NiR Channel Generation from RGB Satellite Imagery

Made by Timur Nabiev
https://github.com/mantesssa/rgb_to_nir

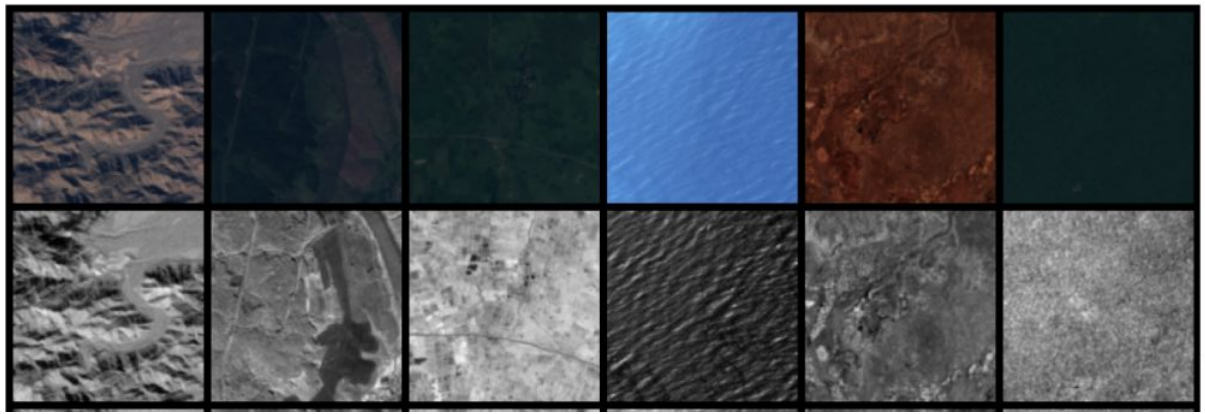
Task

This project focuses on the task of generating Near-Infrared (NIR) images based on corresponding visible spectrum (RGB) Satellite images.

For this task is used dataset https://www.kaggle.com/datasets/enddl22/deepnir-nir-rgb-sen12ms-dataset

Witch is subset of RGB and NIR pairs from Sentinel-2 satellite

SEN12MS dataset of satellite image patches: https://mediatum.ub.tum.de/1474000



RGG images

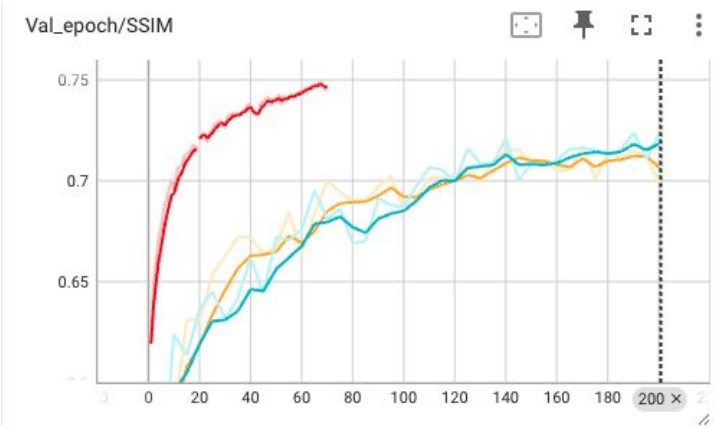
NIR images

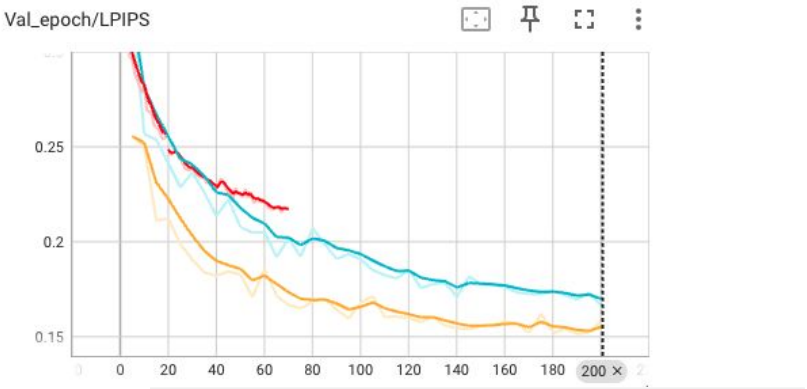
Architectures

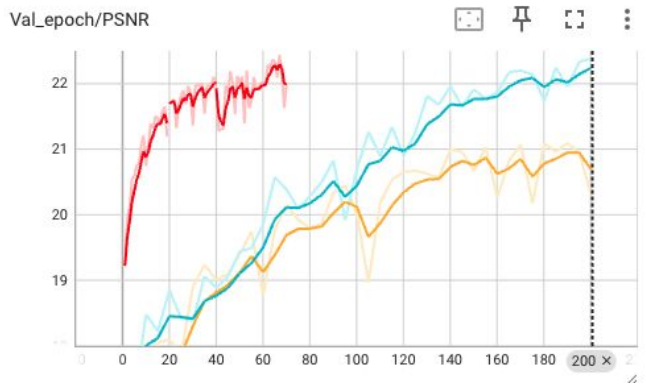
- Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) with U-Net architecture

