74	0.31
20	24.10	0.718	0.24	24.15	0.73	0.30
30	23.25	0.638	0.294	23.65	0.65	0.32
40	22.23	0.557	0.376	23.00	0.65	0.33
50	21.15	0.487	0.459	22.05	0.60	0.40

Spatio-Temporal Perceptual Loss		
PSNR	SSIM	LPIPS
28.6758	0.8797	0.1484



The input blurry noisy image

The reconstructed sequence of sharp images