Resolving Food Security Through Technology

Maria Zubair
Punjab Information Technology Board
Lahore, Pakistan

Introduction

Food, Agriculture, Technology

Food

- Historically Humans secure food:
  - Hunting and Gathering
  - **Agriculture**
- 1/9 people in the world today are under nourished
- Poor nutrition causes nearly 45% of deaths in children under five.

Agriculture

- Agriculture is the single largest employer for providing livelihood to 40% of today's global population.
- 500 million small farms worldwide provide up to 80% food consumed by the developing world.
- Zero Hunger: Sustainable Development Goal

Technology

- Emerging Technologies
  - Big Data
  - Agriculture Drones
  - Deep Learning
  - UAV’s, LIDAR, Radar
  - Multispectral Satellite Imagery
  - Machine Learning
  - Cloud Computing
  - Space Technology

Food, Agriculture and Technology

Food Security → Sustainable Agriculture

- Multispectral Satellite Imagery
- Machine Learning
- Big Data
- Data Modelling
- Volunteer Computing

Why are we doing this?

- Pakistan like many others is an agriculture based economy country.
- Developed countries are using state of the art technologies for agriculture.

So we need to cope up

- Develop a cheap and sustainable solution for agriculture.

Methodology

Conventional method of classification, using GIS tools

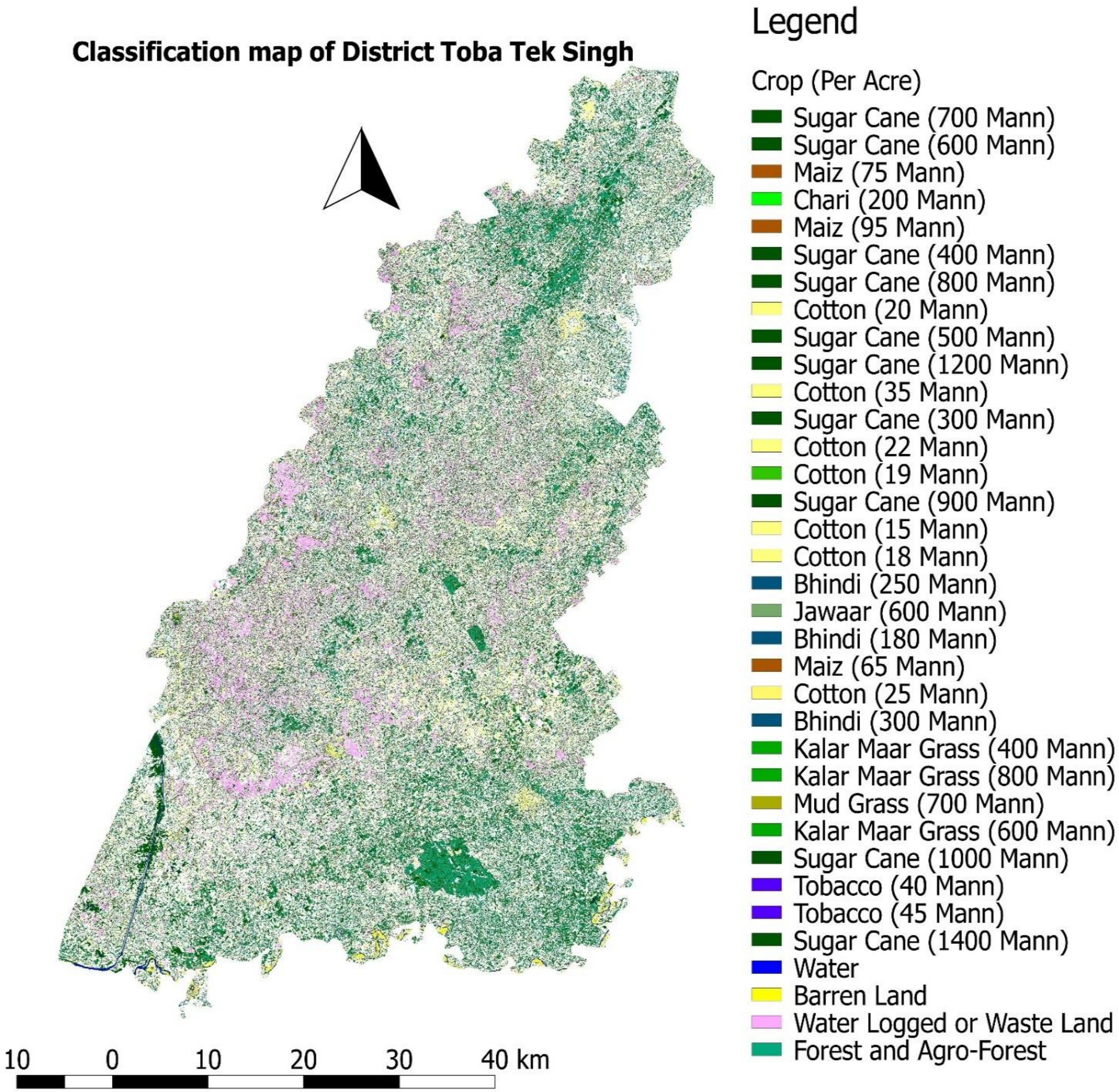
Introduction

- Acquisition Sentinel 2(multispectral imagery) data
- Field Survey
- Sentinel 2 preprocessing
  - Atmospheric Corrections
  - Layer Stacking, Mosaicking and Clipping
  - Filtering and Histogram Correction
- Signature Plotting using field data
  - Crop Library
- Signature based Classification (Minimum Distance Classifier – Euclidean Distance)
- Acreage and yield calculations
- Accuracy Assessment

Results

- 92% Accuracy
- Classified 32 different variety of crops.
- Area based Crop yield prediction model.

