High Resolution Earth Geo-Observatory System for Sentinel-2 Imagery Using Deep Learning (HR-MARS)

Muhammad Roshan Mughees

Abdul Razzaque Jawad

Awais Latif Khatti

Advisor: Dr. Muhammad Shahzad

Co-Advisor: Dr. Faisal Shafait

Submissions Confirmation

- FYP Report reviewed by the Advisor - **YES**
- FYP Report uploaded on PMS - **YES**
- FYP Demo reviewed by the Advisor - **YES**
- FYP Demo uploaded on PMS - **YES**
- Course feedback of all courses submitted on CMS - **YES**

Some facts and figures

Demographics show that the world population is expected to cross 9 billion by 2050 (UNFAO. 2018).

This population increasement will causes

Food scarcity

FAO estimates about 9.3% were suffering from chronic undernourishment in 2018.

There are 11 million people undernourished in developed countries (FAO 2015, see IFPRI 2016 and Rosen et, al, 2016)

Deforestation

Between 1990 and 2016, the world lost 502,00 square miles (1.3 million square kilometers) of forest,

according to the World Bank

Climate Change

The global temperature has increased 1.9 degrees with arctic ice melting to lower 12.85% with 3.3% increase in sea levels, according to

Global Climate Change NASA

Water Shortage

According to UNWater organization:

Over 2 billion people live in countries experiencing high water stress, and about 4 billion people experience severe water scarcity.

(march-2019)

Land Consumption

68% of the World population projected to live in urban areas by 2050, says UN

Today, it is 55%.

Problem statement

Problem statement

Cost effective statistics model for monitoring Earth's geographical changes is required to comprehend the related issues as the annual population growth rate is alarmingly increasing.

Proposed Solution

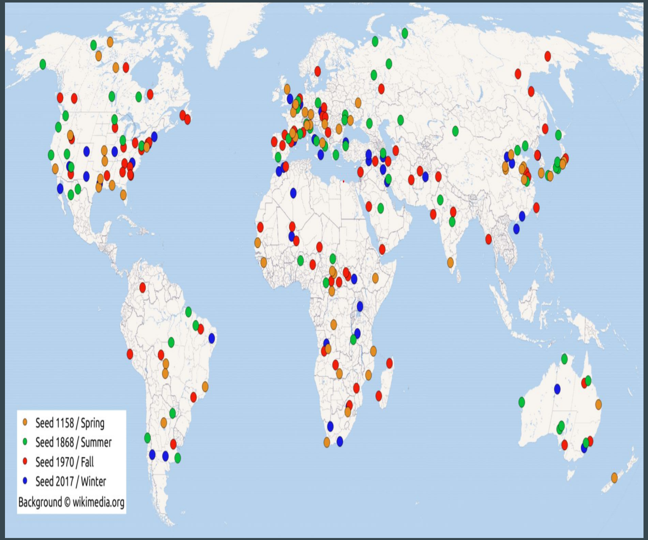
Proposed Solution

We propose a low-cost and generalized solution by extracting Earth's land cover map for Satellite imagery using Deep Learning and Computer Vision to obtain accurate stats through open sourced website for any location

Understanding the datasets

SEN12MS - A Curated Dataset Of Georeferenced Multi-spectral Sentinel-1/2

- 180,662 triplets of dual-pol synthetic aperture radar (SAR) image patches, multi-spectral Sentinel-2 image patches, and MODIS land cover maps
- With geolocation encoded in .tif file
- SAR Image with VV+VH (dual polarization)
- Optical Images of 13 channels at 10m resolution
- Land cover information with 4 schemes at 500m resolution
- Four seasons:
  - ROIs1158 spring
  - ROIs1868 summer
  - ROIs1970 fall
  - ROIs2017 winter



Schmitt, M., Hughes, L.H., Qiu, C. and Zhu, X.X., 2019. SEN12MS--A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. arXiv preprint arXiv:1906.07789.

SEN12MS - A Curated Dataset Of Georeferenced Multi-spectral Sentinel-1/2

Each pixel corresponds to a class from a specific scheme.

- Forests
- Shrublands
- Savanas
- Grasslands
- Wetlands
- Croplands
- Urban and Built-up Lands
- Permanent Snow and Ice
- Barren
- Water Bodies

Class	IGBP value	LCCS LC value	LCCS LU value	LCCS SH value
Evergreen Needleleaf Forests	1	11	-	-
Evergreen Broadleaf Forests	2	12	-	-
Deciduous Needleleaf Forests	3	13	-	-
Deciduous Broadleaf Forests	4	14	-	-
Mixed Broadleaf/Needleleaf Forests	-	15	-	-
Mixed Broadleaf Evergreen/Deciduous Forests	-	16	-	-
Mixed Forests	5	-	-	-
Dense Forests	-	-	10	10
Open Forests	-	21	20	20
Sparse Forests	-	22	-	-
Natural Herbaceous	-	-	30	-
Dense Herbaceous	-	31	-	-
Sparse Herbaceous	-	32	-	-
Shrublands	-	-	40	40
Closed (Dense) Shrublands	6	41	-	-
Open (Sparse) Shrublands	7	43	-	-
Shrubland/Grassland Mosaics	-	42	-	-
Woody Savannas	8	-	-	-
Savannas	9	-	-	-
Grasslands	10	-	-	30
Permanent Wetlands	11	-	-	-
Woody Wetlands	-	-	-	27
Herbaceous Wetlands	-	-	-	50
Herbaceous Croplands	-	-	36	-
Croplands	12	-	-	-
Urban and Built-Up Lands	13	-	9	-
Cropland/Natural Vegetation Mosaics	14	-	-	-
Forest/Cropland Mosaics	-	-	25	-
Natural Herbaceous/Croplands Mosaics	-	-	35	-
Tundra	-	-	-	51
Permanent Snow and Ice	15	2	2	2
Barren	16	1	1	1
Water Bodies	17	3	3	3

Schmitt, M., Hughes, L.H., Qiu, C. and Zhu, X.X., 2019. SEN12MS--A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. arXiv preprint arXiv:1906.07789.

