Unpaired Ultrasound Domain Translation with Multi-Stage Residual GAN Refinement

Authored by
Mark Youssef, BSc

Supervised by
Isabel Drégely-Dávila, PhD

Co-Supervised by
Assist. Prof. Dr. techn. Christian Kollmann, PhD

Year of
2026

Introduction

- Ultrasound is portable, low-cost, and real-time, but speckle, attenuation, and acoustic shadowing can reduce boundary visibility and low-contrast tissue interpretation. [1]
- Handheld systems improve point-of-care access, but often differ from full-size clinical scanners in spatial detail, image fidelity, and device-dependent appearance; acquiring aligned cross-device image pairs of the same anatomy is rarely realistic. [2]
- **Objective**: perform unpaired handheld-to-full-size ultrasound domain translation while preserving diagnostically relevant structures and limiting unsafe changes in dark regions, weak boundaries, and vessel anatomy.

Methods

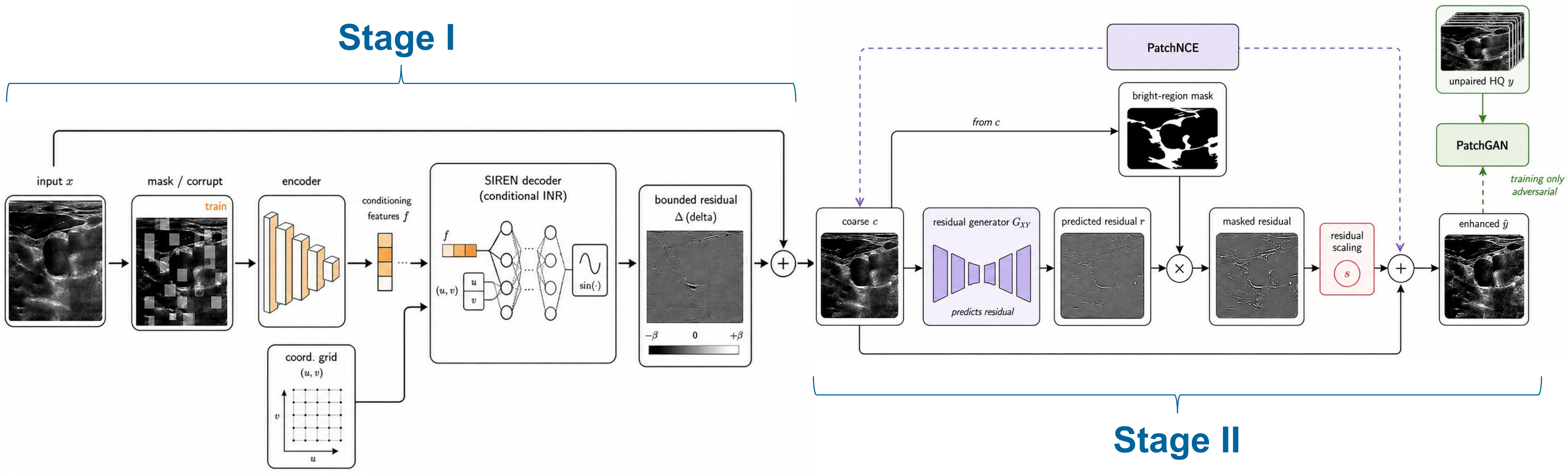

Figure 1: Two-stage framework: Stage I stabilizes the handheld input; Stage II performs residual domain translation toward the full-size clinical scanner domain

Stage I stabilizes the handheld input with a self-supervised Conditional INR; **Stage II** performs the main residual domain translation using PatchNCE, PatchGAN, and bright-region masking. Data: **7,160** full-size clinical-domain frames, **6,730** handheld-domain frames, and **1,024** external validation-style handheld frames.

Results & Discussion

- External FRD improved from **38.0** to **33.6**, an **11.5% reduction** over identity.
- KID stayed **stable** at **0.178**, indicating no clear loss of feature-space alignment.
- Classical **baselines** showed **higher FRD** than the proposed Full Pipeline.
- Stage I **reduced** LPIPS, Edge Coherence, and Dark Preserve compared with training without Stage I.
- **Main finding**: Stage I stabilizes training; Stage II performs the main residual domain translation.
- **Main limitation**: conservative harmonization only — not ground-truth reconstruction.

