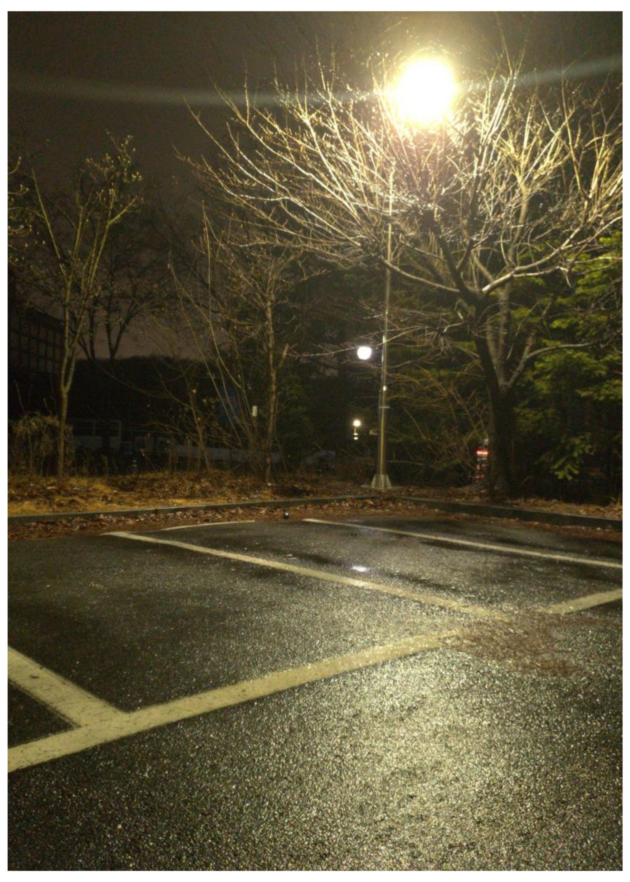
Frequency-Aware Knowledge Distillation for Efficient Under-Display Camera Image Restoration

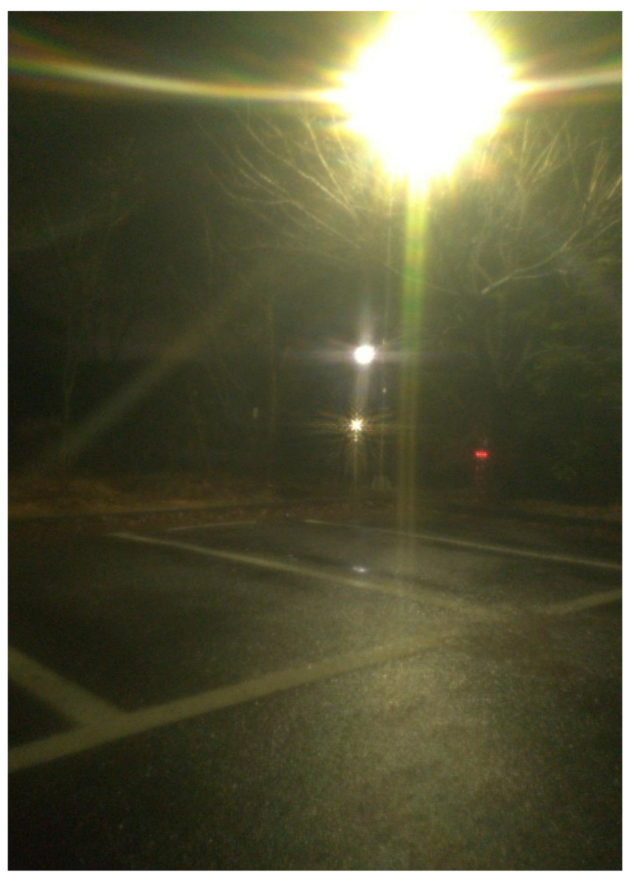
Navdeep Singh, Mohammed Junaid A. Q., Abhinav Shankar H.
Department of Computer Science, University of Toronto

Motivation

- **UDC Degradations**: Under-display cameras suffer from flare, blur, color shifts, and non-uniform noise introduced by the display stack, especially in low-light and high-contrast scenes.
- **Gap in Deployment**: Large, frequency-aware restorers (e.g., MambaIR with FFT blocks) can handle these artifacts, but their memory and latency make them impractical for real-time mobile deployment.
- **Our Goal**: Distill a frequency-aware MambaIR teacher into a compact U-Net student using pixel, perceptual, feature, and FFT-based losses (low-frequency, amplitude + phase, multi-scale in Model 2) so we approach teacher-level quality with an on-device-friendly model.