Crop classification using ground sampling



Class	PixelSum	Percentage %	Area [metre^2]
1.0	46168	0.564241914199	18467200.0
2.0	75983	0.928625744381	30393200.0
3.0	43918	0.536743553712	17567200.0
4.0	96900	1.18426272496	38760000.0
5.0	35755	0.436979501867	14302000.0
6.0	206664	2.52574274294	82665600.0
7.0	42043	0.513828253307	16817200.0
8.0	475773	5.81465665058	190309200.0
9.0	130719	1.59758141531	52287600.0
10.0	59214	0.723683519047	23685600.0
11.0	31980	0.390843363717	12792000.0
12.0	111336	1.36069220584	44534400.0
13.0	91908	1.12325302916	36763200.0
14.0	149997	1.83318736796	59998800.0
15.0	14446	0.176551695818	5778400.0
16.0	69694	0.851764771447	27877600.0
17.0	119253	1.45744977027	47701200.0
18.0	167608	2.04842009086	67043200.0
19.0	117675	1.43816425345	47070000.0
20.0	35263	0.43096652704	14105200.0
21.0	46617	0.569729364803	18646800.0
22.0	7993	0.0976863979421	3197200.0
23.0	64778	0.791683909157	25911200.0
24.0	25912	0.316683340858	10364800.0
25.0	74699	0.912933346663	29879600.0

Above is Classification report of classes related pixels showing each class percentage and area in meter^2 (Each class represent the separate class type)

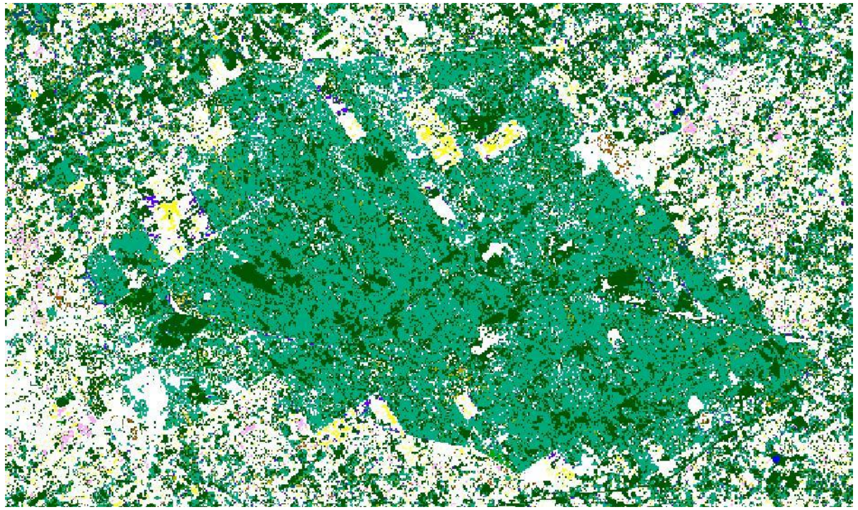
Kamalia (Reserve) Forest

- Water
- Barren Land
- Water Logged or Waste Land
- Forest and Agro-Forest
- Sugar Cane
- Cotton
- Maiz
- Bhindi
- Chari
- Jawaar
- Mud Grass
- Kalar Maar Grass

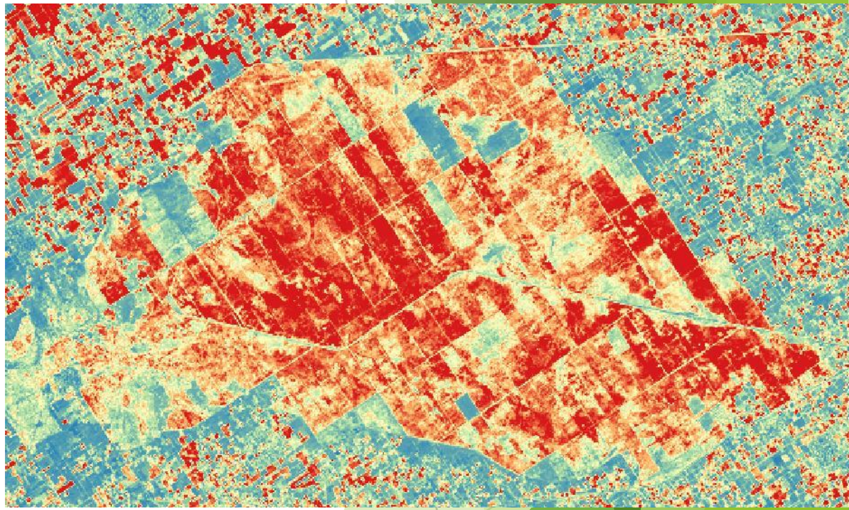
NDVI

0.0564	Low Health Crops
0.175	
0.293	
0.411	
0.529	High Healthy Crops

Natural Color (RGB) image showing Forest area



Classified image showing main forest class (light green) also the few crop patches (dark green) and barren/empty land (yellow) within forest.