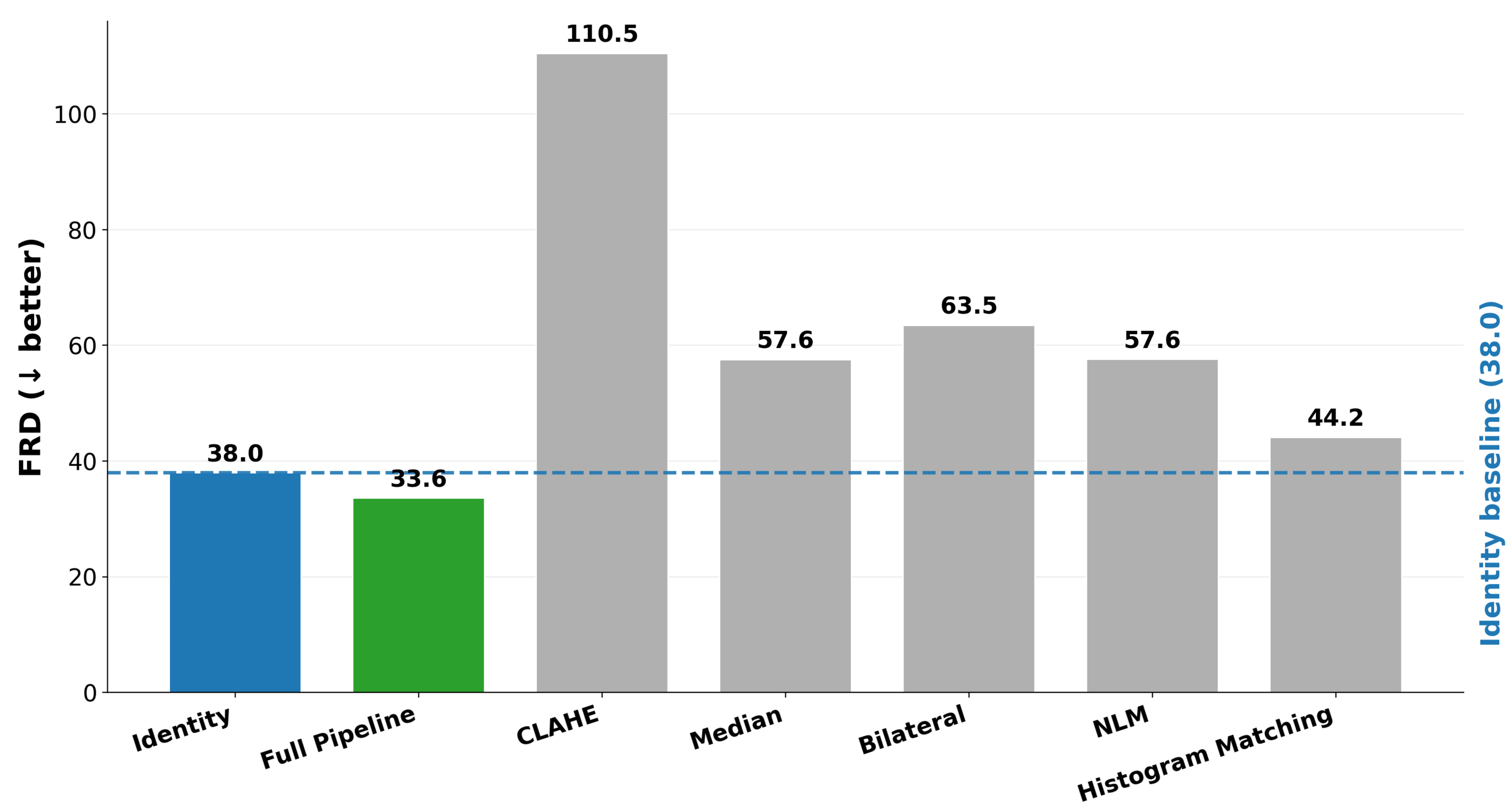

Figure 2: External FRD comparison (n = 673). Full Pipeline reduced FRD from 38.0 to 33.6 while KID remained stable at 0.178; per-image external metrics differed from identity (Wilcoxon + BH, $p < 0.0001$).