DFC 2020 dataset

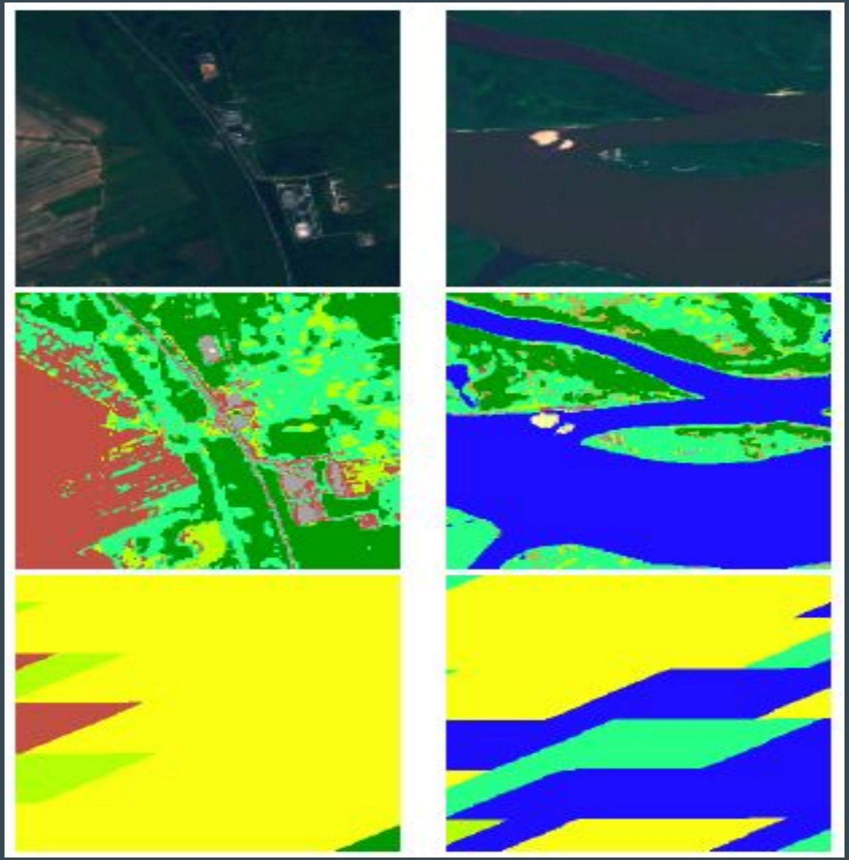
- 6,114 patches of quadruplets, i.e.,
  - Sentinel 1 SAR,
  - Sentinel 2 Optical Images,
  - Modis Landcover Data,
  - Finely labeled Landcover data.
- Its landcover is actually simplified version of IGBP scheme of Modis.
- The .tif files do not have geolocation.

IGBP Class Number	IGBP Class Name	Simplified Class Number	Simplified Class Name
1	Evergreen Needleleaf Forest	1	Forest
2	Evergreen Broadleaf Forest		
3	Deciduous Needleleaf Forest		
4	Deciduous Broadleaf Forest		
5	Mixed Forest		
6	Closed Shrublands	2	Shrubland
7	Open Shrublands		
8	Woody Savannas	3	Savanna
9	Savanna		
10	Grasslands	4	Grassland
11	Permanent Wetlands	5	Wetlands
12	Croplands	6	Croplands
14	Cropland/Natural Vegetation Mosaics		
13	Urban and Built-up Lands	7	Urban/Built-up
15	Permanent Snow and Ice	8	Snow/Ice
16	Barren	9	Barren
17	Water Bodies	10	Water

Schmitt, M., Prexl, J., Ebel, P., Liebel, L. and Zhu, X.X., 2020. Weakly Supervised Semantic Segmentation of Satellite Images for Land Cover Mapping--Challenges and Opportunities. arXiv preprint arXiv:2002.08254.

Differences between Datasets with two examples

- First row shows the Sentinel 2 image with bands 4, 3, 2 stacked with each other as RGB image.
- The second row shows the finely labeled application level segmentation maps of DFC2020 dataset.
- Third row shows coarse labels obtained from Modis satellites following simplified IGBP scheme.

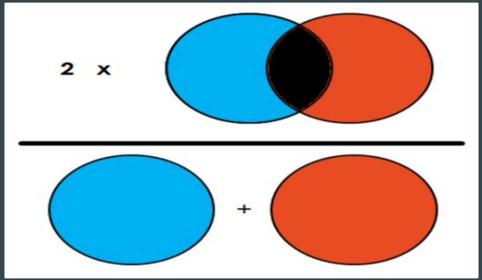


Methodology

Evaluation

- F1 Score/ Dice Coefficient:
    - Weights the true positives more than IoU

$$\frac{2TP}{2TP + FP + FN}$$

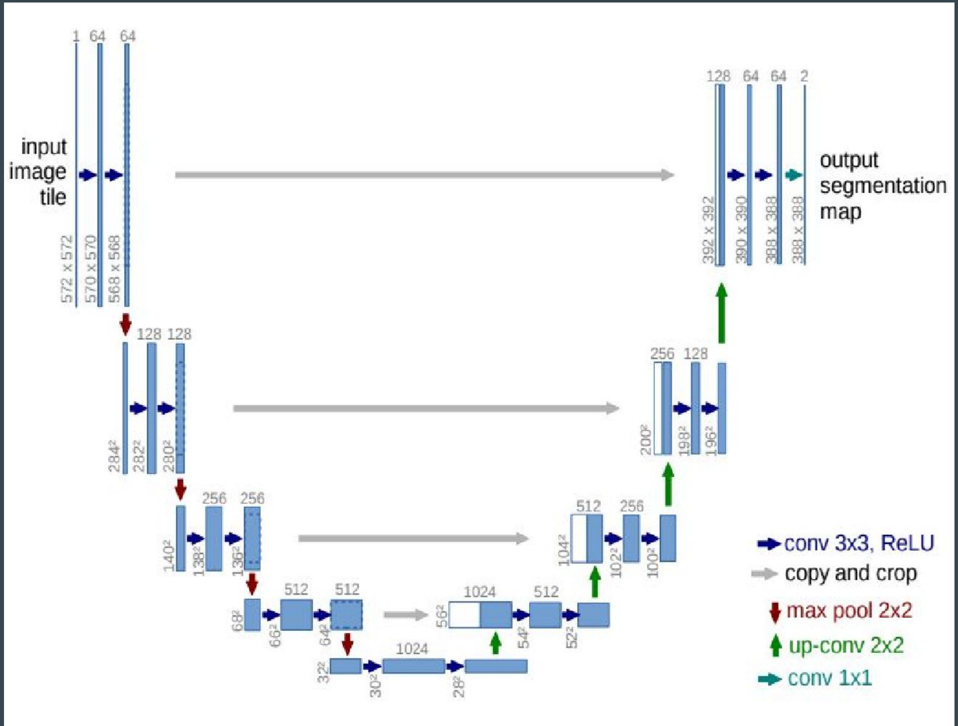


- Intersection over Union/ Jaccard Index:
    - Punishes more for a wrongly labeled pixel

$$\frac{TP}{TP + FP + FN}$$

1 Baseline Semantic Segmentation Network

- Used simple U-Net architecture:
  - Contracting Path: low feature representation of the input.
  - Expanding Path: combines features and spatial info through up-convolutions and concatenations from high resolution features from contracting path.
- Did preprocessing, for extracting four channels of Sentinel 2 i.e., RGB and IR



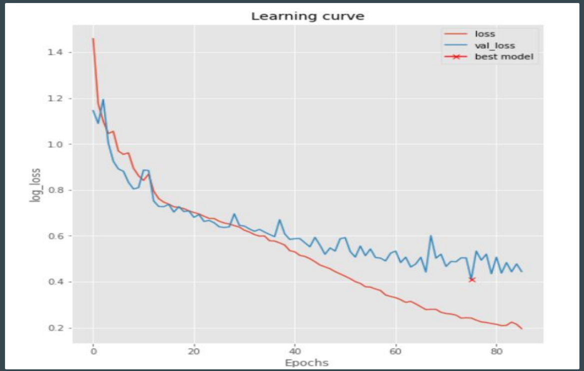
1 Baseline Semantic Segmentation Network

Results:

- From 256x256x4 (input)
- To 256x256x8 (output)

IOU accuracy: 38.12%

F1 score: 42.23%

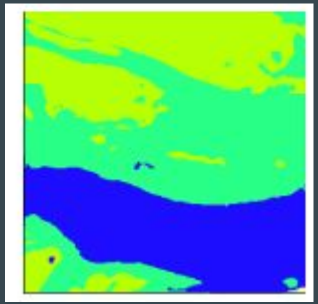


Learning Curve

Plotting:

The result of plot of a random image with its prediction.

