ELE2006 Deep Learning for Image Analysis

Data augmentation for deforestation detection using Pix2Pix

Student: Amanda Lucas Pereira
Responsible Tutors:
Antonio Sombra (aluis.852@gmail.com)
Mabel Ortega (mabelortega.12@gmail.com)

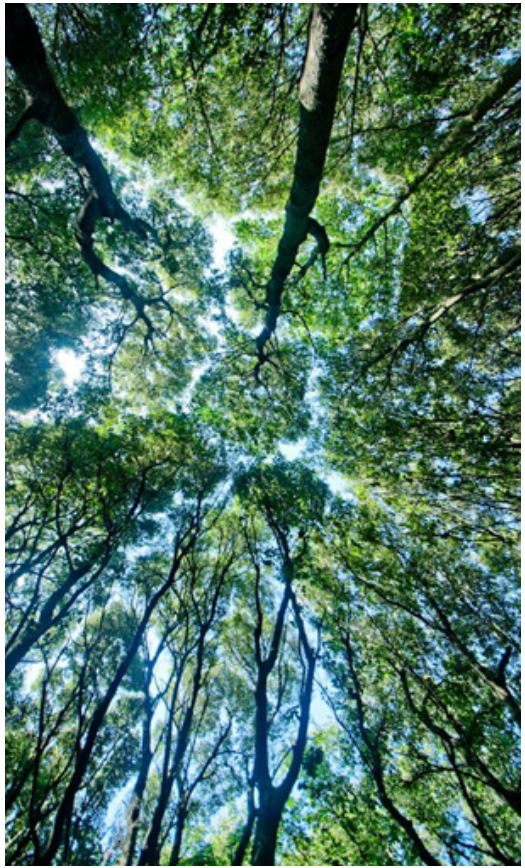
https://github.com/amandalucasp/pix2pix-deforestation (dev)

Context

Forests – Why are they important?

- Cover a **third of all land on Earth**.
- Forests comprise **80%** of all **terrestrial biodiversity**.
- Regulate **biogeochemical cycles** and **climate regulations.**
- Maintain the **ecological balance**.

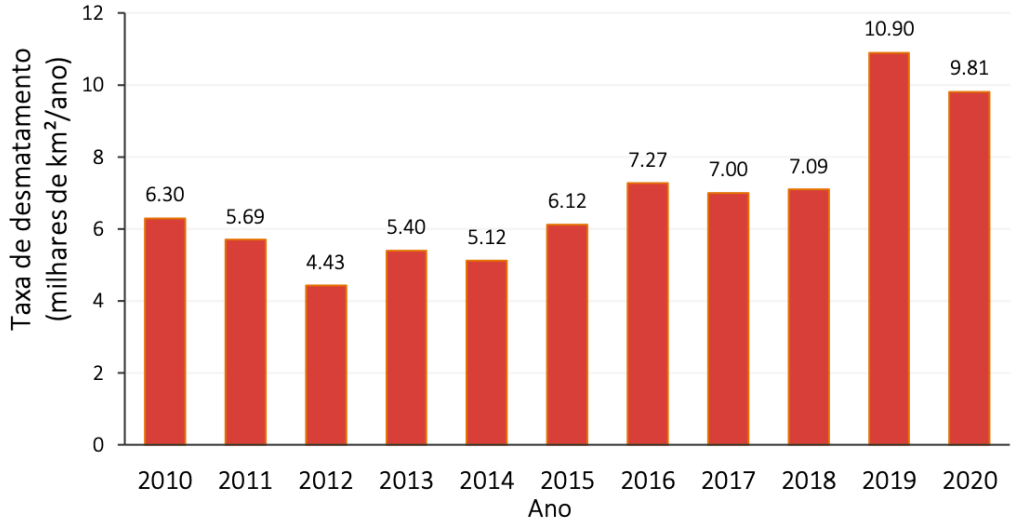
Brazil holds about **one-third** of the world's remaining **rainforests**



Pontifícia Universidade Católica do Rio de Janeiro

Context

Deforestation in the Amazon Biome



Source: Fatos da Amazônia 2021

Limitations

- Deep learning models require a lot of training samples
- Deforestation detection is an application that present unbalanced training data
- Labels are costly

Project Idea

This project aims to develop a method based on GANs for data augmentation in order to mitigate the problem of unbalanced data in deforestation detection and improve the training of the deep learning models used in this application.

Objectives

General objective:

- To train a pix2pix model to generate synthetic deforested samples to alleviate class imbalance in deforestation mapping applications.

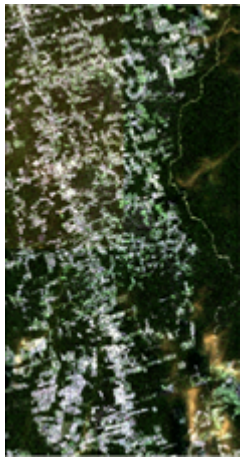
Specific objectives:

- Synthesize images with deforested spots at a later date from an image of an earlier date and a deforestation mask;
- Assess the performance improvement obtained with the augmented training set using synthesized images.

