"Mining The Mines
Using Data Mining"

Bilawal Singh Deu
Dhirain Vij
Sakarth Singh Brar

Problem Statement

Context

Mineral resources are used in various fields, from daily necessities to cutting-edge technology, and have become an indispensable element for the development of modern society.
In Africa and Southeast Asia, these mineral resources are abundant and contribute to economic growth and trade expansion as major exports in **many African and Southeast Asian countries.**

On the other hand, these countries **face issues, such as the inability to properly monitor mine development, often resulting from a lack of their resources.**

Problem Statement

Our Goal

We aim to develop a technology for **detecting mining sites using images from the optical satellite Sentinel-2**. Specifically, it involves classifying images that contain mining sites and those that do not.

Potential Impact

- **Environmental Conservation** by preventing unauthorized mining
- **Economic Benefits** to help in sustaining the local economies that depend on mining, such as in Senegal
- Alleviate **social issues** like conflicts over land use, labor exploitation.
- **Mitigate health risks** associated with poor mining practices, such as exposure to mercury and increased substance abuse.

Literature Review

Mine void identification using Object-based Image Analysis (OBIA) of satellite imagery Sentinel 2 data

This is a comprehensive study on mine void identification using Object-based Image Analysis (OBIA) of satellite imagery, specifically Sentinel-2 data.

Main drawback of this method is the inability to identify voids with an area <1 ha.

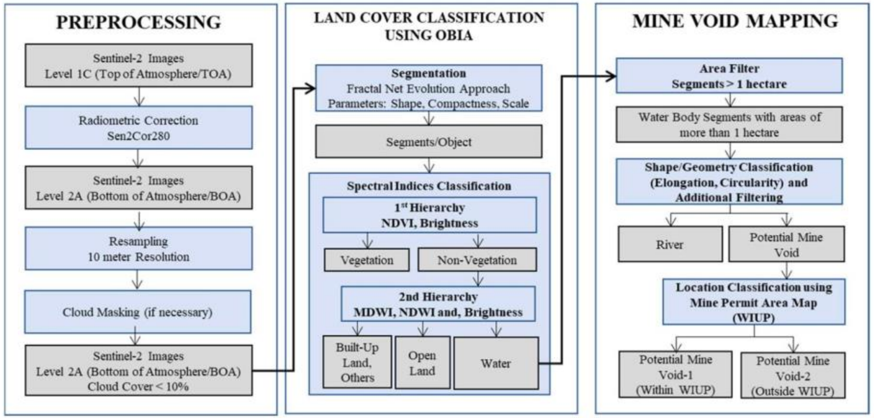
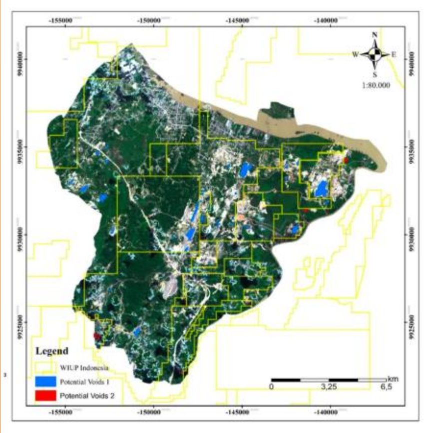
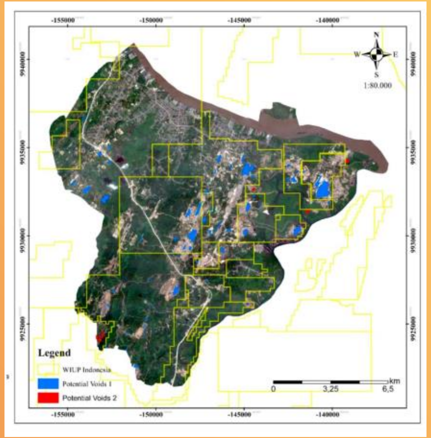


Figure 2. The workflow of methods used in this study.