Sentinel 2 Image



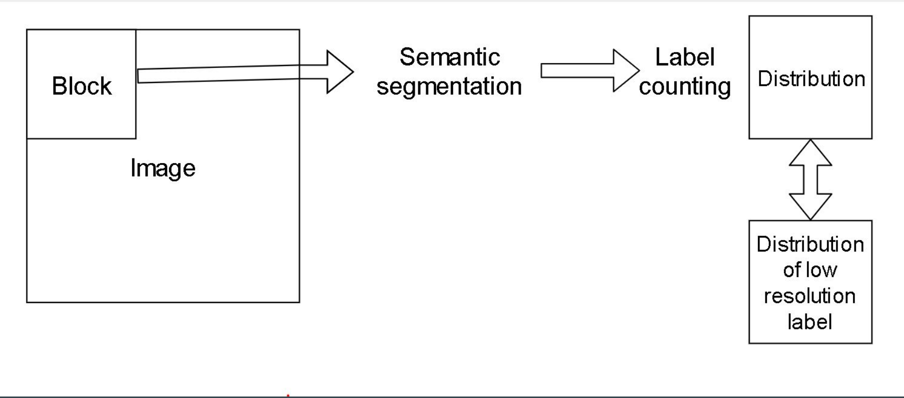
Predicted LC map

2 Label Super-Resolution Network

- Minimizes distance between
  - Distribution of actual low to high class segmentation maps
  - Distribution of predicted labels
- Achieved with Super-Resolution Loss.

$$D(p_{\text{net}}, p_{\text{coarse}}) = \log p_{\text{net}}(C_\ell | X) \sim \text{const} - \frac{1}{2} \frac{(\mu_\ell - \eta_{\ell,z})^2}{\sigma_\ell^2 + \rho_{\ell,z}^2}$$

Simple Summary:



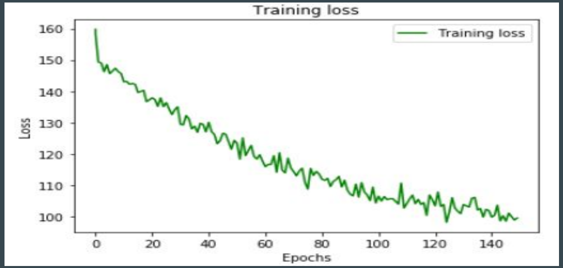
Malkin, K., Robinson, C., Hou, L., Soobitsky, R., Czawlytko, J., Samaras, D., Saltz, J., Joppa, L. and Jojic, N., 2018. Label super-resolution networks.

2 Label Super-Resolution Network

Results:

IOU accuracy: 19.71%

F1 score: 28.76%

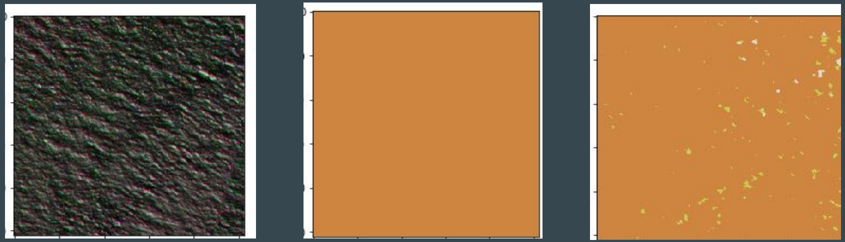


Shortcomings:

- Super-resolving from 30m to 10m.

Plotting:

The result of plot of a random image with its prediction.

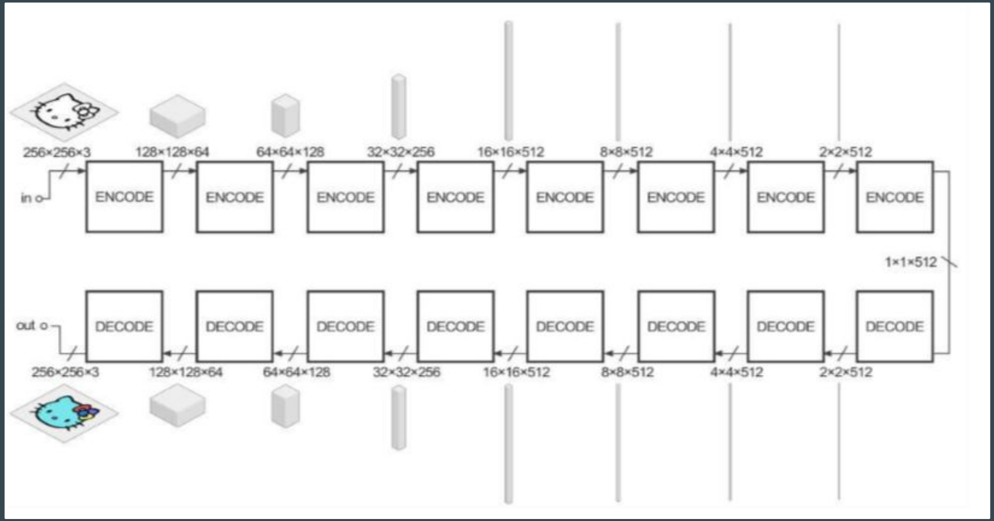


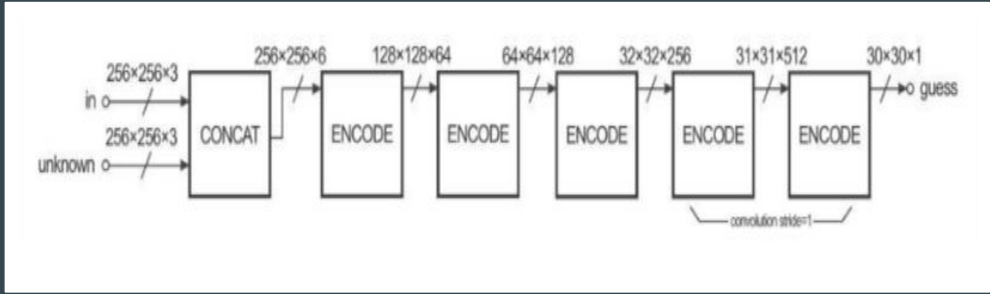
Sentinel 2 Image   Actual LC   Predicted LC

Malkin, K., Robinson, C., Hou, L., Soobitsky, R., Czawlytko, J., Samaras, D., Saltz, J., Joppa, L. and Jojic, N., 2018. Label super-resolution networks.