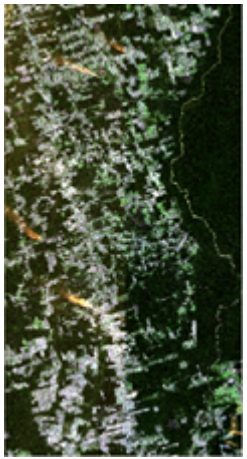
Dataset and Code

- Two HR optical images (9200 x 17720 x 13) from a region of the Amazon Biome.
- Reference (Deforestation 2018-2019)
  - 3 classes
    - 0 - No deforestation
    - 1 - New deforestation
    - 2 - Past deforestation

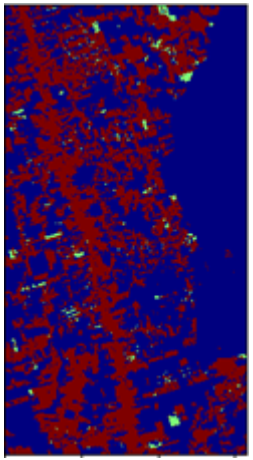
T1 (2018)



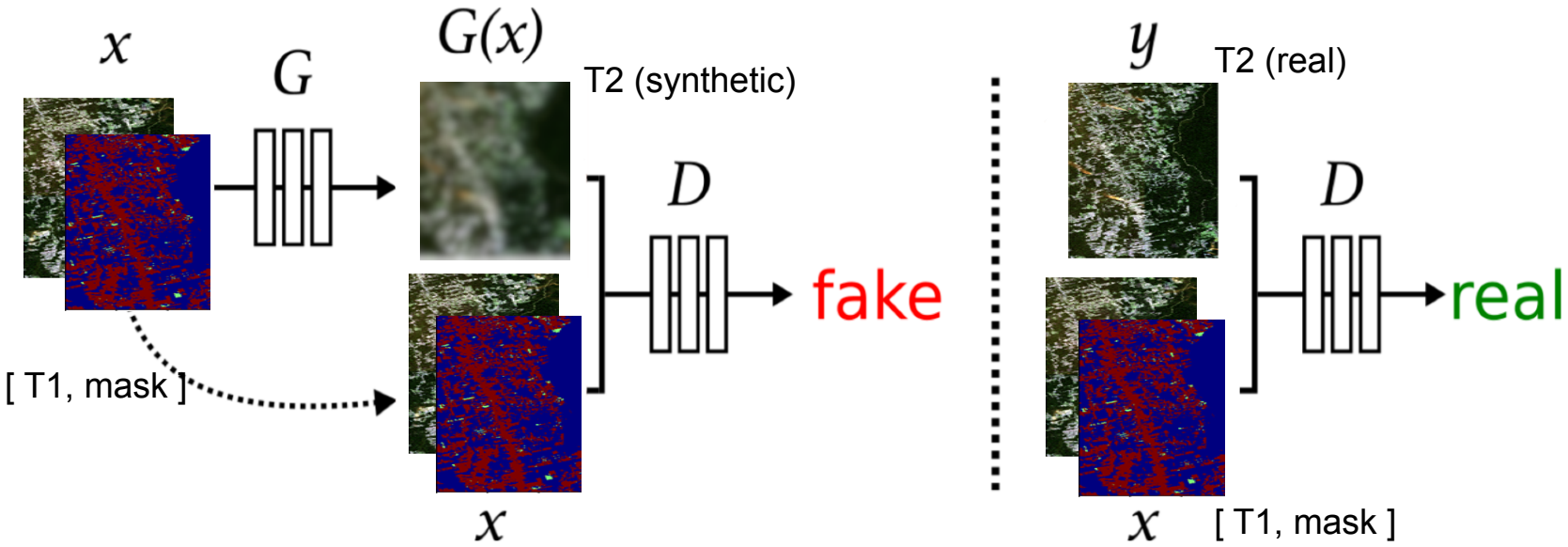
T2 (2019)



Reference (Deforestation 2018-2019)



Method: Training Pix2pix with T1, T2



Method: Generating new deforestation synthetic samples

[T1, new mask]