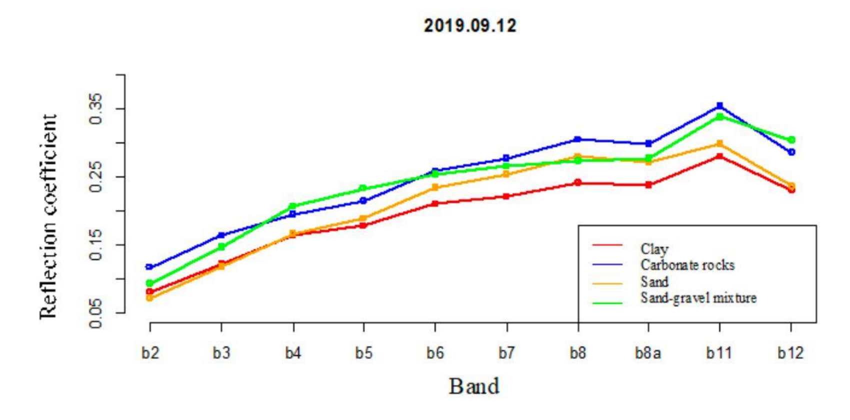




Lestari, L., Kusuma, G. J., Badhurahman, A., Dwiki, S., & Gautama, R. S. (2023). Mine void identification using Object-based Image Analysis (OBIA) of satellite imagery Sentinel 2 data. *Journal of Degraded and Mining Lands Management, 10(2), 4129-4142. https://doi.org/10.15243/jdmlm.2023.102.4129

Literature Review

Automated detection of illegal nonmetallic minerals mining places according to Sentinel-2 data



Reflectance values in 10 spectral channels of the Sentinel 2.

1. Spatial Resolution Adjustment to the best resolution of 10 meters using GDAL library utilities, resulting in a single GeoTiff file per image.
2. 12 Spectral Indices Calculation for each image, leading to the creation of two composites consisting of 22 layers each (10 spectral channels + 12 spectral indices).
3. Utilized 2019 field observation data to create vector layers marking reference quarry locations.
4. Training samples for four types of materials (sand, sand-gravel mix, clays, carbonate rocks) were formed by collecting pixel values corresponding to these locations across 22 channels from two Sentinel-2 composites.
5. Mahalanobis Distance Classification to get the probability of each pixel belonging to a particular quarry type.
6. Maximum Entropy Method to produce probability distribution models for each quarry type. MaxEnt is used for its ability to generate models and predictions based on presence-only data.

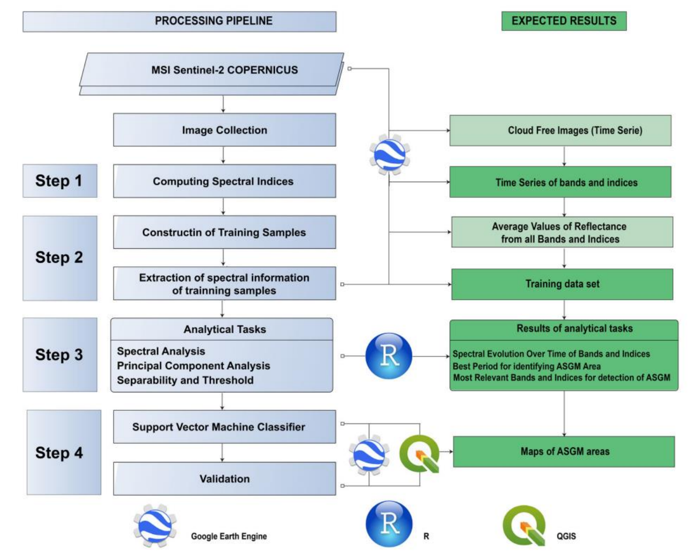
Composites of multiple indexes must be used.

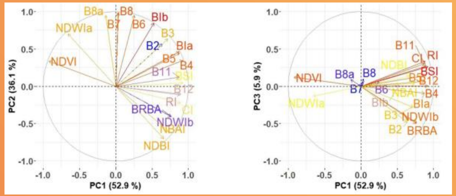
Usmanov, B. M., Isakova, L. S., Mukharamova, S. S., Akhmetzyanova, L. G., & Kuritsin, I. N. (2021). Automated detection of illegal nonmetallic minerals mining places according to Sentinel-2 data. In Proceedings of SPIE 11863, Earth Resources and Environmental Remote Sensing/GIS Applications XII, 118631C. SPIE. https://doi.org/10.1117/12.2600315

