

SAR to EO Image Translation

Team Neural Vanguards



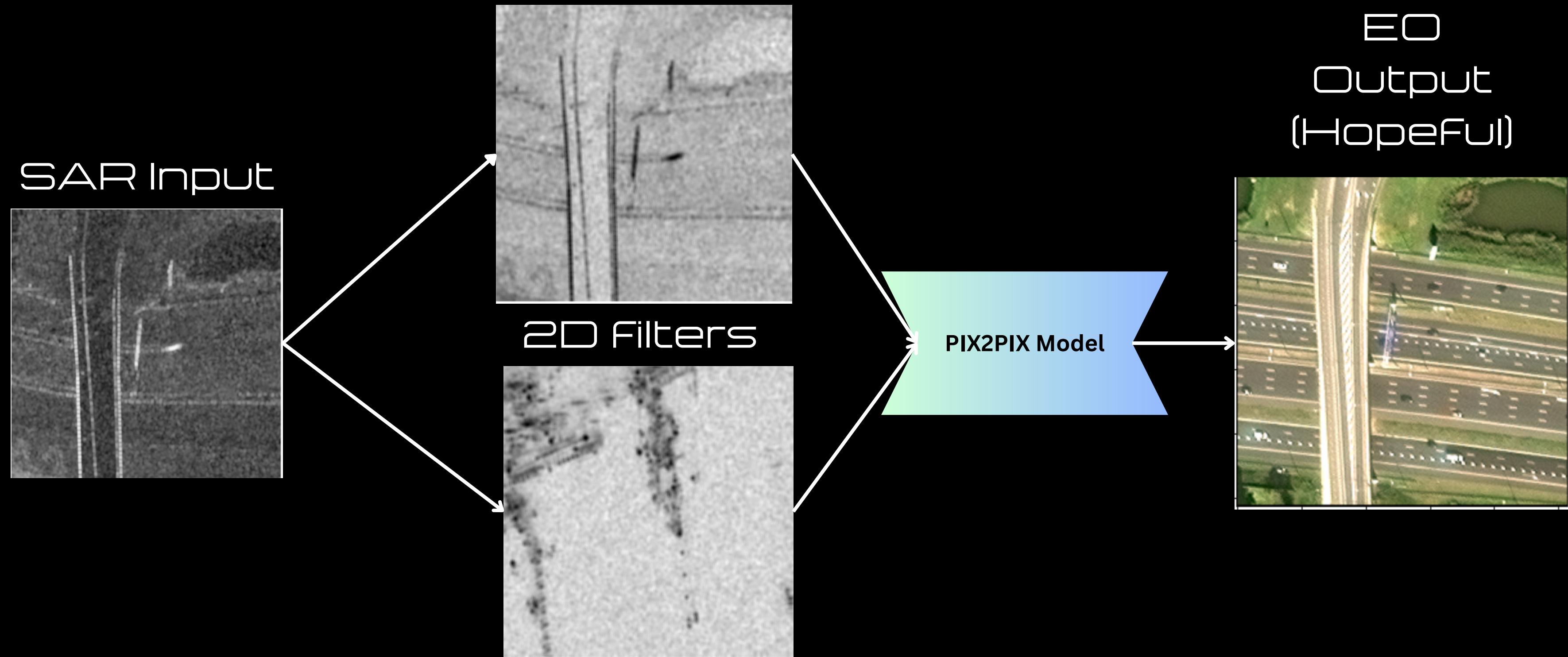
Methodology (Pix2Pix)

- SAR2EO Approach: Using couple filters to add more info about the SAR image which is then fed to the Pix2Pix Model and trained to output the EO image (ground truth).

Model: Pix2Pix, short for Picture-to-Picture, is a type of conditional Generative Adversarial Network (cGAN) designed for tasks like image-to-image translation, where the goal is to learn a mapping from an input image to an output image.

- Preprocessing: Resizing , random cropping and augmentation (called random jittering) are performed on the training data to make the model learn the patterns and task more robustly.

Methodology:



Evaluation

Pix2Pix (and us) uses a combined loss function, incorporating Adversarial Loss and L1 Loss.