NDVI image showing health profile of forest

Issues

- Process is Computation intensive and takes time
- Algorithm only works with certain number of data points.
- No simplified representation of classified data.
- Data Storage
- Frequent availability of data

Food Security

Food Security

- **Food security** exists when all people have physical and economic access to sufficient, safe and nutritious **food** that meets their dietary needs for an active and healthy life.
- It is a measure of the availability of **Food** and Individuals accessibility to it, where accessibility includes affordability.
- FAO identified the four pillars of Food Security as
  - Availability
  - Access
  - Utilization
  - Stability

Agri@HOME (Methodology 2)

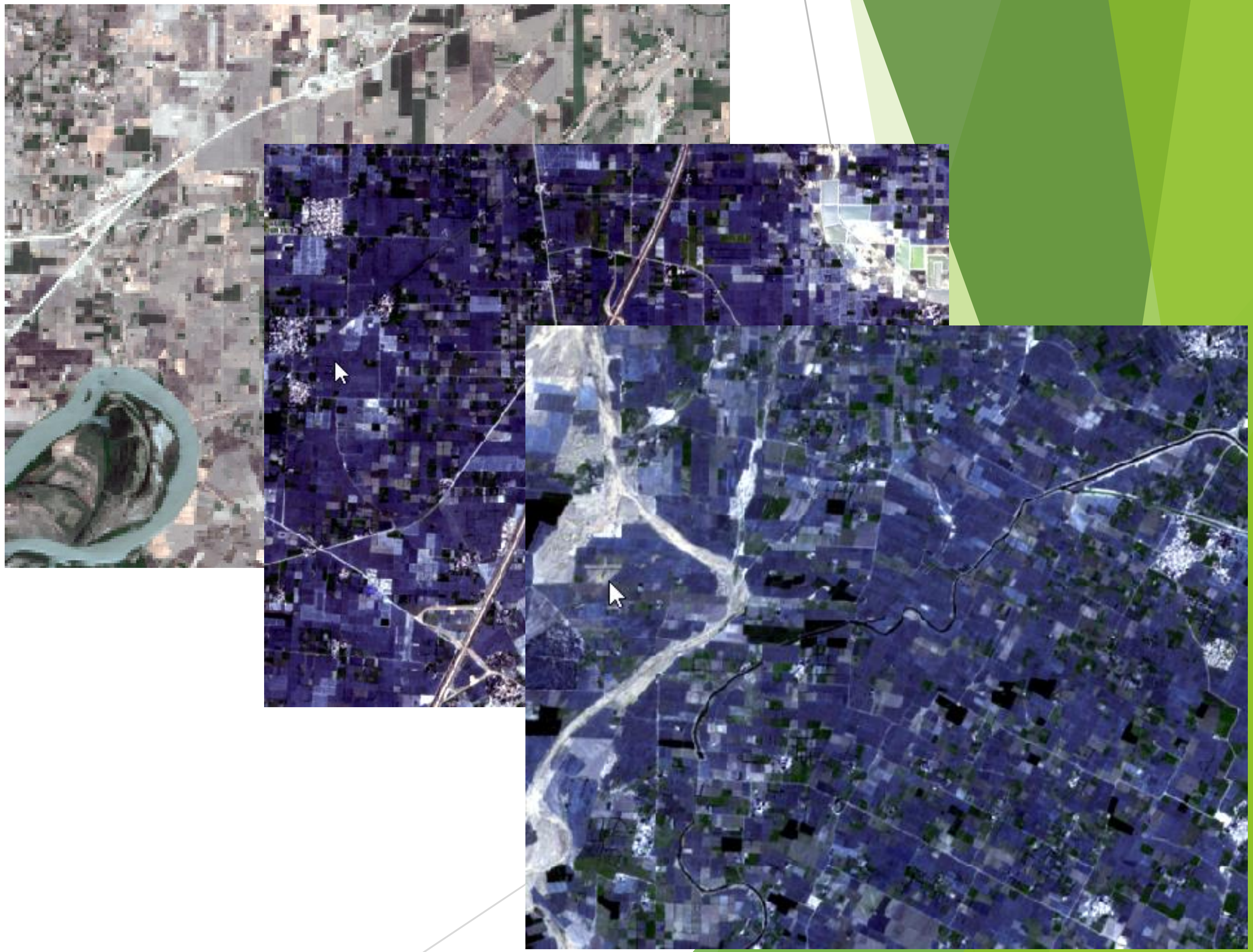
Volunteer Computing based Machine Learning trained Crop Classification Solution for Agriculture

Project

A. Crop Signature Library and Satellite Imagery
B. Crop classification Algorithm
C. Crop Yield calculation models
D. Volunteer Computing