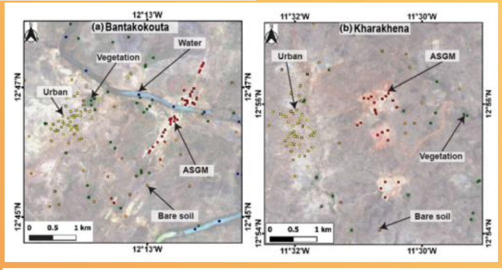
Literature Review

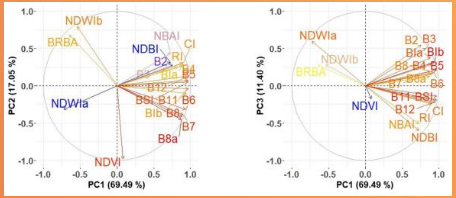
Mapping Artisanal and Small-Scale Gold Mining in Senegal Using Sentinel 2 Data

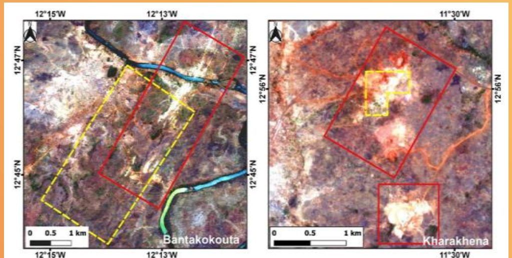
- The objective of this study was to develop a method for detecting and mapping artisanal and small-scale gold mining (ASGM) sites in Senegal using Sentinel 2 data.
- The final result of this study is therefore a map of ASGM sites. By cross validating with the field reality, an overall accuracy of 0.89 is found for the Bantakokouta site (Kappa = 0.85) and 0.80 for Kharakhena (kappa = 0.73).











Lestari, L., Kusuma, G. J., Badhurahman, A., Dwiki, S., & Gautama, R. S. (2023). Mine void identification using Object-based Image Analysis (OBIA) of satellite imagery Sentinel 2 data. *Journal of Degraded and Mining Lands Management, 10(2), 4129-4142. https://doi.org/10.15243/jdmlm.2023.102.4129

Dataset

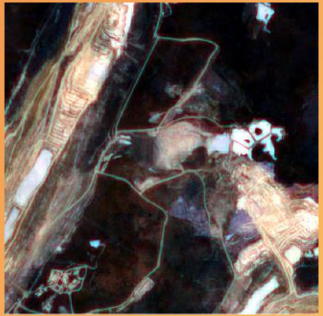
The Sentinel 2 multispectral data, publicly released by the European Space Agency (ESA), are used for this study.
They are acquired through two twin satellites (Sentinel 2A and Sentinel 2B), launched separately in synchronous polar orbit at an altitude of 786 km and evolving 180° apart. Each satellite is equipped with a Multi-Spectral Imager (MSI) sensor including 13 spectral bands (from 443 to 2,190 nm) with a field of view of 290 km and a spatial resolutions of 10 m (four bands in the visible and near infrared domains), 20 m (six bands in the near and short wavelength infrared domain, NIR and SWIR), and 60 m (three atmospheric correction bands)