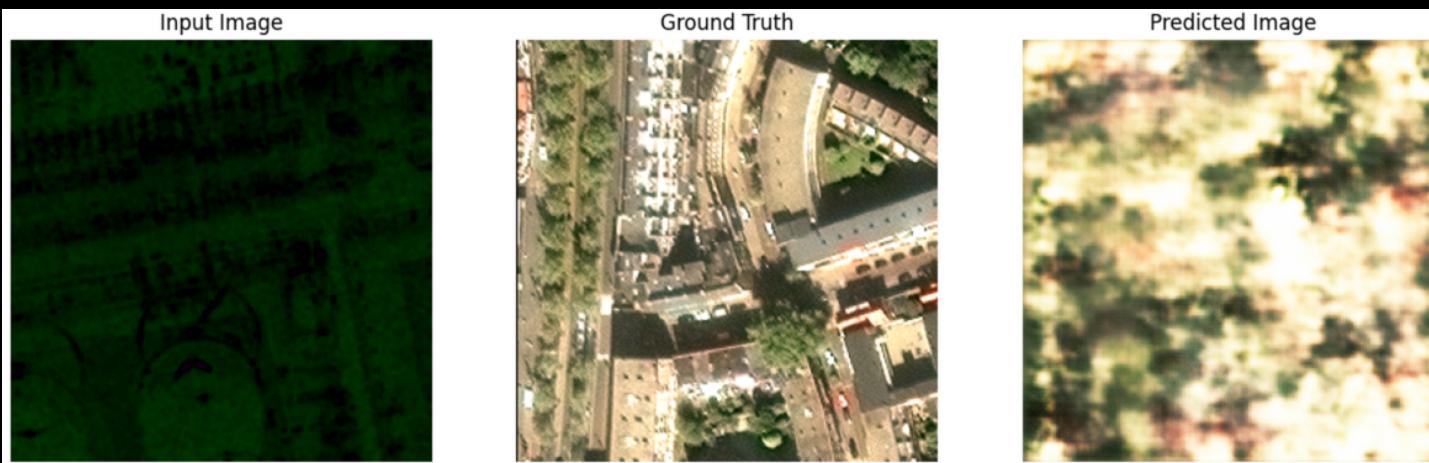
Adversarial Loss: Measures how well the Generator can fool the Discriminator into thinking generated images are real. A lower adversarial loss indicates more realistic images, enhancing the model's ability to create believable outputs.

L1 Loss: Calculates the mean absolute difference between the generated images and the real target images. Encourages the model to maintain structural and content accuracy, making sure generated images closely match the target.

Combined loss ensures generated images are not only realistic (Adversarial Loss) but also accurate representations of the input images (L1 Loss).

