# Suppressing Streak Artifacts Generated by the Interference of Imaging and Therapy Fields: Initial Findings Using a Hybrid U-Net and Diffusion Model

IEEE SAUS 2024
IEEE SOUTH ASIAN ULTRASONICS SYMPOSIUM

Shreyans Jain, Ruchika Dhawan, Nishitakumari Prafulchandra Mistry, Himanshu Shekhar

Discipline of Electrical Engineering, Indian Institute of Technology Gandhinagar, Gujarat, India
agarwalmihir@iitgn.ac.in, shreyans.jain@iitgn.ac.in, ruchikadhawan@iitgn.ac.in, nishita.m@iitgn.ac.in, himanshu.shekhar@iitgn.ac.in

MUSE LAB
MEDICAL ULTRASOUND ENGINEERING LAB

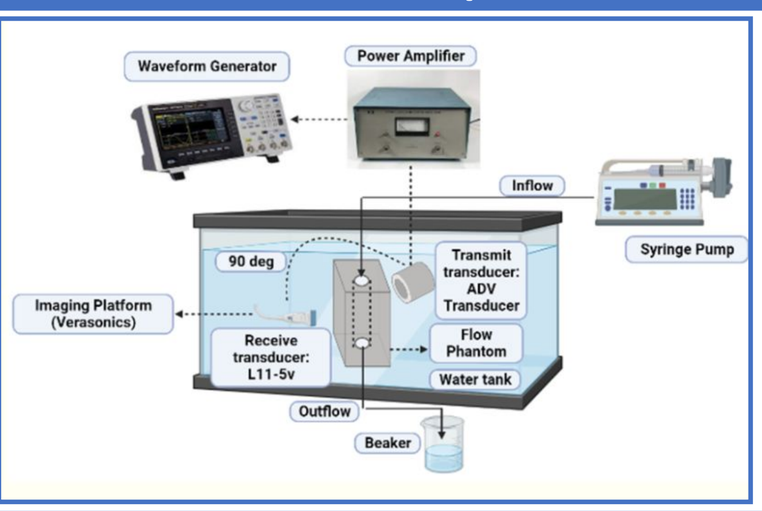
## Background and Motivation

- Therapeutic ultrasound applications include ablation, blood brain barrier opening, and drug delivery.
- B-mode ultrasound is used typically for image guidance and monitoring of therapy.
- When the therapy transducer is on, streak artifacts are produced in images because of interference between therapy and imaging fields.

## Methodology

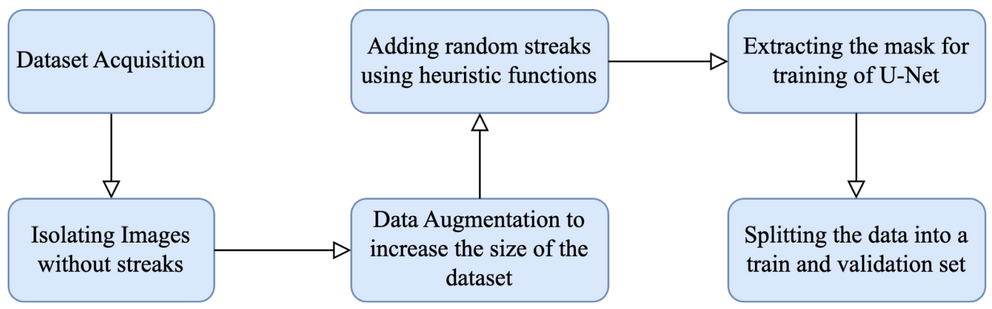
Data Acquisition from experiments → Synthetic Data Generation → Mask Generation from U-Net → Inpainting using Diffusion Model

## Experimental Setup



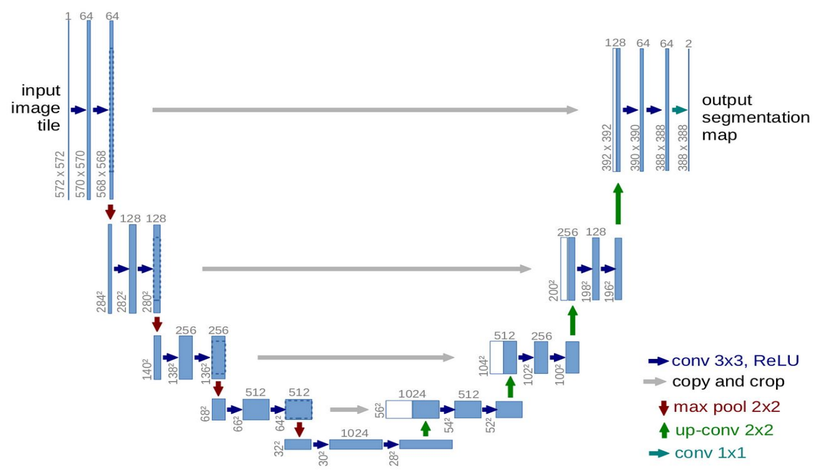

Insonating a wall-less polyvinyl alcohol phantom perfused with perfluorohexane droplets at 5 x $10^9$ droplets/ml undergoing phase transition.

## Generation of Synthetic Dataset




Model trained by artificially introducing streaks into 20 streak-free ultrasound images and obtained 1638 images.

## U-Net Architecture for Segmentation

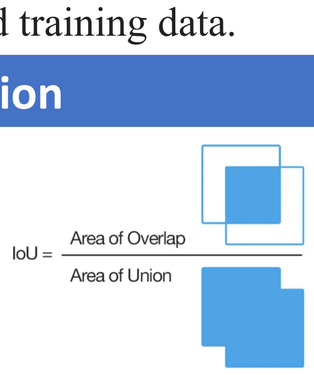



U-Net Architecture [1]

- The UNet is a neural network based on an Encoder-Decoder structure.
- The encoder (contracting path) extracts high-level features, while the decoder (expanding path) reconstructs segmentation maps.
- Suitable for biomedical images with limited training data.