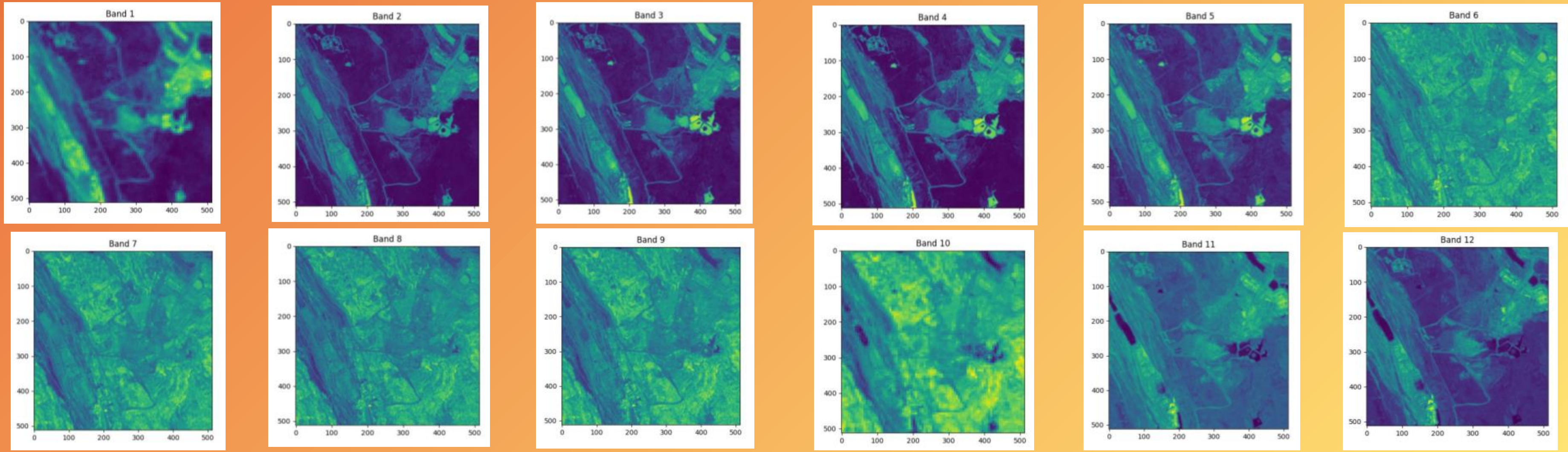
Sentinel-2 bands	Central wavelength (μm)	Resolution (m)
Band 1 – Coastal aerosol	0.443	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation red edge	0.705	20
Band 6 – Vegetation red edge	0.740	20
Band 7 – Vegetation red edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Vegetation red edge	0.865	20
Band 9 – Water vapour	0.945	60
Band 10 – SWIR – Cirrius	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

1. Classification model

Our Data

- The Sentinel-2 images contain information on 12 bands:
  'B1', 'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B8A', 'B9', 'B11', 'B12'
- Each band is of size 512x512 pixels.
- The images have been processed to mask clouds from 2022/1/1 to 2023/12/31, and the median of all images is used as the image for that location.





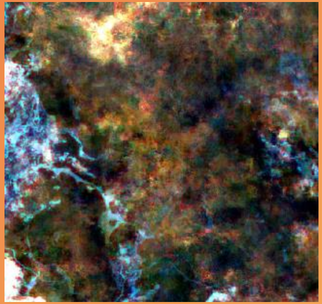
Our Data

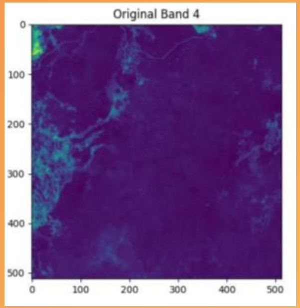
- We have 1242 images in our train dataset.
- Corresponding to each image, we have either a label '1' (Contains Mining Site) or '0' (Doesn't contain Mining Site) in a csv file.
- The competition dataset is not balanced as there are only 256 positive samples out of 1242 which is only around 20%.
- We have 1243 images in our evaluation dataset.

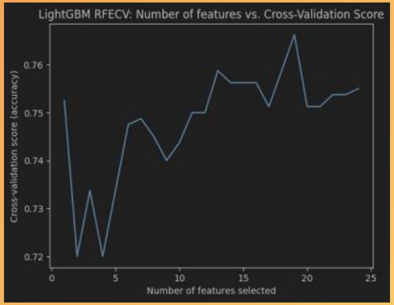
	A	B	C
1	train_0.tif	0	
2	train_1.tif	0	
3	train_2.tif	0	
4	train_3.tif	1	
5	train_4.tif	0	
6	train_5.tif	0	
7	train_6.tif	0	
8	train_7.tif	0	
9	train_8.tif	0	
10	train_9.tif	0	
11	train_10.tif	0	
12	train_11.tif	0	
13	train_12.tif	1	
14	train_13.tif	0	
15	train_14.tif	1	

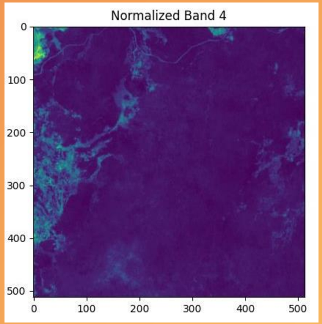
Data Pre-processing

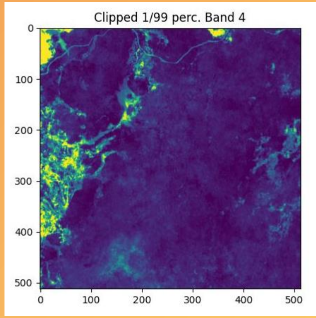
- Normalization/Clipping
- Adding spectral Indices
- Adding spectral Indices
- Feature extraction
- Data Augmentation