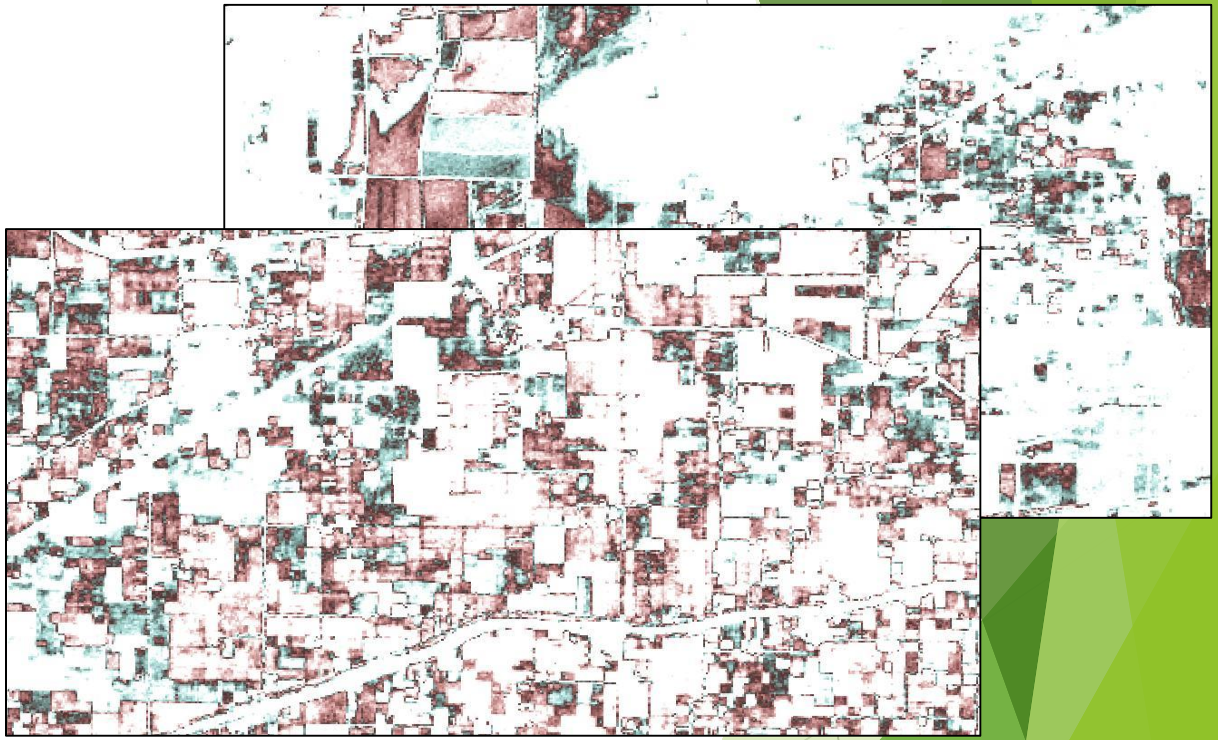
Crop Signature Library and Satellite Imagery

- Sentinel 2A
- Spatial resolution of **10** m
- Multi-spectral data with **11** bands
- Crop Signature Collection using extension workers who already work in the field as advisory to farmers.
- Signatures are marked per season and data is collected periodically.
- Data is saved in database with a unique nomenclature as each tile is subdivided into 100 cell tiles for processing.



Crop Classification Algorithm

- Crop Health Estimation using **NDVI**
- Built Up area Elimination using **NDBI**
- Field Segmentation
  - Based on signatures (NDVI)
  - Existing segments(fields)
- **Random Forest** using **Minimum Distance** Classifier



Crop Yield Model

1. NDVI(Normalized Difference Vegetation Index) based Model

   Formula: Yield = (**acreage** x **seed type**) x %(scaled **NDVI**)

   acreage is area used by a specific crop variety,

   seed type is the estimated yield produced by a seed and

   ***NDVI** is a health vegetation index

2. Burke and Lobell Model (SYCM)
   - **Simulations** to output
     - Leaf Area Index(**LAI**),
     - biomass,
     - Green chlorophyll vegetation index(**GCVI**)
     - water stress and weather.
   - **Statistical model** using above variables
   - **Regression model** based on calculated yield to satellite based actual yield data.