Ground Truth



UDC Input

Related Work

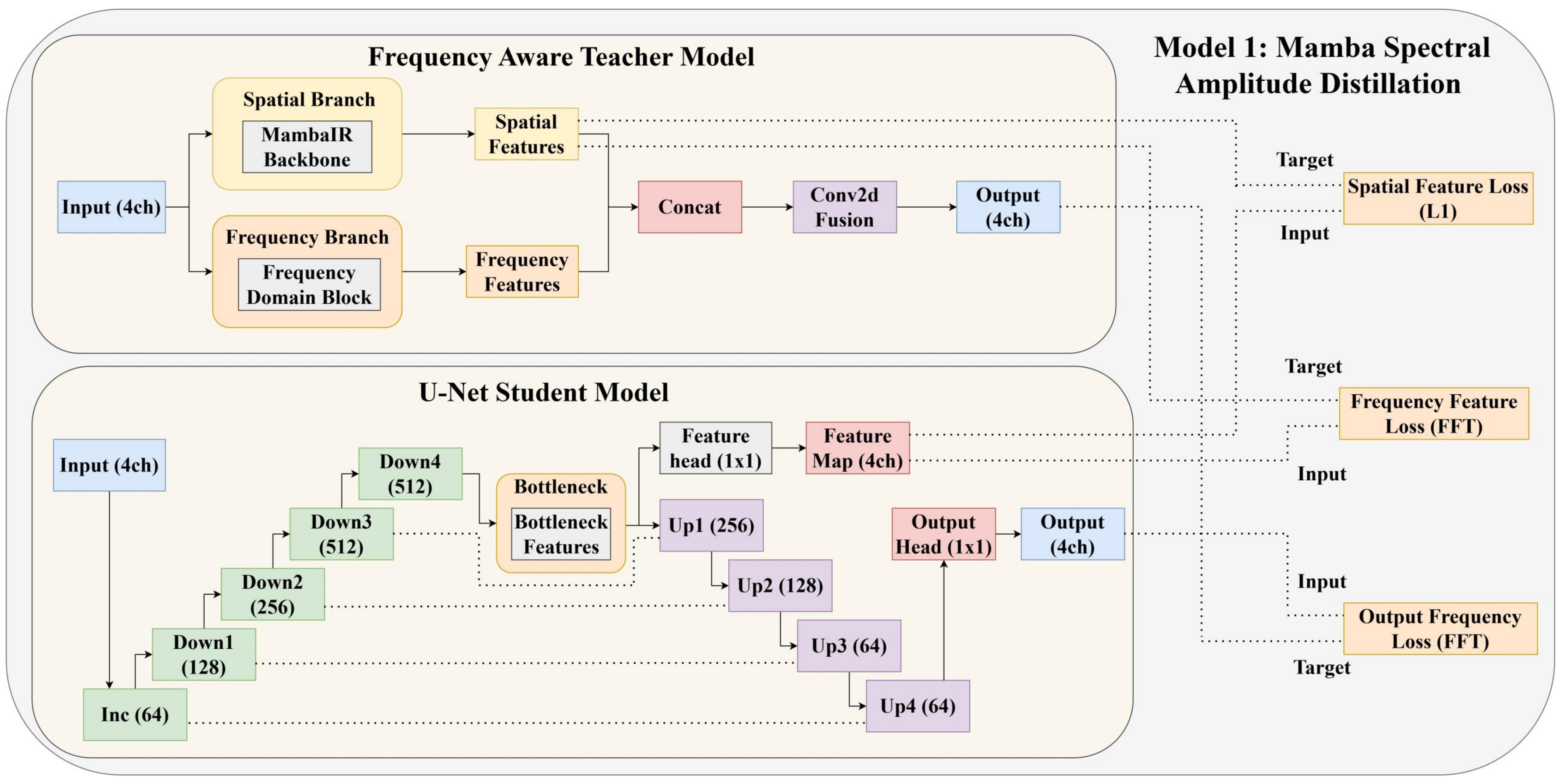
- **UDC Restoration Models**: Prior UDC methods use CNN/ Transformer architectures trained on paired UDC-clean data, achieving strong PSNR/SSIM but typically with heavy backbones and high compute cost.
- **State-space & Frequency Modeling**: Models like MambaIR [1] and frequency-aware modules show that modeling long-range context and spectral patterns improves restoration, but these designs are mostly used in large teacher-style networks rather than lightweight students.
- **Knowledge Distillation**: Existing distillation for restoration mainly matches output pixels and spatial features; frequency-domain and phase-aware KD are rare. Our, work combines a frequency-aware teacher with low-frequency, multi-scale FFT distillation to transfer flare-handling ability into a lightweight U-Net student