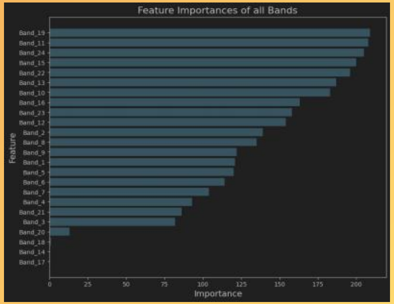













```
Optimal number of features : 19
Model Accuracy with Selected Features: 0.75
['Band_1', 'Band_2', 'Band_3', 'Band_4', 'Band_5', 'Band_6', 'Band_7', 'Band_8', 'Band_9', 'Band_10', 'Band_11', 'Band_12',
'Band_13', 'Band_15', 'Band_16', 'Band_19', 'Band_22', 'Band_23', 'Band_24']
```

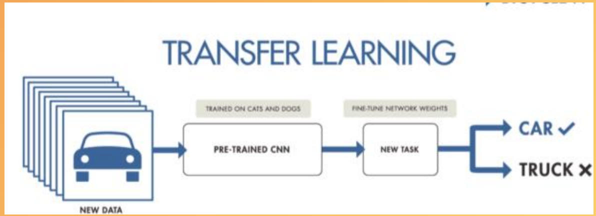
Usmanov, B. M., Isakova, L. S., Mukharamova, S. S., Akhmetzyanova, L. G., & Kuritsin, I. N. (2021). Automated detection of illegal nonmetallic minerals mining places according to Sentinel-2 data. In Proceedings of SPIE 11863, Earth Resources and Environmental Remote Sensing/GIS Applications XII, 118631C. SPIE. https://doi.org/10.1117/12.2600315

ML-Methodology

Convolutional Neural Network (CNN)

With Transfer Learning

- The MaxViT model, specifically the "maxvit_tiny_tf_512" configuration, is being employed that is relevant to classification tasks.
- MaxViT is a type of Vision Transformer that integrates both convolutional and transformer architectures.
- Maxvit_tiny_tf_512 model is initialized with 14 input channels and is set to predict two classes.

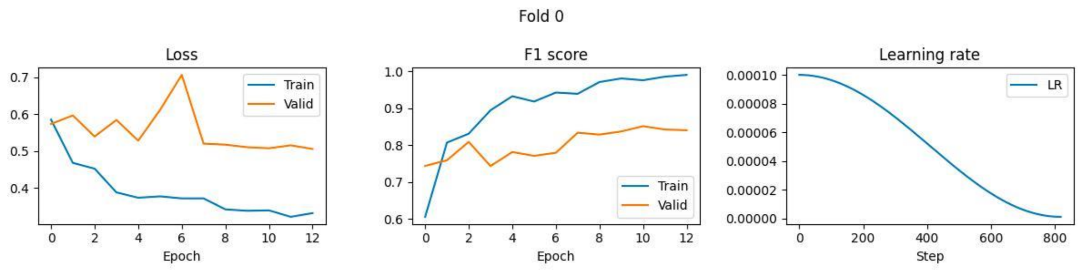


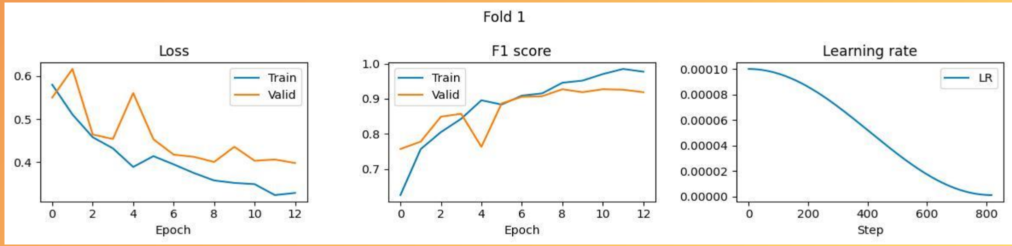
Why?

- MaxViT can benefit from transfer learning, where a model pretrained on a large dataset (such as ImageNet) is fine-tuned on a smaller, domain-specific dataset.
- MaxViT combines the strengths of both convolutional neural networks (CNNs) and transformers.