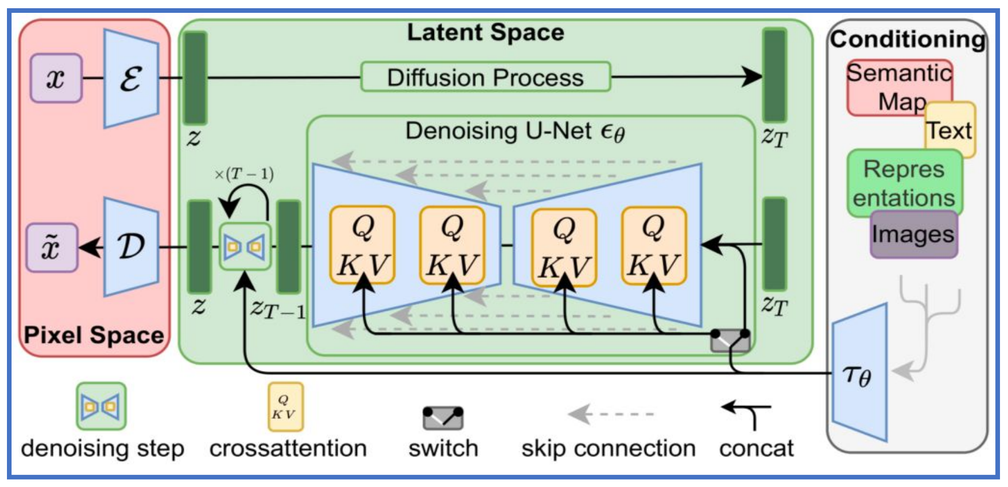
## Metric for Segmentation

IOU score (Intersection of Union), a metric quantifying overlap accuracy between predicted and actual segmentation




IOU Score [2]

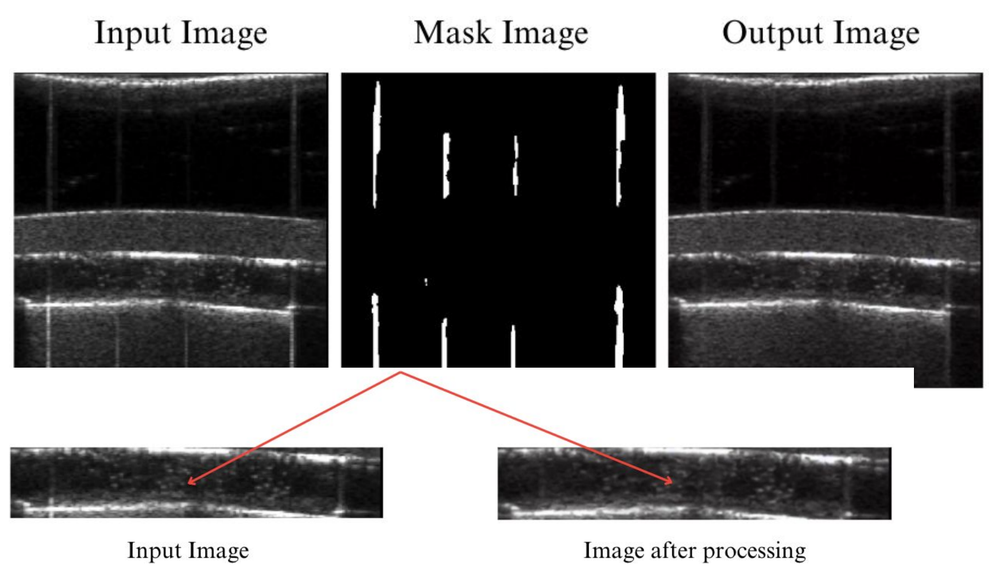
## Diffusion Models




Latent Diffusion Model Architecture [5]

- It is a denoising U-Net architecture
- The model takes three inputs: a noisy image, a mask and a conditioning signal (mostly text).
- It processes a latent space and a text-based prompt as inputs, denoising the image via a diffusion process.
- The conditioning signal provides additional information to guide the denoising process.
- Skip connections in U-Net link encoder and decoder layers, preserve fine details in the image during the denoising process.
- We used text prompts like “blend” to achieve the best inpainting results when a mask was given.

## Results




- Our results demonstrate the ability of our model to remove streak artifacts while preserving contrast from microbubbles.
- We see an increase in the Signal-to-Noise ratio by 6.0 ± 3.4 dB (n=154 frames)
- The model's rapid convergence underscores its potential for real-time clinical application.
- We reported an IOU score of 0.698.

## Summary and Future Work

- We integrated U-Net with diffusion models which provides a potential solution for streak artifact removal in therapeutic ultrasound imaging.
- Future work will focus on validating the model with more extensive and diverse datasets and reducing inference times.

## References

1. Source: Ronneberger et al., U-net, MICCAI 2015
2. Source: https://medium.com/@ankuraga/iou-score-and-its-variants-for-deep-learning-bd09ecf14832
3. Ronneberger et al., U-net, MICCAI 2015
4. Rombach et al., Latent Diffusion, CVPR 2022
5. Source: Ulhaq, Anwaar & Akhtar, Naveed & Pogrebna, Ganna. (2022). Efficient Diffusion Models for Vision: A Survey. 10.48550/arXiv.2210.09292.