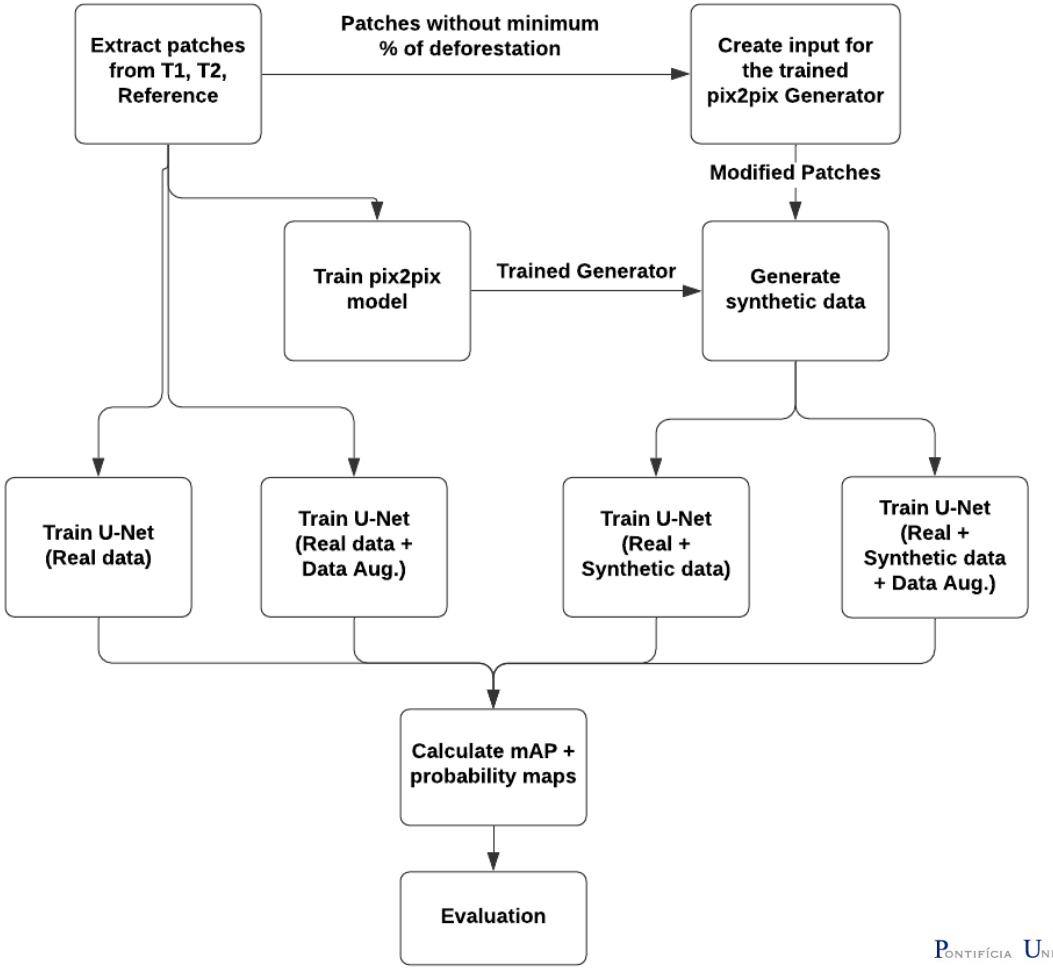
128x128x(C+1)

G

T2 (synthetic)

128x128xC

Pipeline


Extract patches from T1, T2, Reference
Patches without minimum % of deforestation
Create input for the trained pix2pix Generator
Modified Patches
Train pix2pix model
Trained Generator
Generate synthetic data
Train U-Net (Real data)
Train U-Net (Real data + Data Aug.)
Train U-Net (Real + Synthetic data)
Train U-Net (Real + Synthetic data + Data Aug.)
Calculate mAP + probability maps
Evaluation

Experimental Setup

Experimental Setup

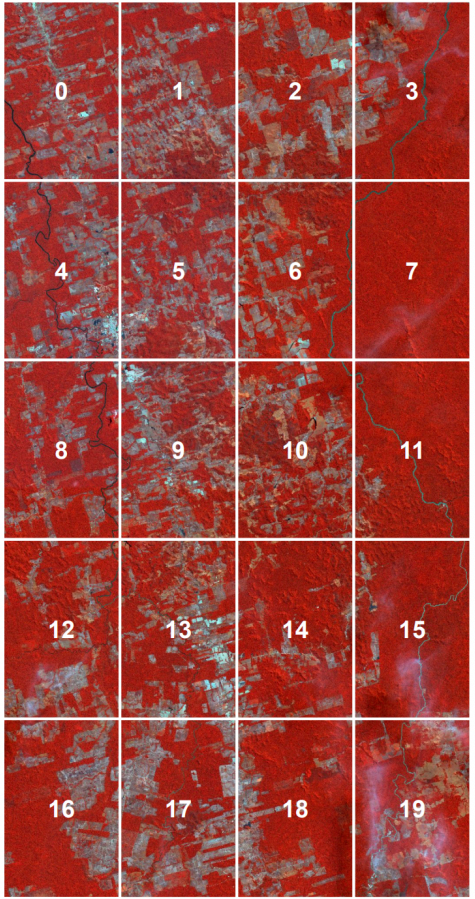
- Concatenate T1 and T2
- Bands choice: [Blue, Green, NIR]*
- Extract 128x128 patches with a minimum of 2% **New Deforestation**