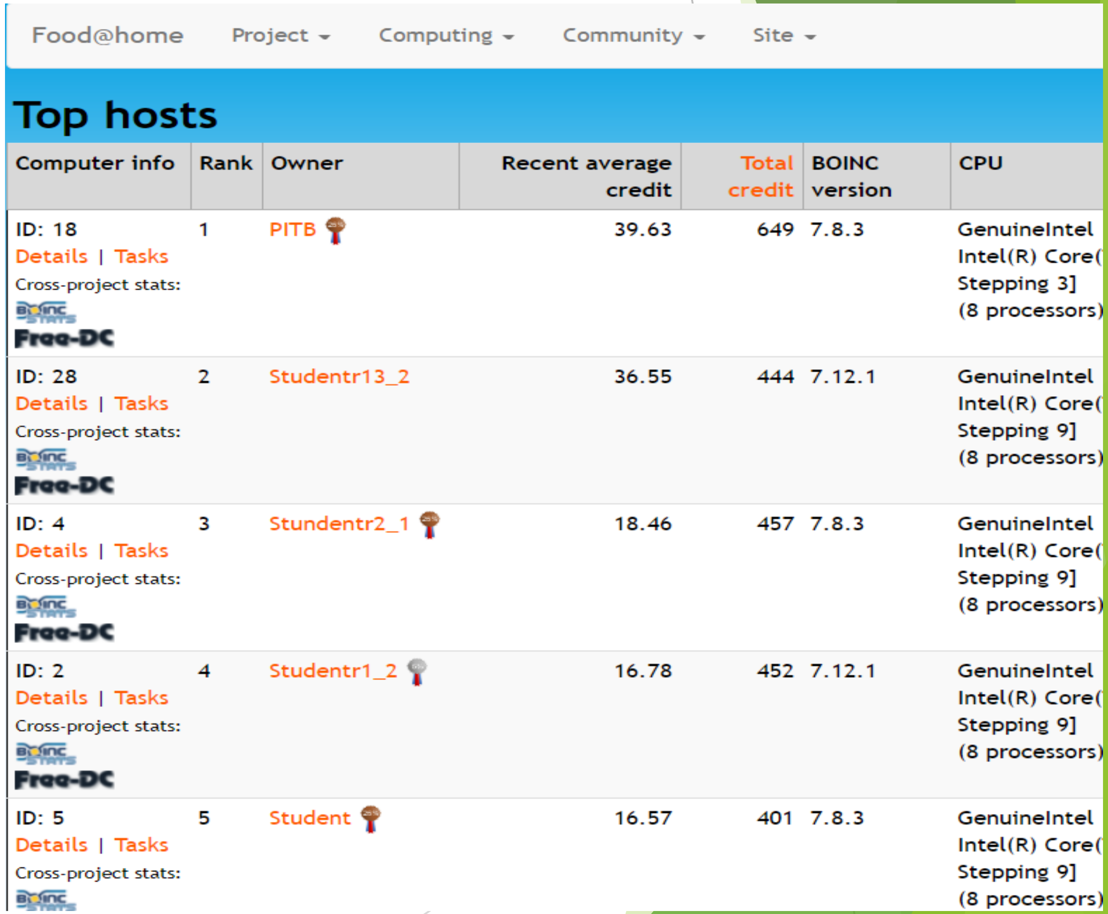
Implementing this

*NDVI = $\rho NIR - \rho Red / \rho NIR + \rho Red$

Volunteer Computing

- Using BOINC (Berkeley Open Infrastructure for Network Computing)
- Customized architecture for Agri@HOME to use idle resources for crop classification and yield estimation
- 10 hours task takes approximately 25 minutes.
- Multiple users across the world can contribute by installing this piece of code in their machines.
- A WEBSITE www.foodathome.com is a “work in progress”. It will work as a platform to see classified crops and their yield and contribute by lending computation cycles by installing the software on your machine.

