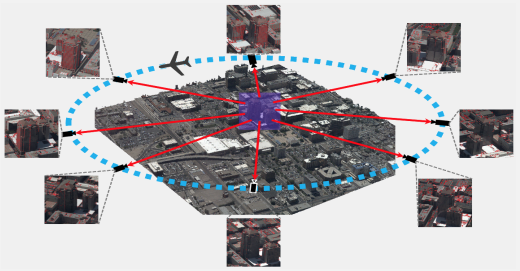
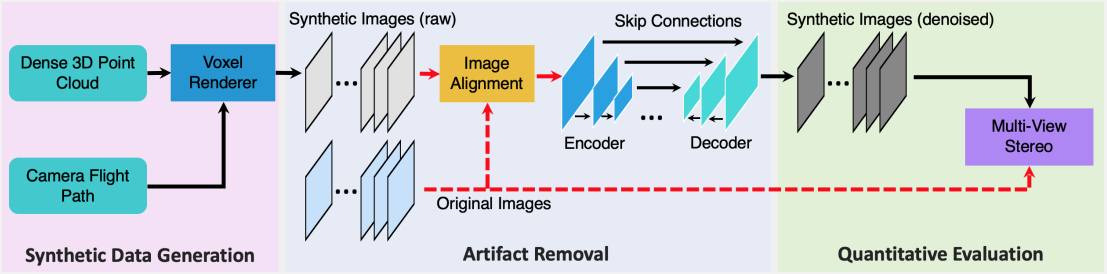
Deep Realistic Novel View Generation for City-Scale Aerial Images

Koundinya Nouduri*, Ke Gao*, Joshua Fraser, Shizeng Yao, Hadi AliAkbarpour, Filiz Bunyak, Kannappan Palaniappan
University of Missouri-Columbia, MO, USA

Overview

- Introduced a novel end-to-end framework for generation of large scale synthetic aerial image sequences with associated precise ground truth camera metadata
- Proposed a novel deep learning network with an explicit edgemap processing stream to remove image artifacts





ARNet-Edge: Edge-Augmented Artifact Removal Network

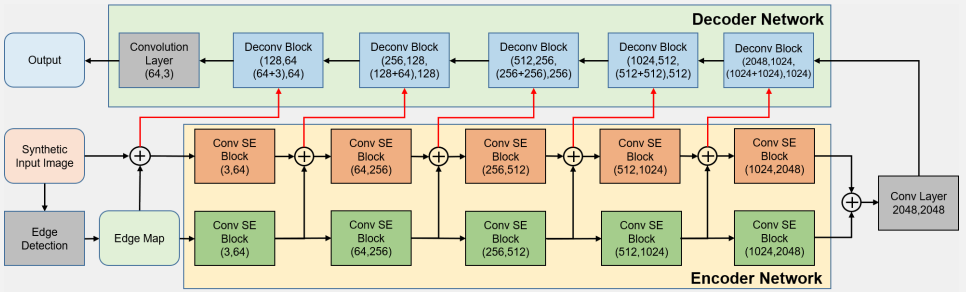
Custom U-Net like encoder-decoder network. Learns to map 2D raw synthetic images to realistic-looking synthetic images with reduced artifact

- **Encoder:** two parallel series of Squeeze and Excitation Resnet (SE-Resnet Blocks). **Stream-1:** extracts features from RGB input; **Stream-2:** extracts features from edge-map
- **Decoder:** deconvolution layers followed by two convolution layers attempting to reconstruct the original image from encoders features
- **Combined loss:** mean square error + structural similarity

$$SSIM(gt, gi) = \frac{(2\mu_{gt}\mu_{gi} + C_1) + (2\sigma_{gtgi} + C_2)}{(\mu_{gt}^2 + \mu_{gi}^2 + C_1)(\sigma_{gt}^2 + \sigma_{gi}^2 + C_2)}$$

$$L_{ssim}(gt, gi) = \frac{1}{N}\sum^{N}(1 - SSIM(gt, gi))$$

$$L_{mse}(gt, gi) = \frac{1}{N}\sum_{i=0}^{N}(gt - gi)^2$$



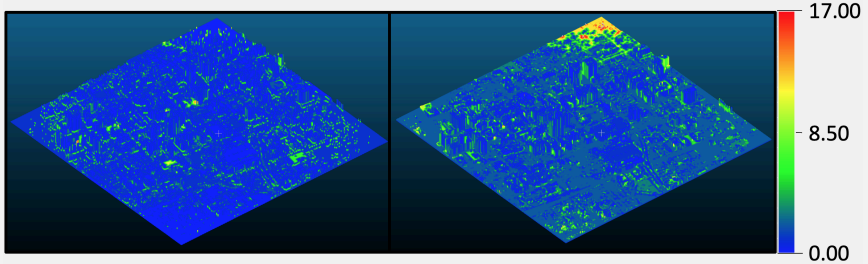
Experimental Results

1. Image Quality

Model	SSIM		PSNR	
	ABQ	LA	ABQ	LA
Raw Synthetic	0.573	0.651	20.65	23.53
REDNet	0.629	0.686	22.29	24.35
Deep Image Prior	0.619	0.710	21.12	25.21
ARNet	0.638	0.704	22.40	25.15
ARNet-Edge	0.665	0.721	23.21	25.89

2. Dense 3D Point Cloud Reconstruction

- Reconstructed a dense point cloud from the denoised ABQ image sequence (PC_d) and from the raw ABQ image sequence (PC_r); point cloud generated from the original aerial image sequence as the ground truth (PC_o)
- Computed cloud-to-cloud distance



PC_d vs PC_o PC_r vs PC_o