Bands	μ	Resolution (m)
1 Coastal Aerosol	0.443	60
2 Blue	0.490	10
3 Green	0.560	10
4 Red	0.665	10
5 Vegetation Red Edge	0.705	20
6 Vegetation Red Edge	0.740	20
7 Vegetation Red Edge	0.783	20
8 Near infrared (NIR)	0.842	10
8A Vegetation Red Edge	0.865	20
9 Water Vapour	0.945	60
10 Shortwave IR/Cirrus	1.375	60
11 Shortwave IR 1	1.610	20
12 Shortwave IR 2	2.190	20

Table 1. Sentinel-2 Bands. The column μ refers to the central wavelength. Source: (Jovanovska Kaplan and Avdan, 2017)

Experimental Setup

- Split the image in 20 tiles
  - 8 tiles for training, 2 for validation and 10 for test
  - Same split used in (M. X. Ortega, R. Q. Feitosa, J. D. Bermudez, P. N. Happ and C. A. De Almeida, 2021)

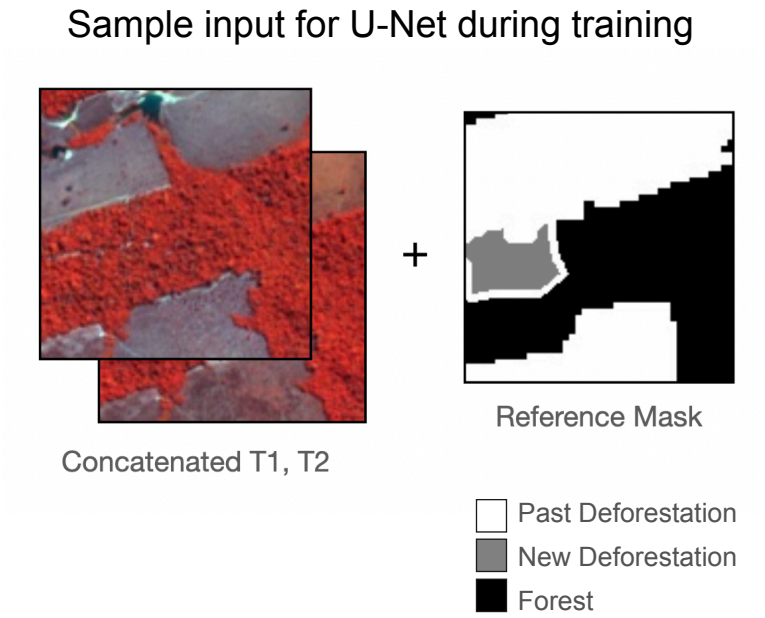


Experimental Setup

- **Patches for U-Net** (stride = 0.7):
  - Training patches: 2800
  - Validation patches: 806
  - Testing patches: 3143
- **Patches for pix2pix** (stride = 0.8):
  - Training patches: 6432
  - Testing patches: 7175

Experimental Setup

- U-Net: each input pair is composed by an image of T1+T2, and its corresponding reference mask.

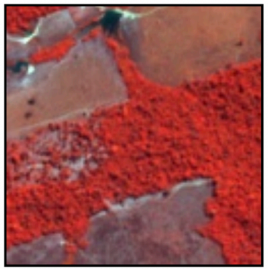


Experimental Setup

- Pix2pix: each input pair is composed by T1+mask, and the reference is the image from T2.

Sample input for the pix2pix model during training

Concatenated T1, Mask is fed to the Generator



T2 is fed to the Discriminator and compared to the Generator output

Past Deforestation
New Deforestation
Forest

Pontifícia Universidade Católica do Rio de Janeiro

Experimental Setup

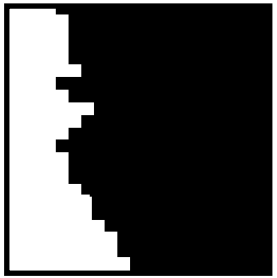
- During this stage, the patches that are rejected due to not meeting the minimum % of New Deforestation criteria are later **modified/combined** in order to provide input for the trained pix2pix
  - A few alternatives were tested, such as
    - **Combining patches without deforestation with masks that contain new deforestation**
    - Modifying reference masks to add more deforestation (through dilation)

Creating input for the trained pix2pix Generator

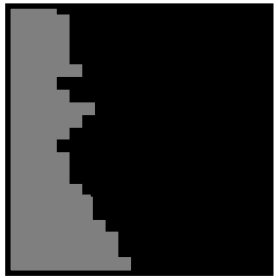
- For this report, the process of creating input for the **trained pix2pix Generator** was done as it follows:
  - Get **images with No Deforestation in T1**
  - Get **masks with *only Past Deforestation*** and change the pixel value from the old deforestation pixels to the value corresponding to ***New Deforestation***
  - **Randomly combine T1 + new mask to form the input pairs**
  - Feed the **T1 + new mask** to the trained Generator
- Using this method of combining no deforestation image patches with modified deforested reference patches, a total of 10000 patches were generated.