Volunteer Computing Stats

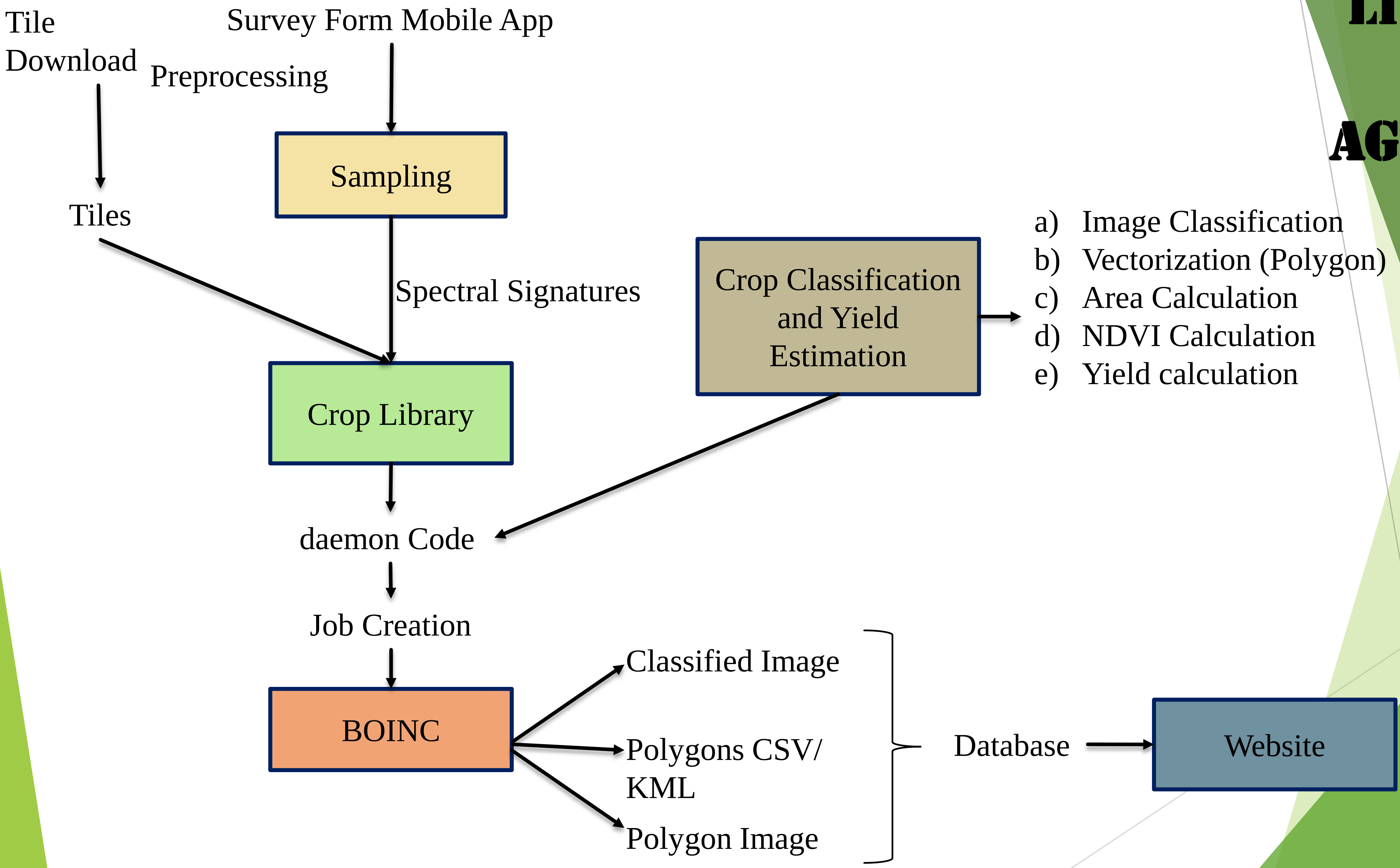


Food@home Project Computing Community Site

Top hosts

Computer info	Rank	Owner	Recent average credit	Total credit	BOINC version	CPU
ID: 18 Details | Tasks Cross-project stats:	1	PITB	39.63	649	7.8.3	GenuineIntel Intel(R) Core(Stepping 3] (8 processors)
ID: 28 Details | Tasks Cross-project stats:	2	Studentr13_2	36.55	444	7.12.1	GenuineIntel Intel(R) Core(Stepping 9] (8 processors)
ID: 4 Details | Tasks Cross-project stats:	3	Stundentr2_1	18.46	457	7.8.3	GenuineIntel Intel(R) Core(Stepping 9] (8 processors)
ID: 2 Details | Tasks Cross-project stats:	4	Studentr1_2	16.78	452	7.12.1	GenuineIntel Intel(R) Core(Stepping 9] (8 processors)
ID: 5 Details | Tasks Cross-project stats:	5	Student	16.57	401	7.8.3	GenuineIntel Intel(R) Core(Stepping 9] (8 processors)