Results

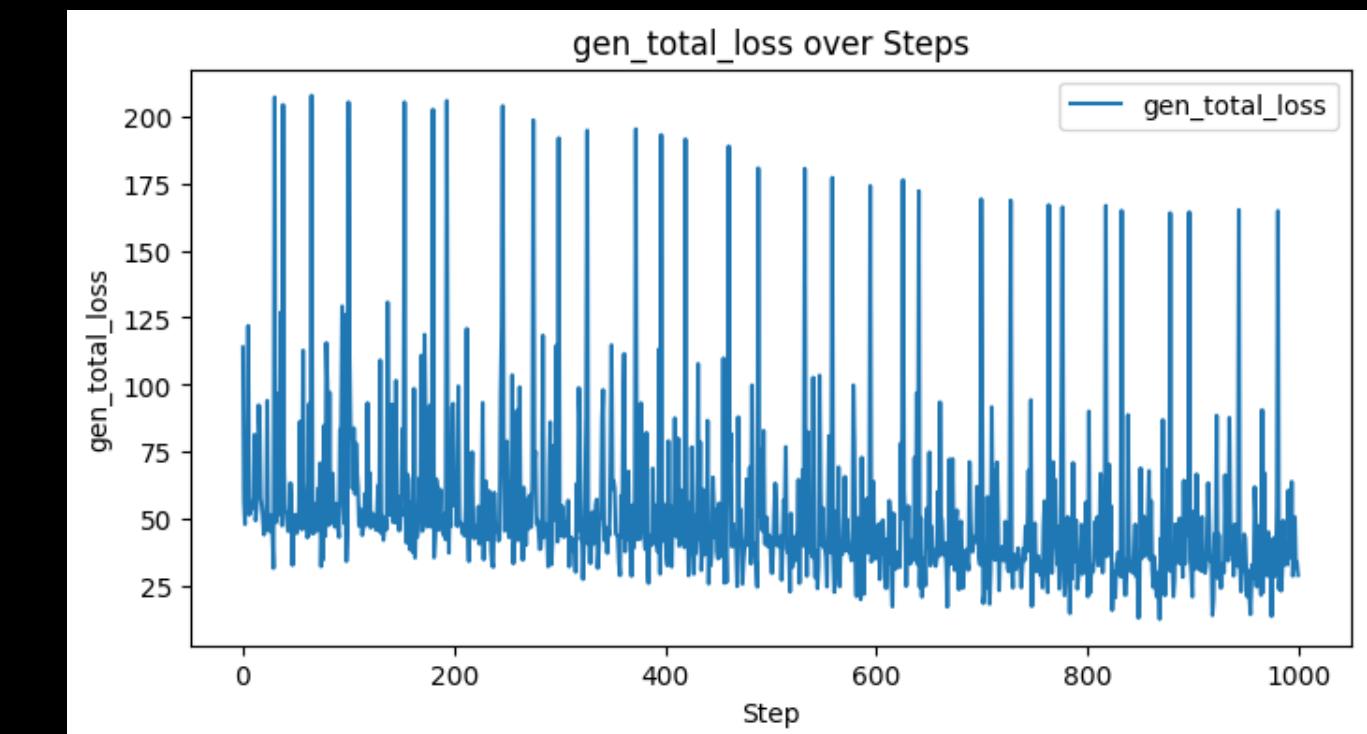
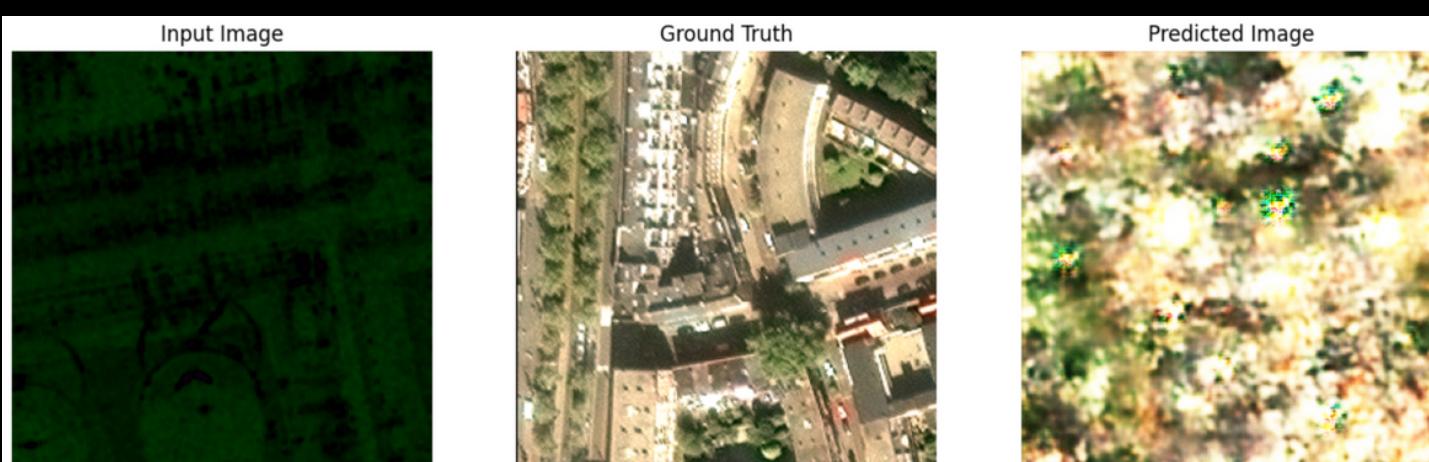
STEP 0



STEP 1000



STEP 2000



The Model performance deteriorates over time wrt visual feedback/observation (overfitting most likely). While the total loss oscillates heavily while trending downward. Input SAR image above is shown as a RGB image wrongly thus creating the green shade/tint.

Future Directions

- Better thresholding techniques to make the image sharper prior to feeding into the model might help the performance of the model.
- ViT (Visual Transformer Methods may be tried (they are the SoTA for now...).
- Adding more info about the SAR image such as the Map data of the captured area (from OpenStreetMap) might give more context and thus better results.