Gallwey, J., Robiati, C., Coggan, J., Vogt, D., & Eyre, M. (2020). A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining. *Remote Sensing of Environment, 248, 111970. https://doi.org/10.1016/j.rse.2020.111970

Performance Metrics



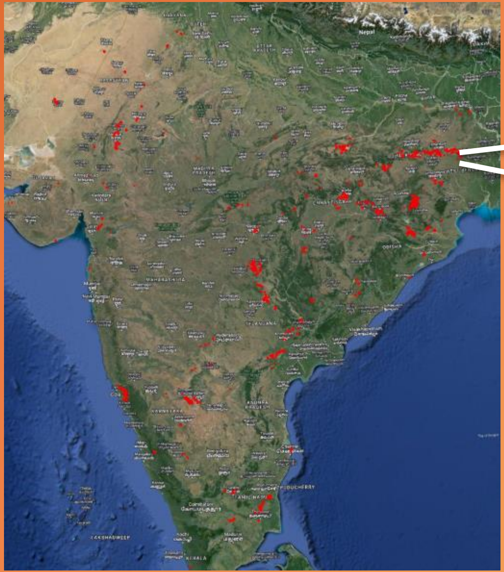


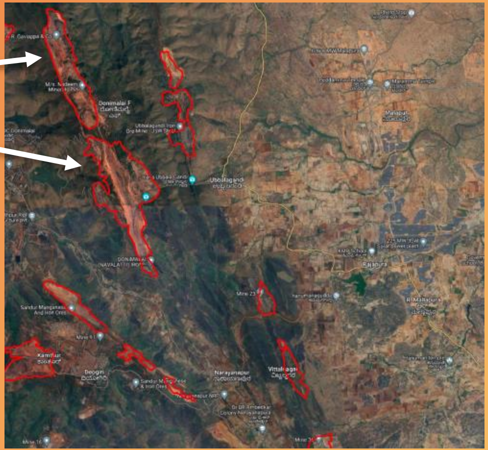

```
Epoch 12 training 63/63 [LR 0.000001] - loss: 0.3319: 100%|██████████| 63/63 [00:25<00:00,  2.51it/s]

tp: 203, tn: 786, fp: 2, fn: 2
Epoch 12 train loss = 0.3319, base f1 score (0.5 threshold) = 0.9828 (best threhold: 0.39 -> f1 0.9902)


100%|██████████| 42/42 [00:04<00:00,  9.22it/s]

tp: 42, tn: 191, fp: 7, fn: 9
Epoch 12 validation loss = 0.5055, base f1 score (0.5 threshold) = 0.8400 (best threhold: 0.49 -> f1 0.8400)
Accuracy: 0.9357
Confusion Matrix:
tensor([[191,   9],
        [  7,  42]], dtype=torch.int32)
```

```
Epoch 12 training 63/63 [LR 0.000001] - loss: 0.3292: 100%|██████████| 63/63 [00:26<00:00,  2.38it/s]

tp: 195, tn: 789, fp: 0, fn: 9
Epoch 12 train loss = 0.3292, base f1 score (0.5 threshold) = 0.9754 (best threhold: 0.66 -> f1 0.9774)


100%|██████████| 42/42 [00:04<00:00,  9.46it/s]

tp: 51, tn: 189, fp: 8, fn: 1
Epoch 12 validation loss = 0.3984, base f1 score (0.5 threshold) = 0.9174 (best threhold: 0.43 -> f1 0.9189)
Accuracy: 0.9639
Confusion Matrix:
tensor([[189,   1],
        [  8,  51]], dtype=torch.int32)
```