LIFE CYCLE OF AGRI@HOME



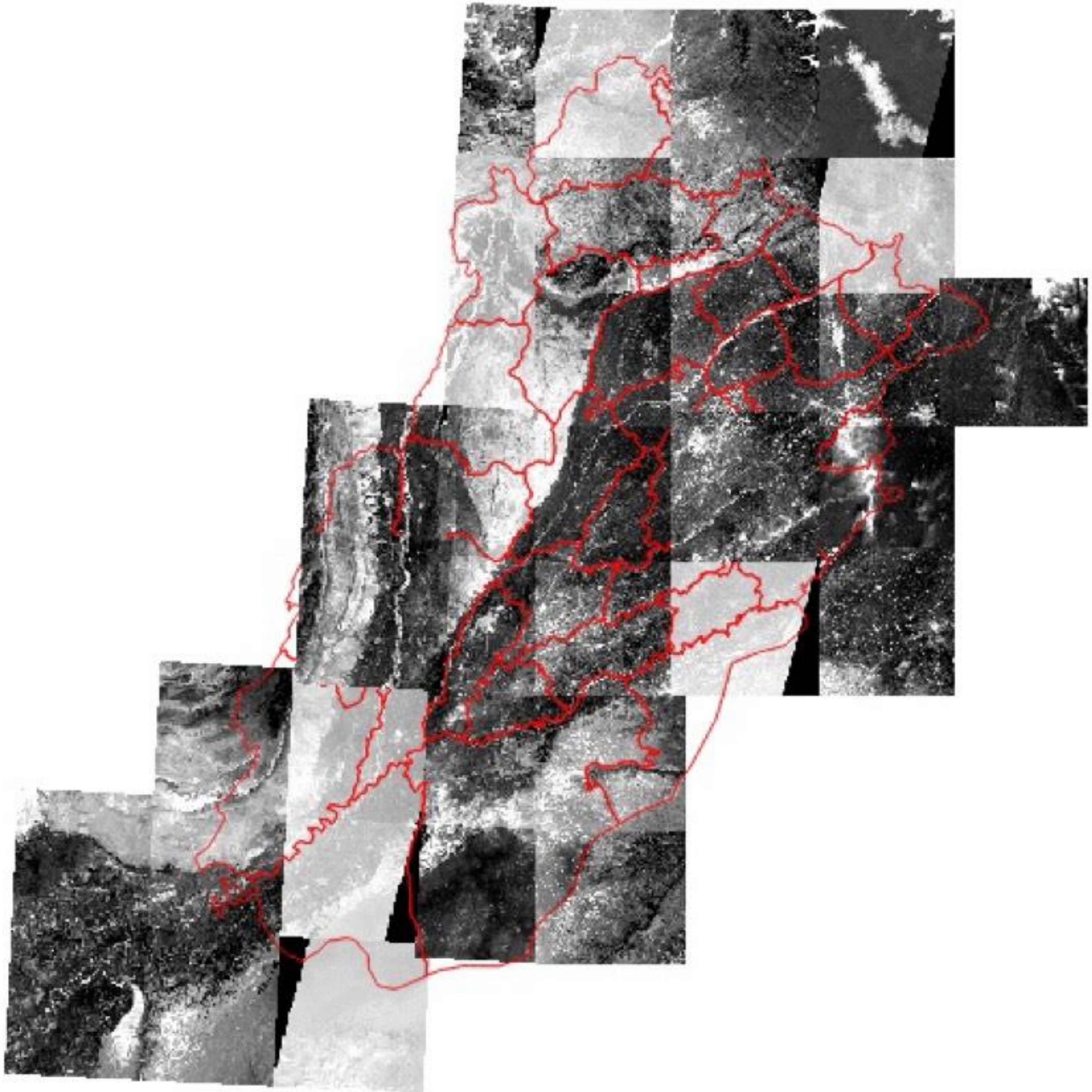
Example: Punjab

Largest Province of Pakistan

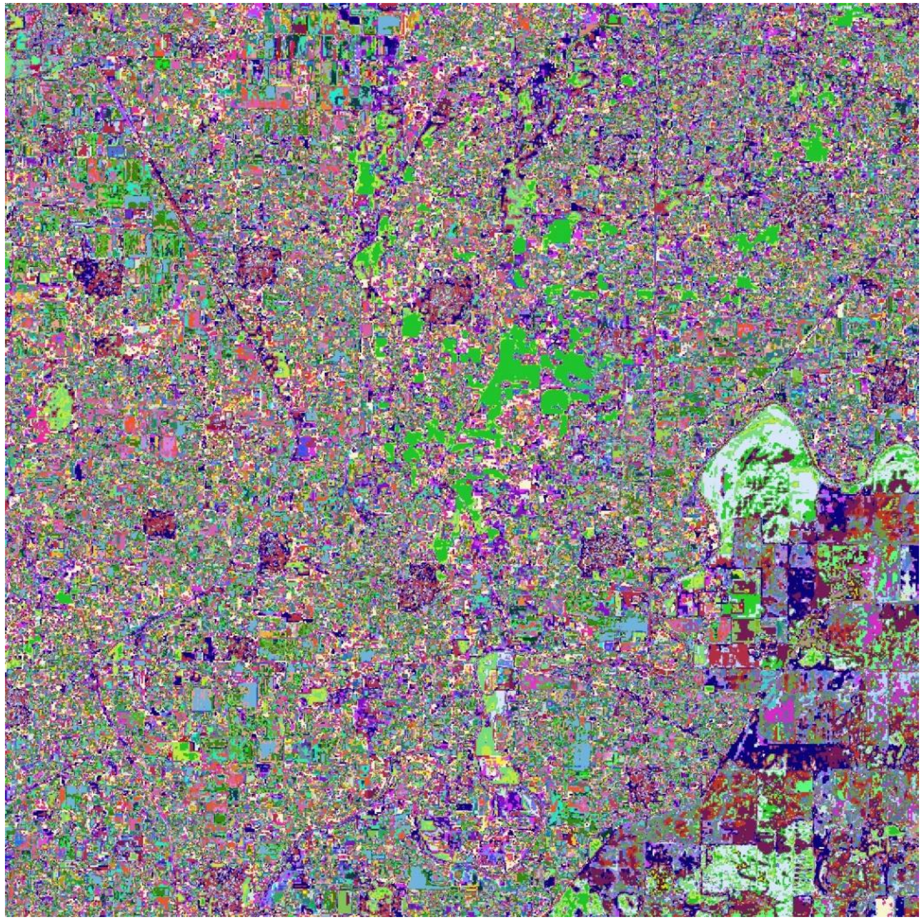
Punjab Wheat Yield Estimation

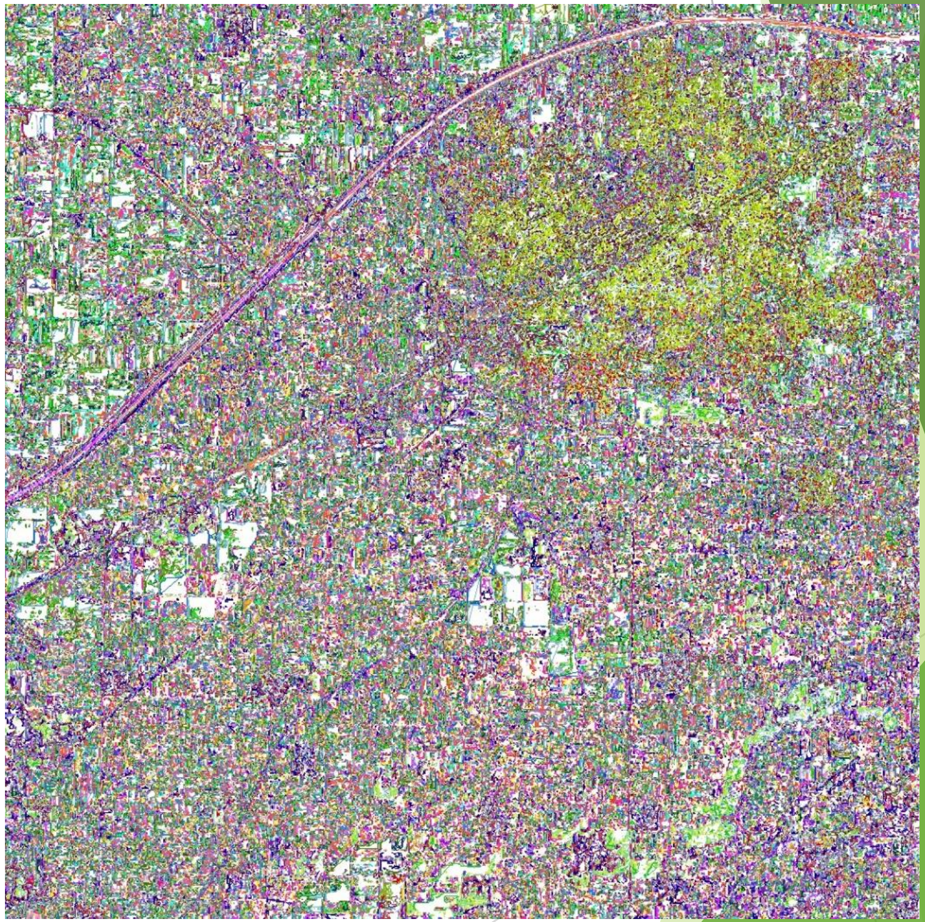
- Downloaded **38** tiles from **7** vertical strips of Sentinel 2A for **Punjab**.
- Unprocessed tile size of 38 tiles is approx. 30GB.
- Processed tile size is approx. 60GB.