Image from T1 with only “No Deforestation” on its mask

Change pixel value from “Past Deforestation” (White) to “New Deforestation” (Grey)



Mask with only “Past Deforestation” (White) and Forest (Black)



New Mask

Concatenated T1 and New Mask

Past Deforestation
New Deforestation
Forest

Pontifícia Universidade Católica do Rio de Janeiro

Creating input for the trained pix2pix Generator

- With the pairs composed by these rejected patches, input is provided to the trained pix2pix Generator

Experiments with pix2pix

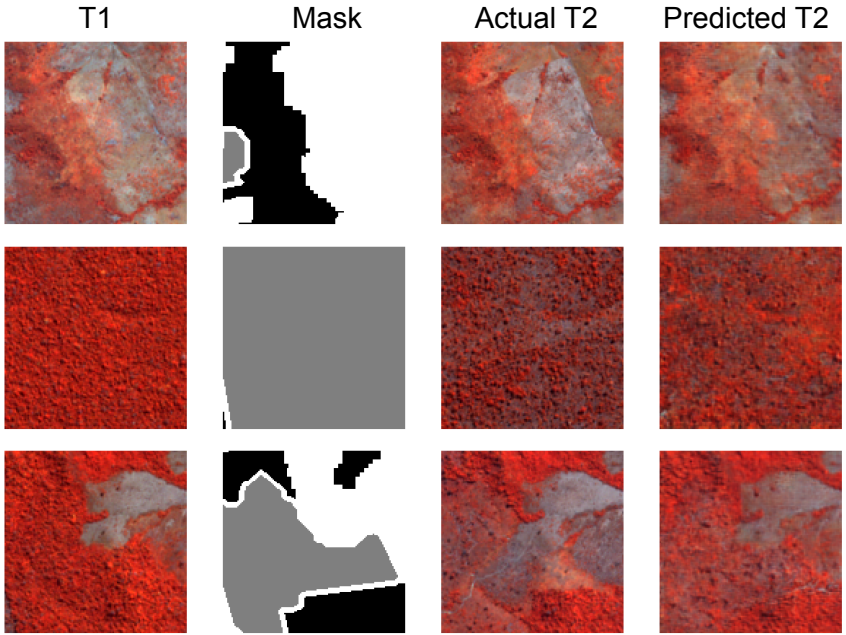
Experiments with pix2pix

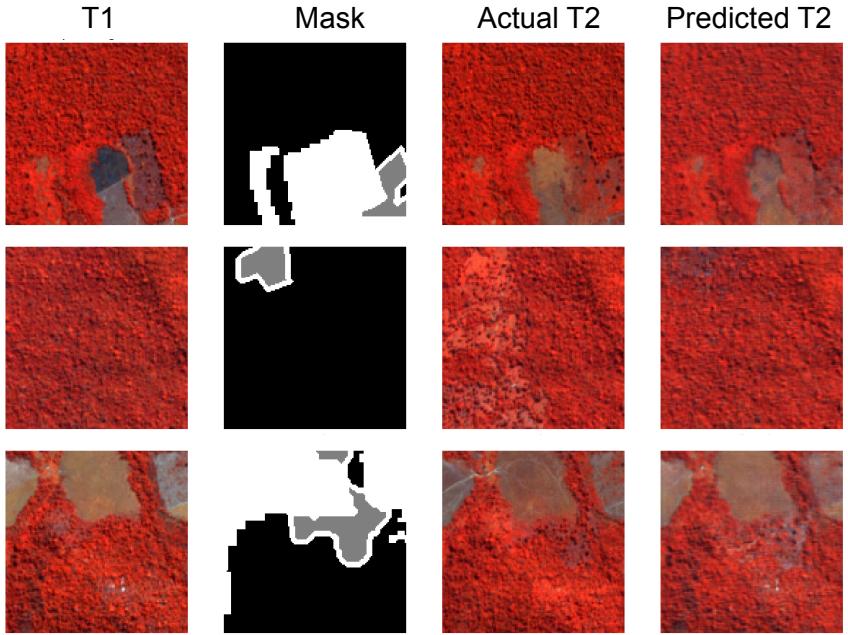
- In the experiments of synthetic images using pix2pix it will be presented only a qualitative analysis due to the lack of metrics to validate its performance
  - Testing patches used for visual evaluation
  - Quantitative results are obtained later using the generated data on U-Net
- For training the pix2pix models, patches obtained using 0.8 of overlap:
  - Training patches: 6432 (using 8 training tiles)
  - Testing patches: 7175 (using 10 testing tiles)

Experiments with pix2pix

- **Experiment 1A: trying different number of filters**
  - Original pix2pix architecture uses patches 256x256. Since we're using 128x128xc we tried using different number of filters in the input layer and also removing layers in the bottleneck of the Generator
  - Satisfactory results were obtained using
    - 64 filters in the input layer
    - Removing encoder's final layer
    - Removing decoder's first layer