References

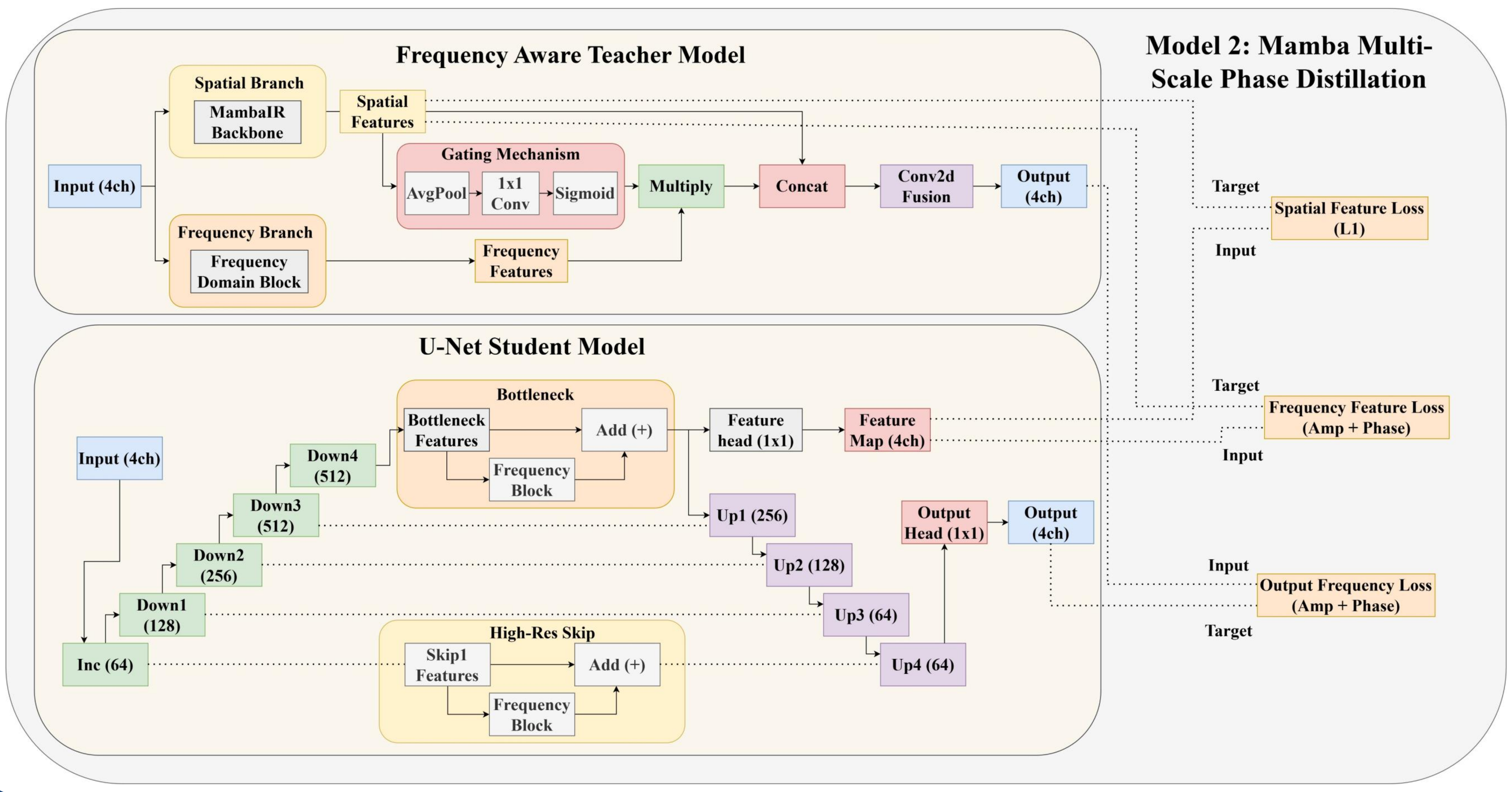
[1] Guo et al. Mambair: A simple baseline for image restoration with state-space model. In European conference on computer vision 2024, Springer Nature Switzerland.
[2] Ahn et al. UDC-SIT: A real-world dataset for under-display cameras. Advances in Neural Information Processing Systems. 2023.
[3] Ahn et al. Integrating spatial and frequency information for under-display camera image restoration. Pattern Analysis and Applications. 2025.
[4] Liu et al. Fsi: Frequency and spatial interactive learning for image restoration in under-display cameras. InProceedings of the IEEE/CVF International Conference on Computer Vision 2023.