3 Conditional Generative Adversarial Networks

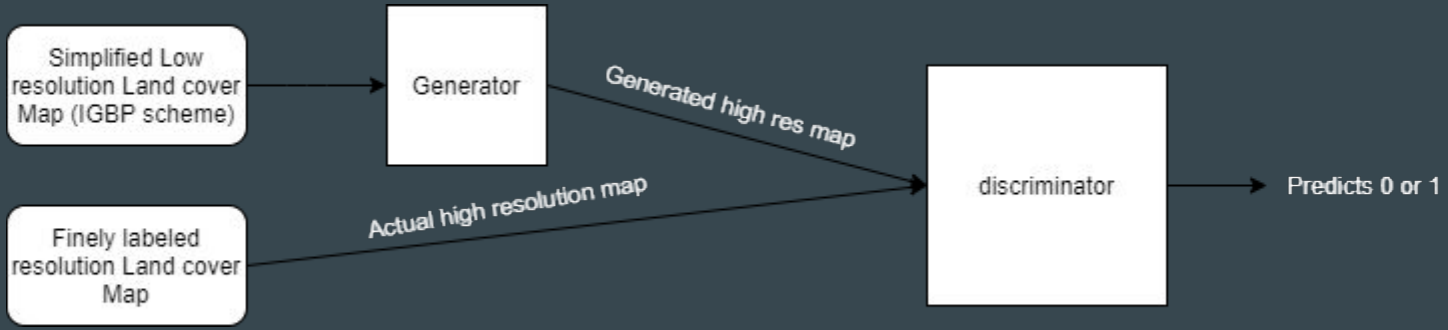
- GANs typically have generator and a discriminator
- Conditional GANs are conditioned on the Input Image.
- The generator has a semantic segmentation network for a low dimensional feature representation of the input image and then reconstructs it.
- The discriminators distinguishes between the generated output and the actual output with the help of binary_crossentropy.





3 Conditional Generative Adversarial Networks (CGANs) Low to high

- First the discriminator is trained with actual finely labeled LC map with label 1 (True)
- Then for the next half time it is given generated LC map with label 0 (False)
- Then it is set to untrainable
- Now the generator is trained with discriminator model with generated LC and label 1 (True).
- The cycle goes on to better fool the discriminator.



Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

3 Conditional Generative Adversarial Networks (CGANs) Low to high

Results:

- From 256x256x8 (input) Low res
- To 256x256x8 (output) High res

IOU accuracy: 54.02%

F1 score: 62.99%

Shortcomings:

- Limited scope as no visual information of image

Plotting:

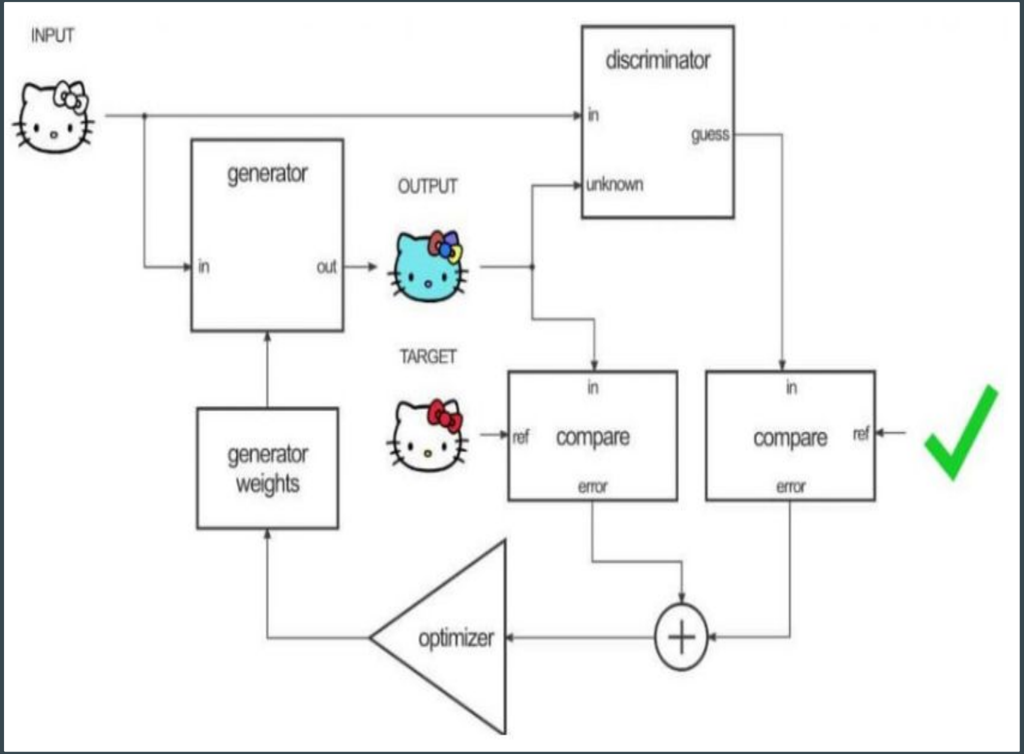
The result of plot of three randomly chosen samples with its prediction.

1_Predicted high resolution Land Covers

2_Actual high resolution Land Covers

4 Conditional Generative Adversarial Networks (Pix2Pix)

- Used for Image to Image translation
- The generator has skip connections
- The input was the RGB and IR channels
- It predicted the Land cover maps of 8 channels (i.e., equal to the number of classes)



Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

4 Conditional Generative Adversarial Networks (Pix2Pix)

Results:

Its testing accuracies are as follows:

IOU accuracy: 68.77%

F1 score: 77.57%

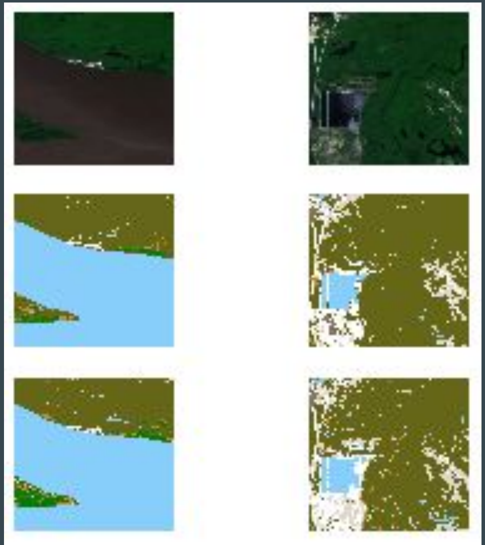
Plotting:

The result of plot of two randomly chosen samples with its prediction.