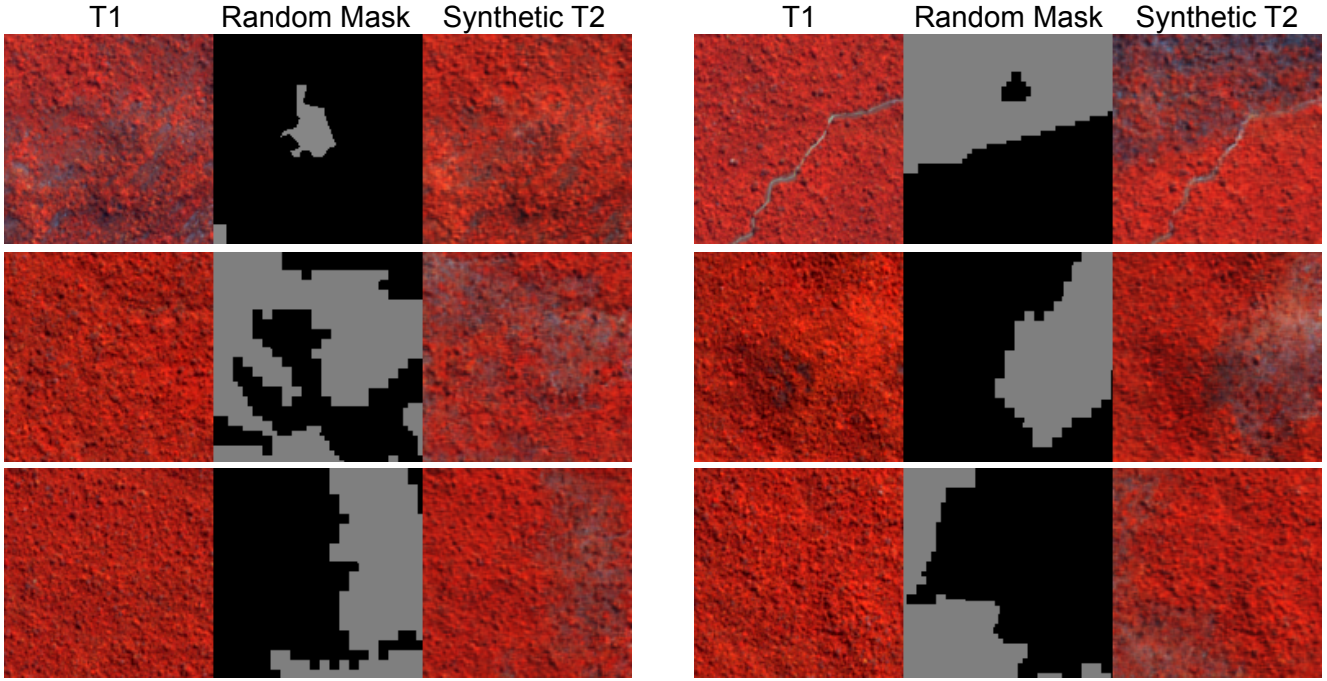
Experiment 1A: Visual Inspection on Test Data