New Technique

Restores image sharpness by reconstructing frequency intensity, coupling Mamba's global spatial context with spectral learning.



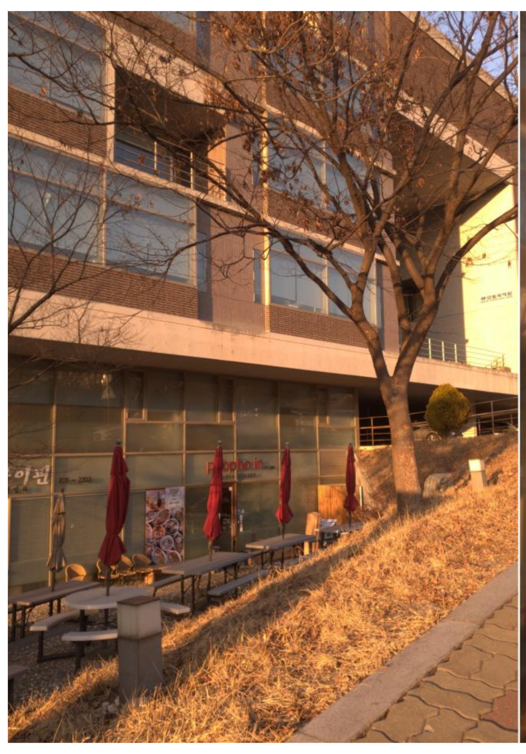
Recovers structural fidelity and fine details (like flares) by aligning both frequency amplitude and phase across multiple resolutions, gated with Mamba's global context.



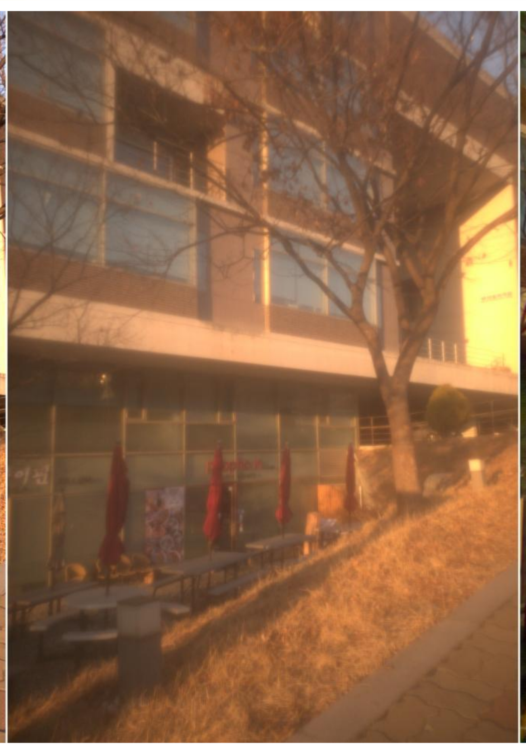
Experimental Results and Discussion

Results on UDC-SIT [2]

Model	PSNR (dB)		SSIM		Inference Time (ms)	
	Teacher	Student	Teacher	Student	Teacher	Student
Model 1	23.05	23.89	0.762	0.802	2586.95	61.20
Model 2	23.46	23.05	0.776	0.739	2587.49	93.49
Model 1 patch	26.74	26.58	0.815	0.810	17.57	3.73
Model 2 patch	26.24	26.42	0.812	0.792	79.77	6.90
FSI [4]	21.72		0.6674		-	
SFIM [3]	29.61		0.8758		-	



Ground Truth



UDC Input

Model 2 Teacher

Model 2 Student

- Model 2's teacher and student both significantly clean up the UDC input, with sharper details compared to the degraded raw image.
- The Model 2 student visually tracks the teacher very closely, while being much faster and suited for fast deployment on real devices.
- Model 2, generally preserves structure, and fine textures better while reducing flares giving more visually consistent restorations across different scenes.