1_RGB Image

2_Predicted Map

3_Actual Map



Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

5 Bag of Color - Algorithm

- Inspired by computer vision technique of Bag of Visual Words (used for classification)
- Patch-Wise Classification with variable sized patches (unlike traditional approach)
- Patch pixels may comprised of any pixels, not necessarily neighbor pixels (unlike traditional approach)
- Patch-wise classification leads to perfect semantic segmentation

Training Scores:

F1: 87.03%, IOU: 79.28%

Testing Scores:

F1: 67.77%, IOU: 77.35%

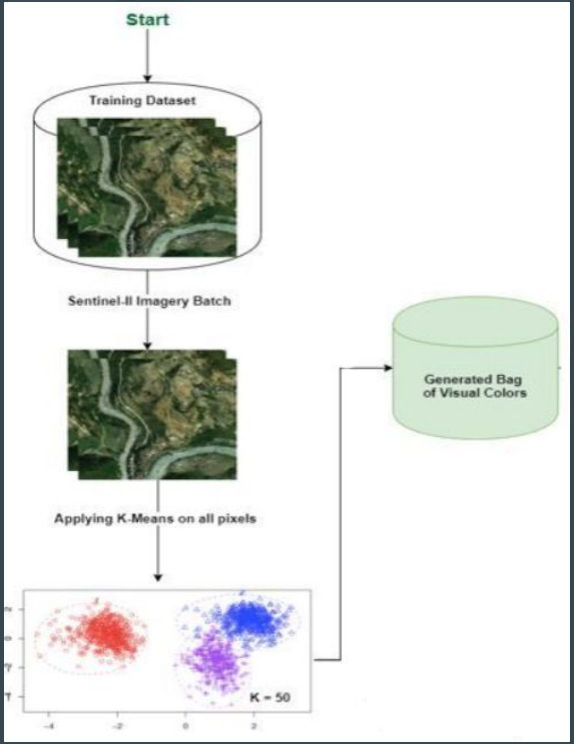
Wengert, C., Douze, M. and Jégou, H., 2011, November. Bag-of-colors for improved image search. In Proceedings of the 19th ACM international conference on Multimedia (pp. 1437-1440).

Bag of Color - Algorithm Steps

- Vocabulary Generation
- Image Division
- Histogram Generation
- Model Training
- Model Prediction

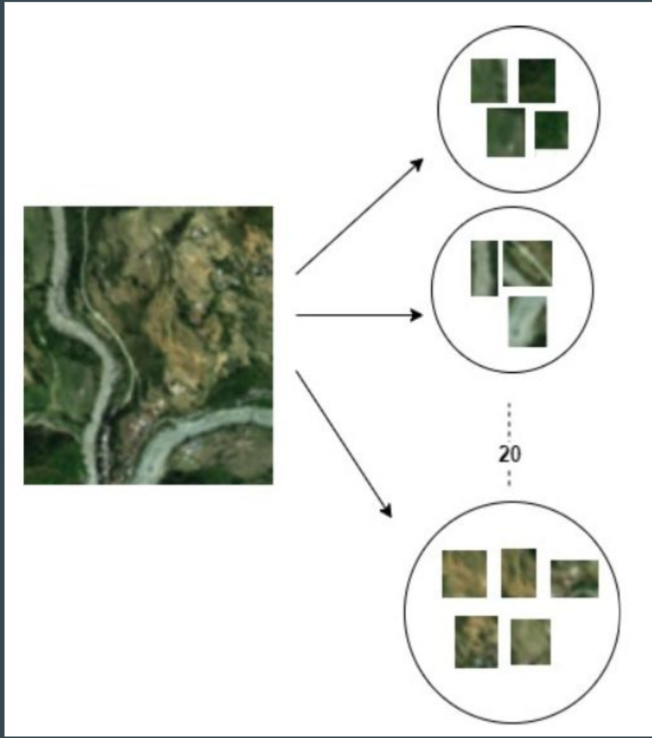
Bag of Color - Vocabulary Generation

- Load all image pixels
- Normalized pixel values (R,G,B,IR) of the sentinel imagery are clustered into n-clusters (we used 80 clusters)
- Makes it less prone to noise, clouds, change in seasons, intra-class variations.



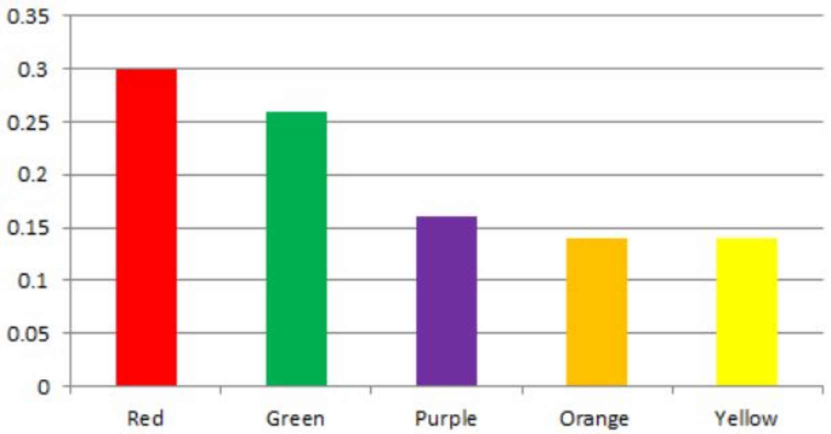
Bag of Color - Image Division

- Divide image into cluster patches (We divided into 20 patches)
- Patch can be of any sized and its pixels can be at any spatial location (variable)
- Maximum labels in image can be 8 (number of classes), so 20 is chosen to have more flexibility in segments



Bag of Color - Histogram Generation

- For each pixel in a patch, we compute its global class (cluster) from vocabulary
- Obtain count of each vocabulary cluster (80 in our case) for each cluster
- Normalize histogram