- Flow Matching models used mainly the https://arxiv.org/pdf/2412.06264 repo
  - Continuous Flow Matching (CFM) with U-Net architecture
  - Discrete Flow Matching (DFM) with U-Net architecture

Training


Val_epoch/SSIM
0.75
0.7
0.65
0
20
40
60
80
100
120
140
160
180
200


Val_epoch/LPIPS
0.25
0.2
0.15
0
20
40
60
80
100
120
140
160
180
200

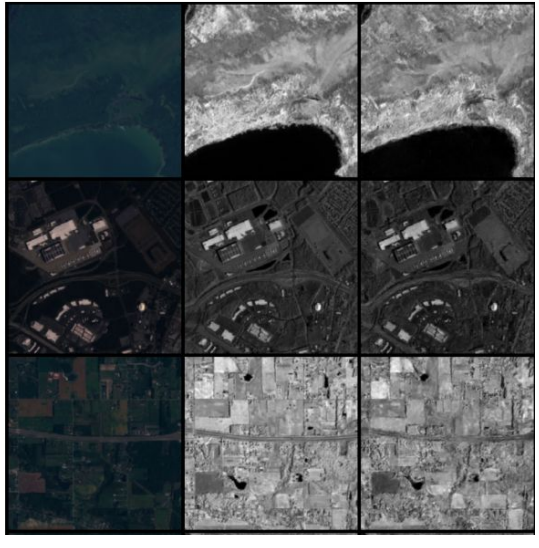

Val_epoch/PSNR
22
21
20
19
0
20
40
60
80
100
120
140
160
180
200

Results

	WGAN	DFM	CFM
SSIM (Higher is better, range [0,1])	0.7402	0.7002	0.7249
PSNR in dB (Higher is better)	22.1696	20.3391	22.4407
LPIPS (AlexNet) (Lower is better)	0.2263	0.1583	0.1659
SAM (Lower is better)	0.1611	0.1837	0.161
Correlation Coefficient (Higher is better, range [-1,1])	0.8385	0.7963	0.8432

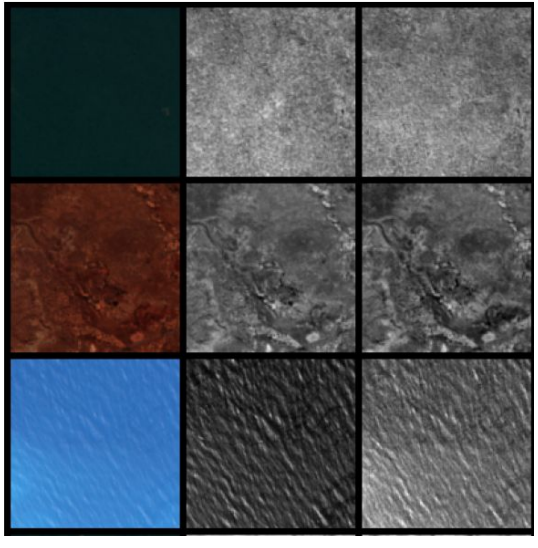
CFM

RGB GT NIR GEN NIR



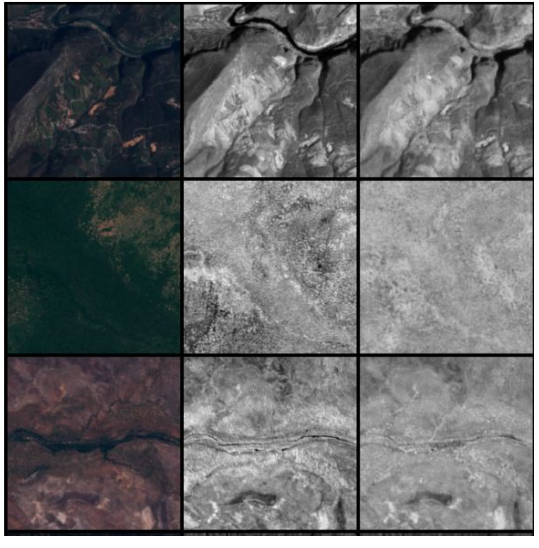
DFM

RGB GT NIR GEN NIR



WGAN

RGB GT NIR GEN NIR



Results

Three distinct deep learning models—WGAN, DFM, and CFM—were evaluated for the task of translating RGB images to the Near-Infrared (NIR) spectrum. All three models demonstrated commendable performance across a suite of standard image quality metrics. Specifically, the WGAN model achieved the highest score for the Structural Similarity Index Measure (SSIM = 0.7402), whereas the DFM model attained the best result for the Learned Perceptual Image Patch Similarity (LPIPS = 0.1583), which reflects perceptual similarity. The CFM model emerged as the most balanced, exhibiting superior performance in Peak Signal-to-Noise Ratio (PSNR = 22.4407 dB), Spectral Angle Mapper (SAM = 0.1610 rad), and Correlation Coefficient (0.8432), alongside competitive SSIM and LPIPS values. It is pertinent to note that, based on a subjective assessment of training speed, WGAN was observed to be the most rapid to train. Considering the high performance levels achieved, it is postulated that the models' potential has not yet been fully realized, and continued training, coupled with hyperparameter optimization, could yield further enhancements in the results.