Past Deforestation
New Deforestation
Forest

Experiment 1A: Synthesis of new T2 images using random modified masks

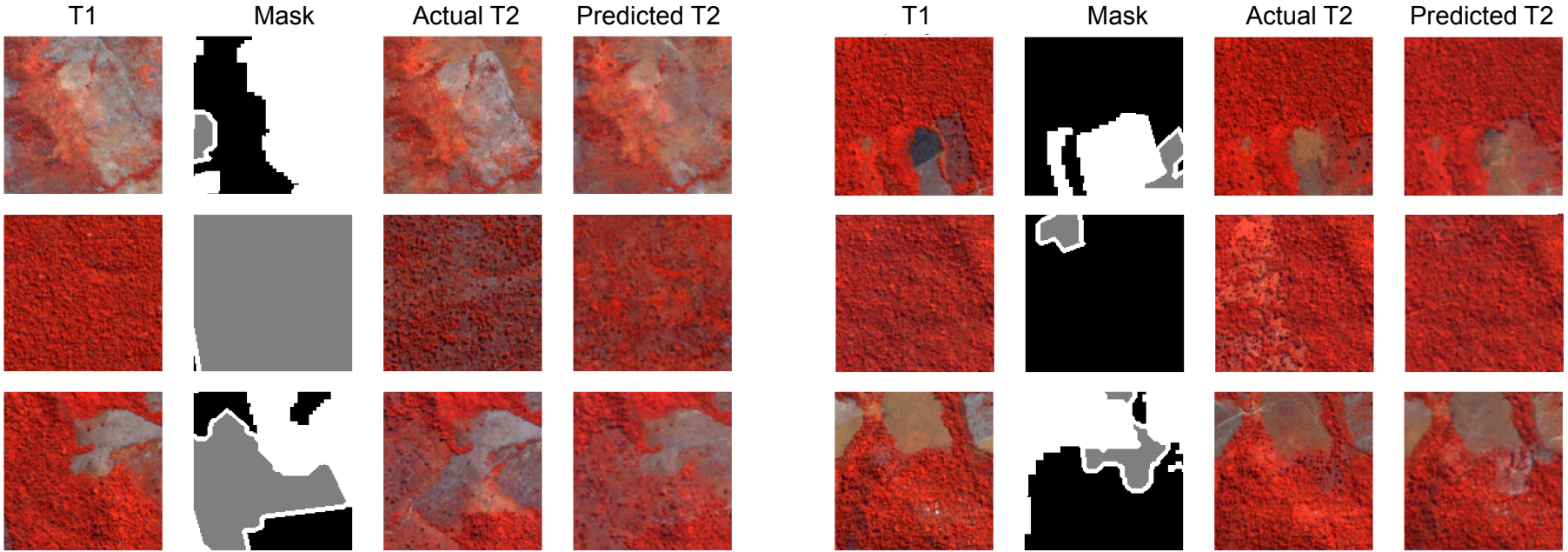


New Deforestation
Forest

Experiments with pix2pix

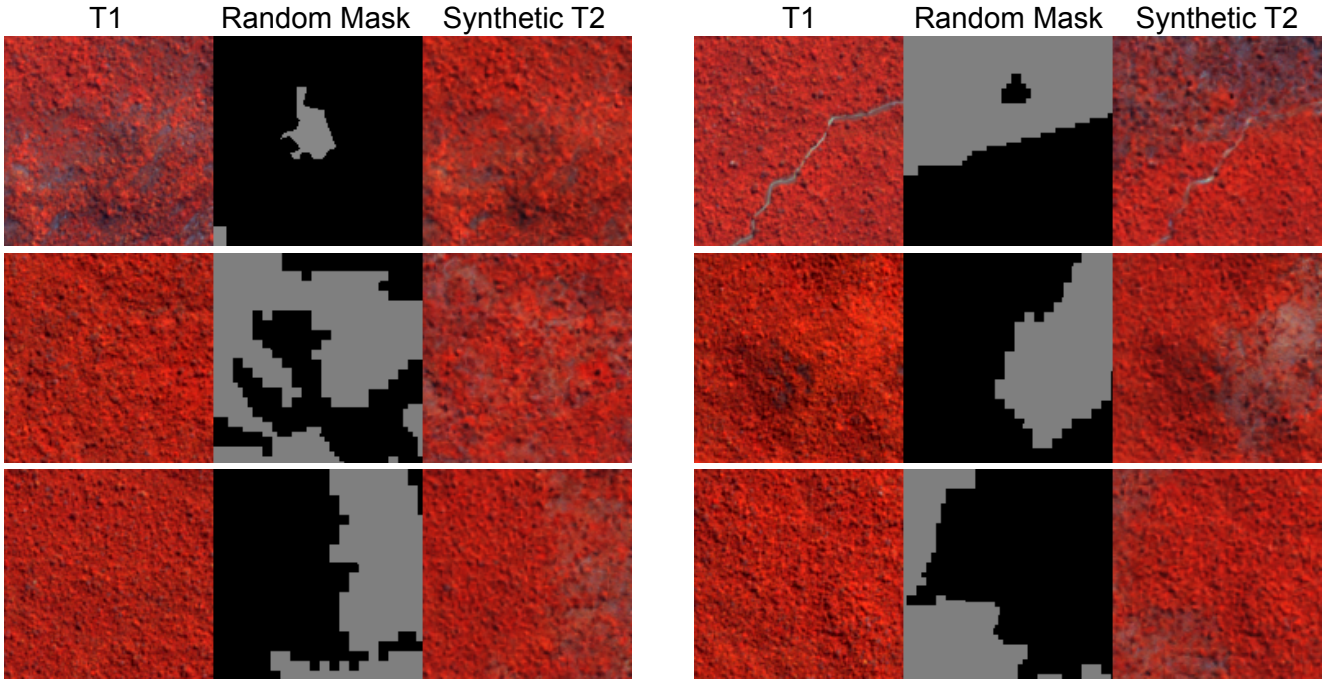
- **Experiment 1B: Modifying parameters from 1A**
  - The initial pix2pix code used for this project had some differences in implementation when comparing to the tensorflow version by affine-layer
  - In this experiment, we modified the way loss is calculated to match the author's:
    - adding an "Epsilon" parameter that might help convergence
    - using a Sigmoid in the output of the Discriminator.
    - coding changes (as in the specific functions used)

Experiment 1B: Visual Inspection on Test Data



Past Deforestation
New Deforestation
Forest

Experiment 1B: Synthesis of new T2 images using random modified masks



New Deforestation
Forest

Experiments with pix2pix

- **Experiment 2: feeding a binary mask to the pix2pix model**
  - The intuition behind this experiment was to try to help the model learn to only modify pixels which corresponded to the “new deforestation” class.
  - Synthetic images obtained for this approach seemed “blurry”
- **Experiment 3: using residual blocks on the bottleneck of Generator**
  - We tested keeping and dropping layers from the Encoder/Decoder for the Generator.
  - When keeping, the training collapsed; when dropping, the images seemed a bit “blurry”.