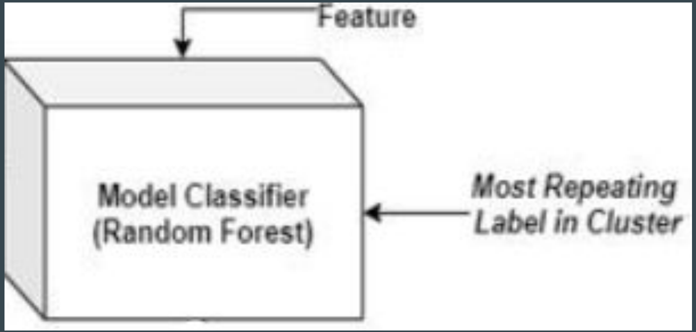


Bag of Color - Classifier

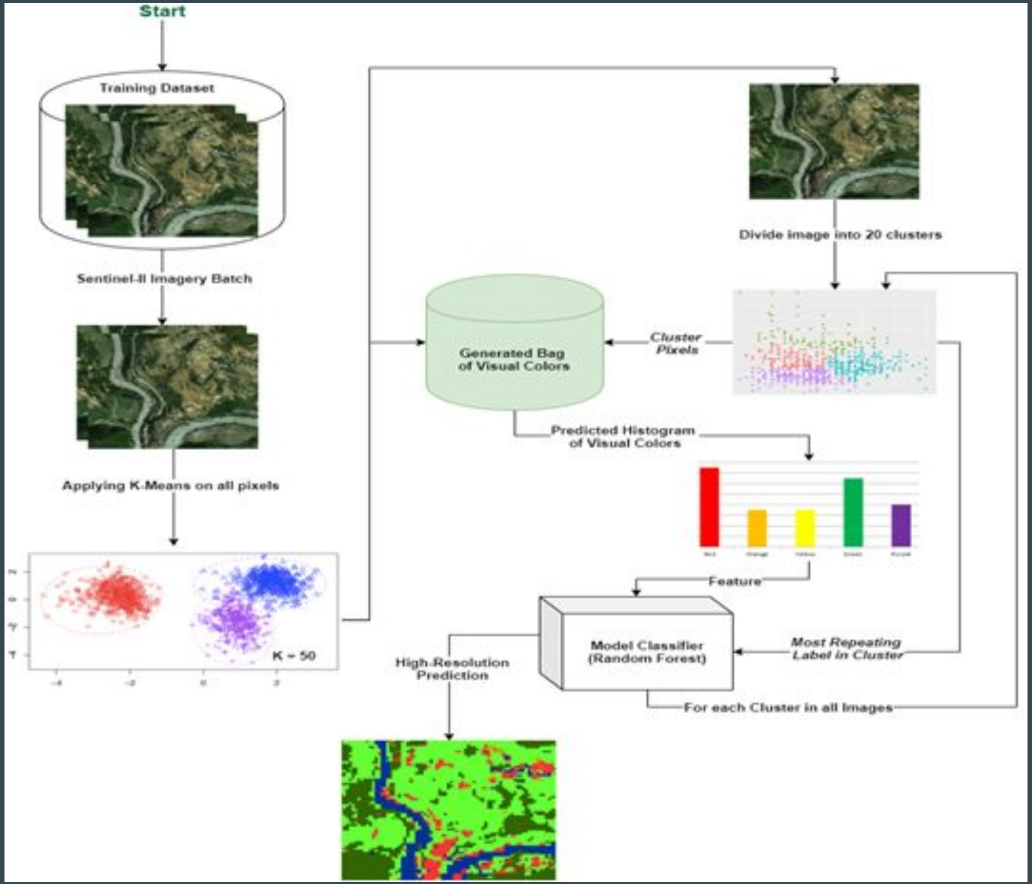
- For each pixel in a patch, we determine their prediction labels and assign maximum repeating label to whole patch
- X (Features): Normalized Histogram

  Y (Labels): Label for cluster

- Train histogram features (X) on computer cluster label (Y) on any classifier
- We used Random Forest with 1000 estimators



Bag of Color - Overall Pipeline



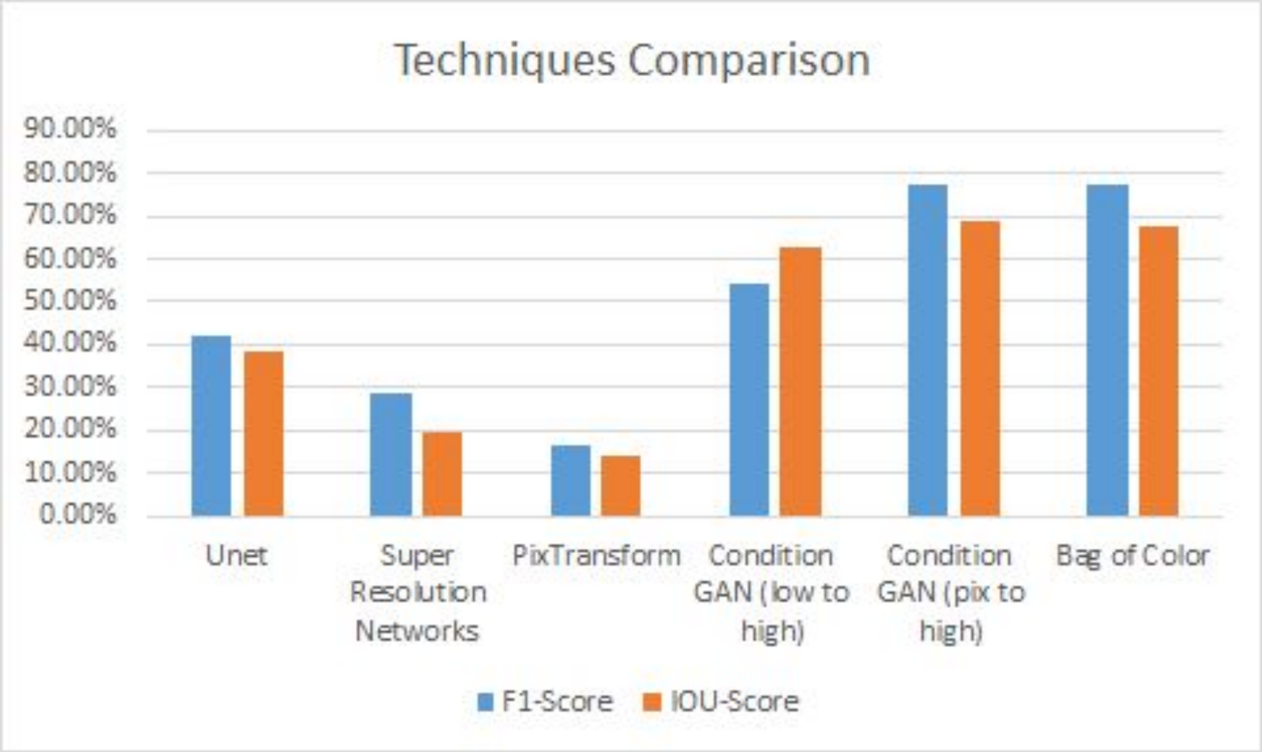
Bag of Color - Comparisons

Ground Truth (High-Resolution Land Cover)

Low-Resolution Land Cover Map

Predicted High-Resolution Land Cover Map

Summarizing all methods



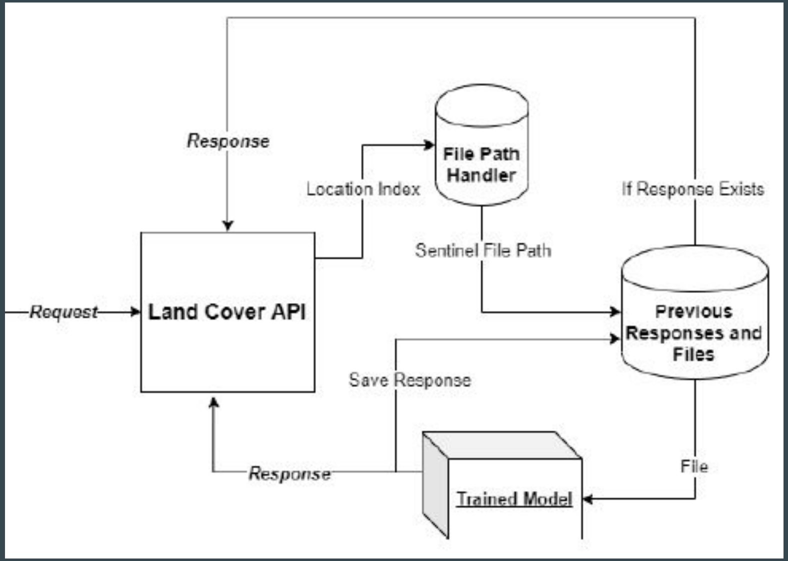
Extras:

- Obtaining Sentinel Images from Google Earth Engine
- Labelling Earth regions using QGIS and ArcMap
- Trying different statistical distributions for super resolution

Web-Based interface for user interaction

Model connection with Website

- Communicates through REST APIs
- Backend Server kept independent
- Built in Flask (Python) that connects with trained model
- To speed up the response, we save the fresh prediction on server to respond quickly (without making prediction again)
- Convert responses to a readable and useful form

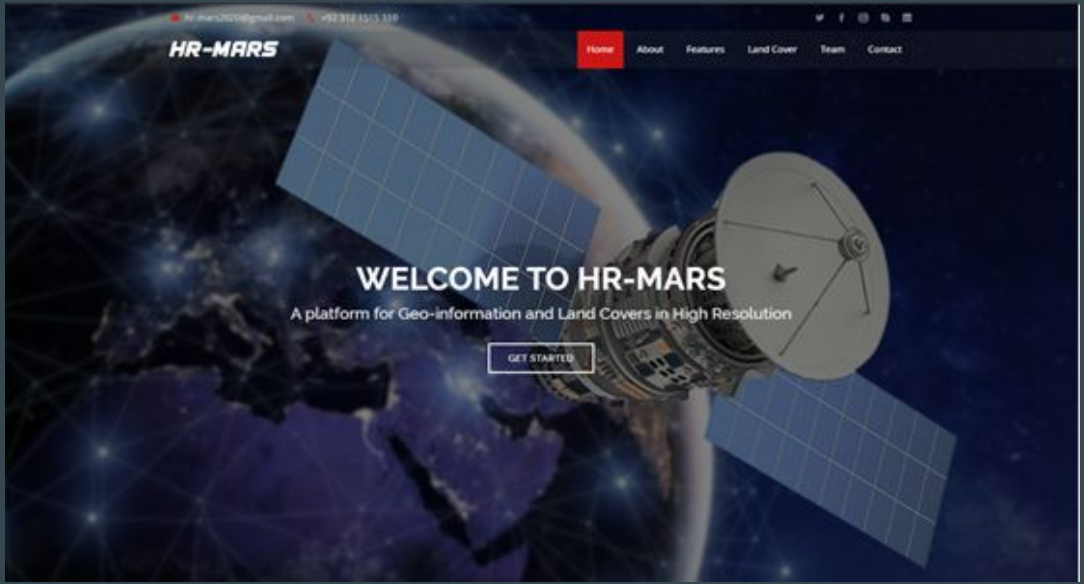


APIs Provided by Server

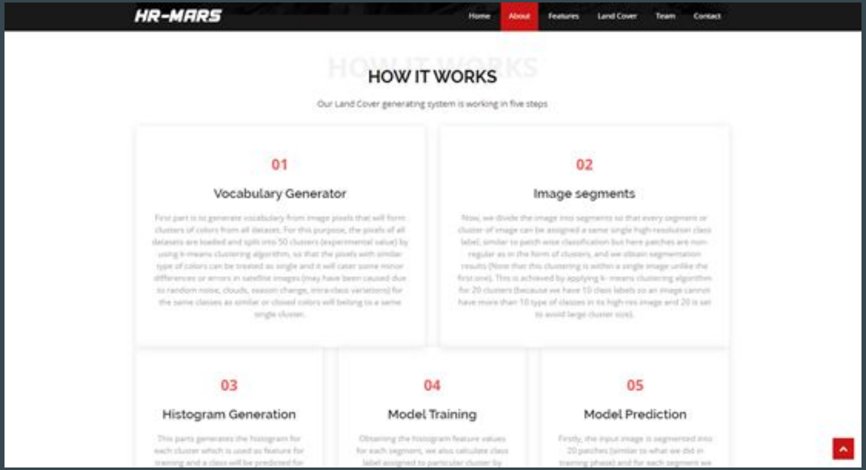
- Extract geographical coordinates of all available land cover maps
- Obtain low-resolution land cover map for selected point
- Obtain high-resolution land cover map for selected point
- Compute statistical distribution for low-resolution map of the selected point
- Compute statistical distribution for high-resolution map of the selected point

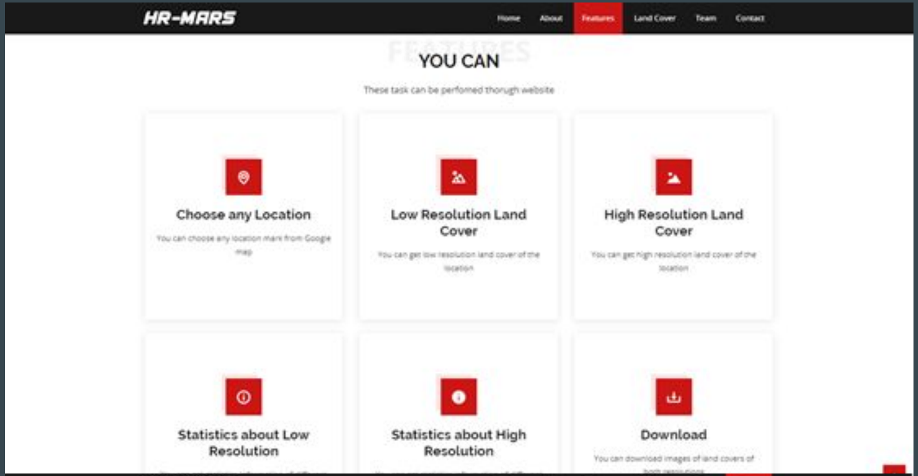
Website

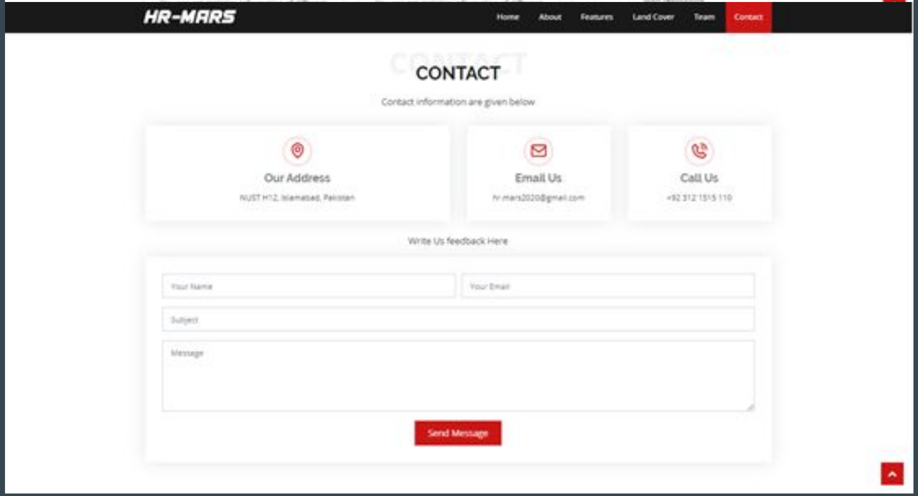
- Landing Page
  - ➔ Basic information
  - ➔ Features descriptions
  - ➔ Team
  - ➔ Contact Information
  - ➔ Feedback