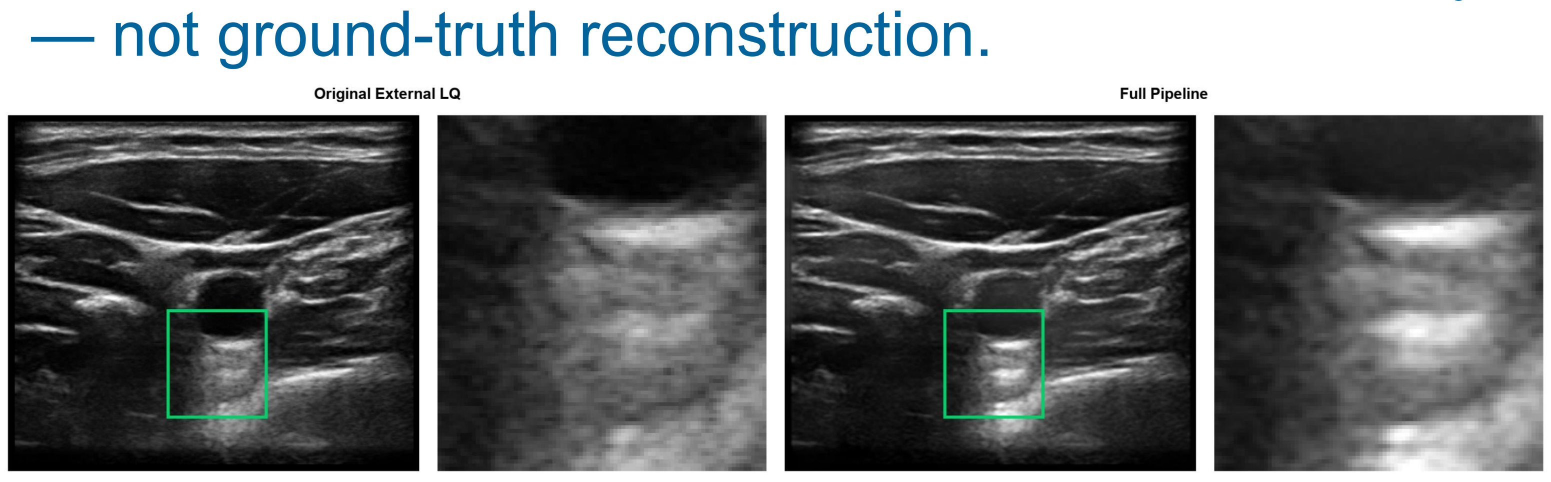

Figure 3: External example before and after enhancement

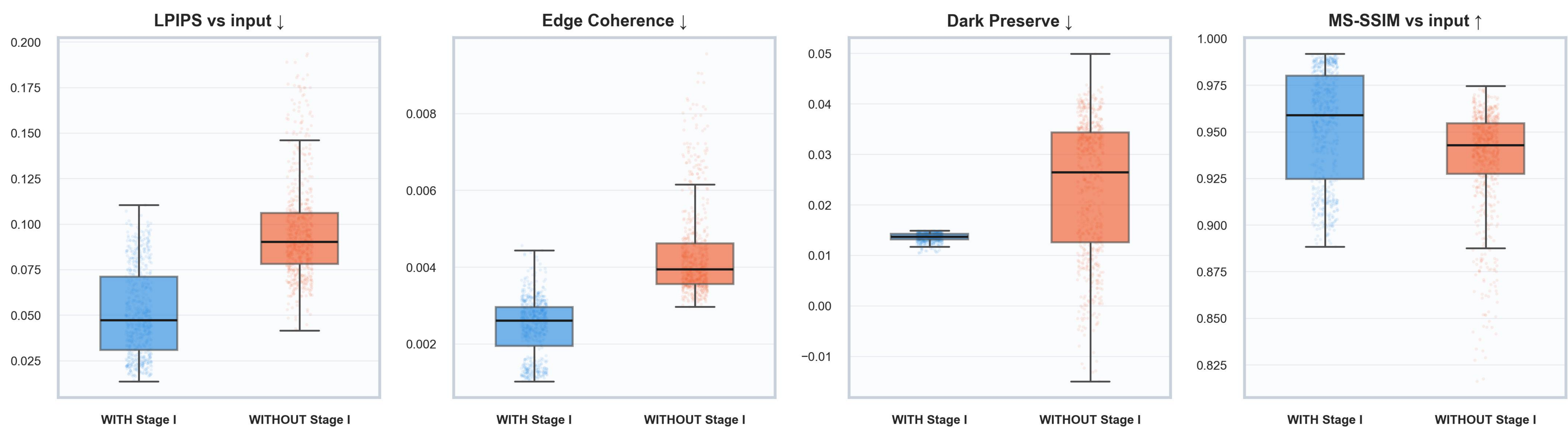

Figure 4: Stage I ablation (n = 673). Training with Stage I reduced LPIPS, unsupported edges, and dark-region artifacts (Wilcoxon + BH correction, $p < 0.0001$)

References

[1] S. H. Contreras Ortiz, T. Chiu, and M. D. Fox, “Ultrasound image enhancement: A review,” Biomedical Signal Processing and Control, vol. 7, no. 5, pp. 419–428, 2012.[2] H. G. A. van der Pol, L. M. van Karnenbeek, M. Wijkhuizen, F. Geldof, and B. Dashtbozorg, “Deep learning for point-of-care ultrasound image quality enhancement: A review,” Applied Sciences, vol. 14, no. 16, 2024.