Experiments with U-Net

Experiments with U-Net

- Experimentation is done by training the U-Net model from scratch in 5 runs, with a patience of 10 and using a Weighted Categorial Cross Entropy, with the weight assigned to the *old* deforestation class set to zero (As in M. X. Ortega, R. Q. Feitosa, J. D. Bermudez, P. N. Happ and C. A. De Almeida, 2021).
- Visualization maps are generated using the mean output from the 5 experiments.

U-Net: Baseline Experiment

- **Only real patches were used. Baseline Dataset:**
  - Training patches: 2800
  - Validation patches: 806
  - Testing patches: 3143

U-Net + Classic Data Augmentation

- **Only real patches were used + Classic Data Augmentation methods**
- Methods are the same as in (M. X. Ortega, R. Q. Feitosa, J. D. Bermudez, P. N. Happ and C. A. De Almeida, 2021)
  - Random rotation of 90 degrees
  - Horizontal and vertical flips

U-Net + Synthetic Data

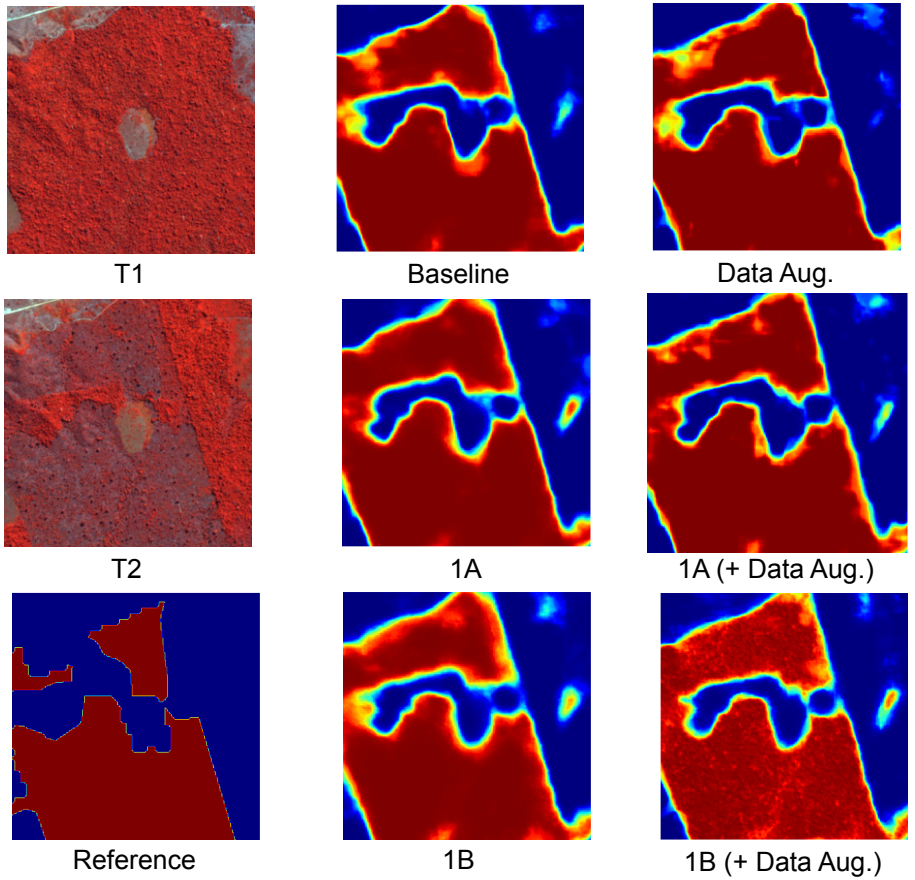
- **For this experiment, real and synthetic patches were used.**
- **Augmented Dataset:**
  - Training patches: 2800 **+ 1000 synthetic samples**
  - Validation patches: 806
  - Testing patches: 3143

U-Net: Results

Exp	Dataset	Classic DA	mAP
1	Real Data	No	91.91
2	Real Data	Yes	92.39
3	Real+Synt (1A)	No	92.38
4	Real+Synt (1A)	Yes	91.26
5	Real+Synt (1B)	No	91.77
6	Real+Synt (1B)	Yes	92.79

Table 1: Results obtained using 10 different thresholds for the calculation of mAP. In the table, (1A) and (1B) refer to pix2pix models trained differently.

Sample results using a testing patch



Conclusion

- Considering the provided dataset, pix2pix seems to able to mimic the most usual kinds of deforestation presented in the dataset, producing “grey-ish” pixels when generating synthetic output
- It might improve the results obtained in pix2pix if the provided Deforestation mask presented a granularity coherent to the number of apparent sub-classes
    - Eg. some areas are deforested in T1 but not in T2, and labeled as “Past deforestation”
    - Maybe this problems will vanish when using more bands

Ideas in Progress

- Experiments using all/different bands
- Changes in the architecture of the pix2pix itself
- Changes in the residual blocks' architecture
- Modifying the normalization applied in the preprocessing stage

Thank you