Website

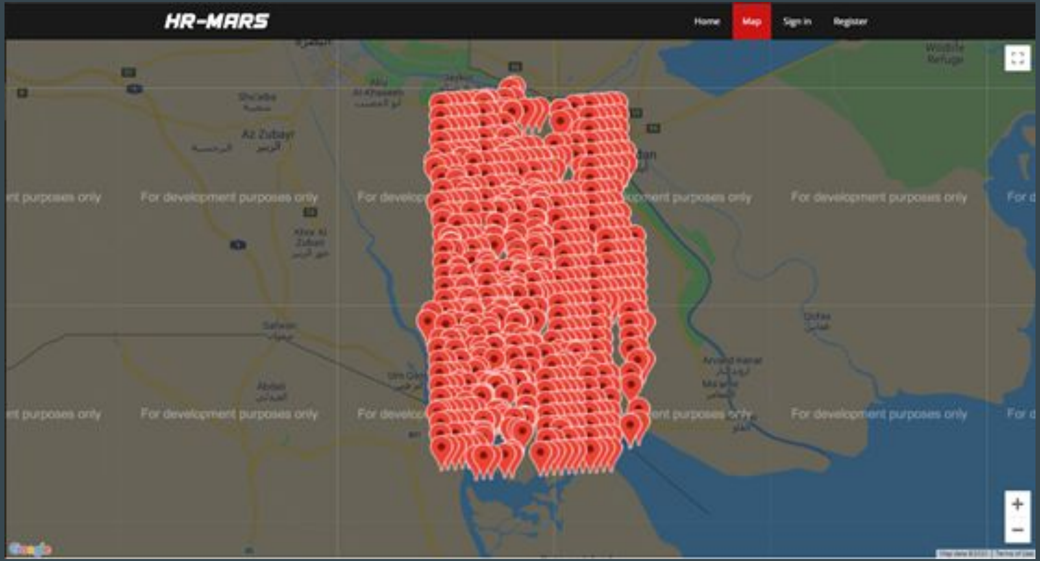

HR-MARS
Home
About
Features
Land Cover
Team
Contact
HOW IT WORKS
Our Land Cover generating system is working in five steps
01
Vocabulary Generator
02
Image segments
03
Histogram Generation
04
Model Training
05
Model Prediction


HR-MARS
Home
About
Features
Land Cover
Team
Contact
YOU CAN
These task can be performed thorugh website
Choose any Location
You can choose any location mark from Google map
Low Resolution Land Cover
High Resolution Land Cover
Statistics about Low Resolution
Statistics about High Resolution
Download

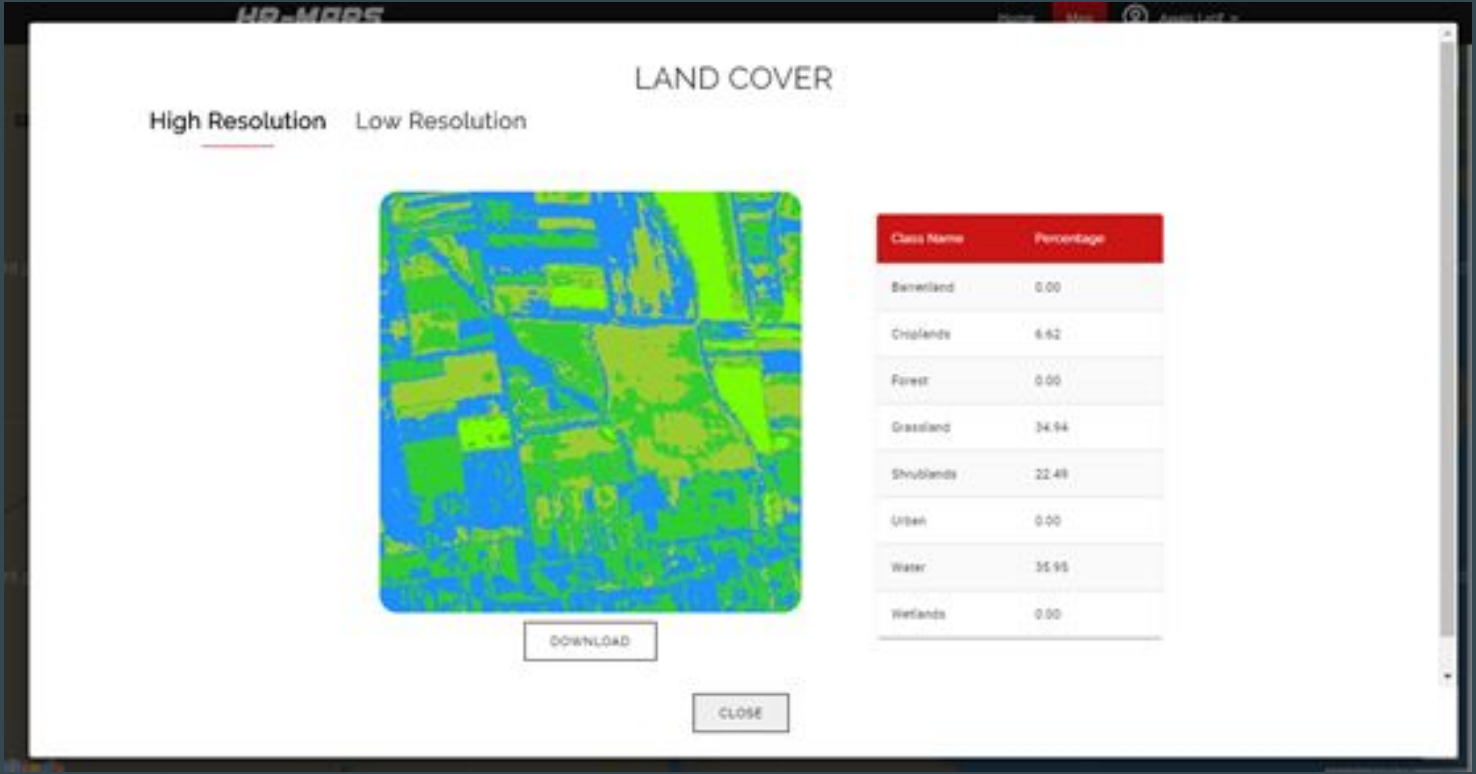

HR-MARS
Home
About
Features
Land Cover
Team
Contact
CONTACT
Contact information are given below
Our Address
NUST H12, Islamabad, Pakistan
Email Us
Call Us
Write Us feedback Here
Your Name
Your Email
Subject
Message
Send Message

Website

- Google Map Page
  - ➔ Select Location
  - ➔ High and low resolution land cover
  - ➔ Statistics Information
  - ➔ Download land cover



Website



LAND COVER
High Resolution
Low Resolution
Class Name
Percentage
Barrenland
0.00
Croplands
6.62
Forest
0.00
Grassland
34.94
Shrublands
22.49
Urban
0.00
Water
35.95
Wetlands
0.00
DOWNLOAD
CLOSE

Future work

- Forum for discussion
- Upload satellite image manually to get land cover
- Host Competitions (Different users upload their trained model for prediction of land cover).

The End