2. Binary Segmentation model

Our Data

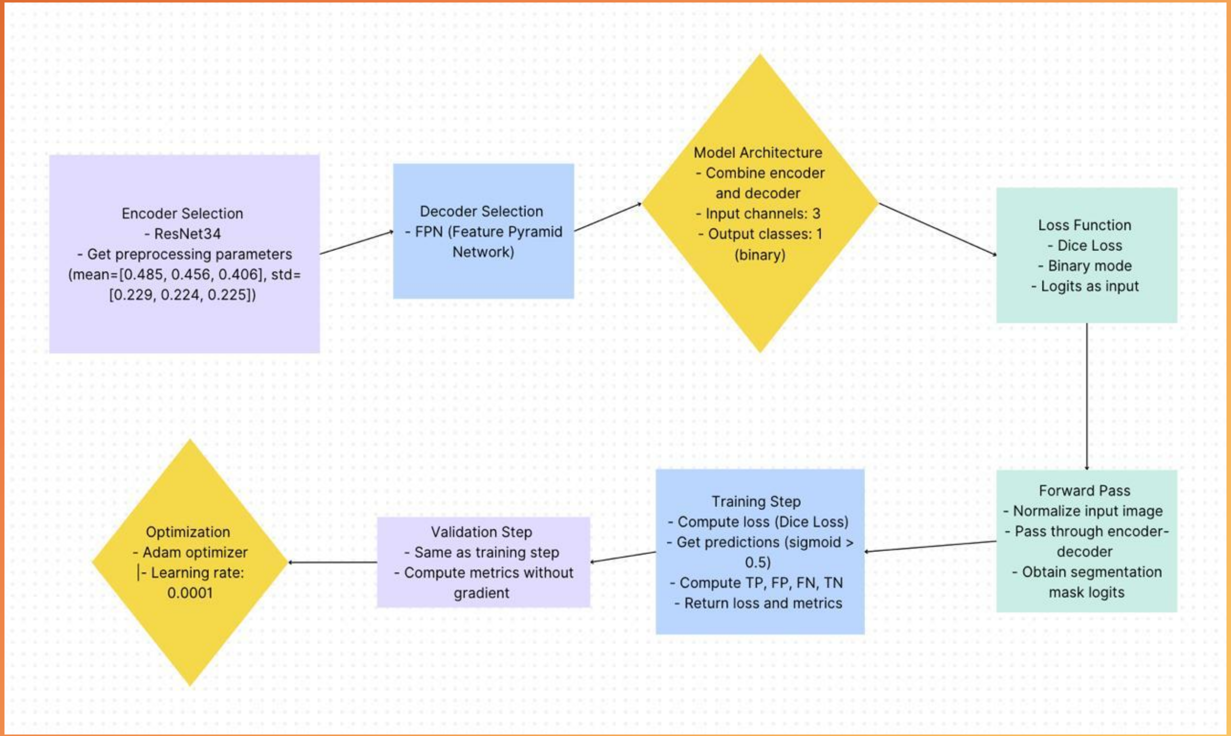




QGIS 3.6 Noosa

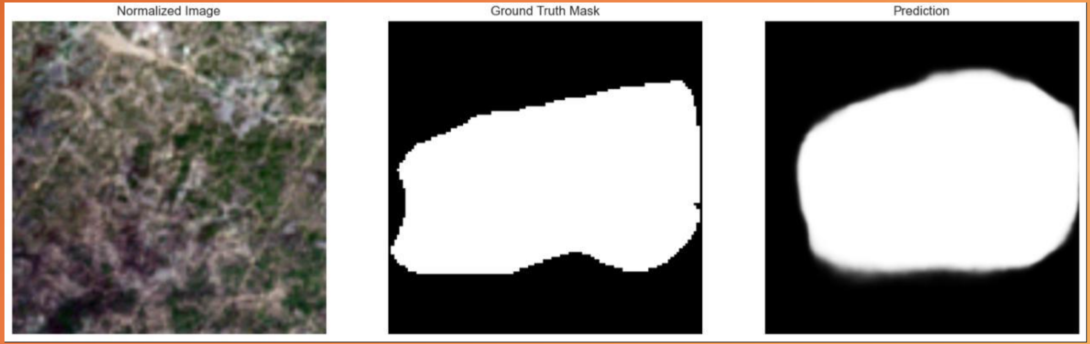
Google Earth Engine

ML-Methodology



Madhuanand, L., Sadavarte, P., Visschedijk, A.J.H., Denier Van Der Gon, H.A.C., Aben, I., & Osei, F.B. (2021). Deep convolutional neural networks for surface coal mines determination from sentinel-2 images. *European Journal of Remote Sensing, 54(1), 296-309. https://doi.org/10.1080/22797254.2021.1920341

Performance Metrics




```
[{'valid_dataset_iou': 0.5550410747528076,
  'valid_per_image_iou': 0.5450248718261719}]
```


```
[{'test_dataset_iou': 0.5458776950836182,
  'test_per_image_iou': 0.5388392210006714}]
```

Thanks!