- Tiling (dividing each tile to 100 small “cell tiles”) and adding nomenclature for cell tile, converting 38 tiles to 3800 cell tiles.
- A complete tile on average takes 10 - 14 hours, where as our code process cell tile in 4 to 8 minutes(depending on the system).
- Processed Complete Punjab in less than 2 days, using 15-20 machines(available from 6pm – 8am)

Punjab Tiles

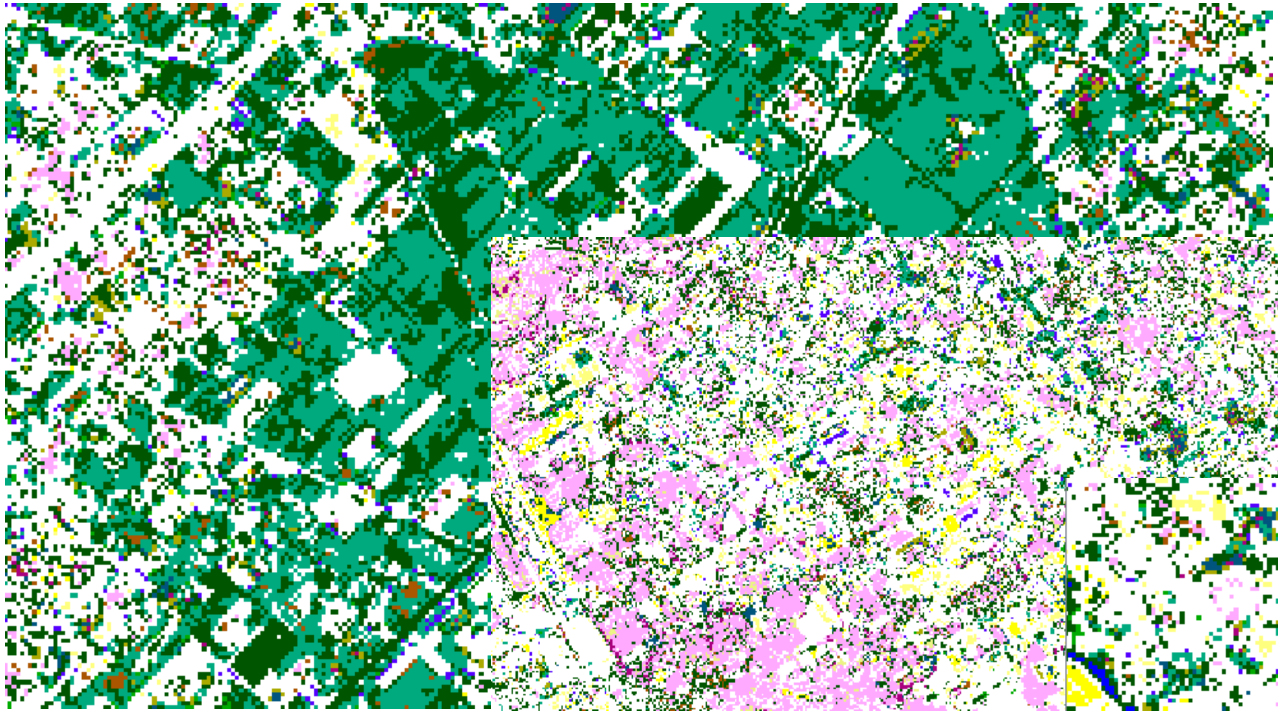


Results

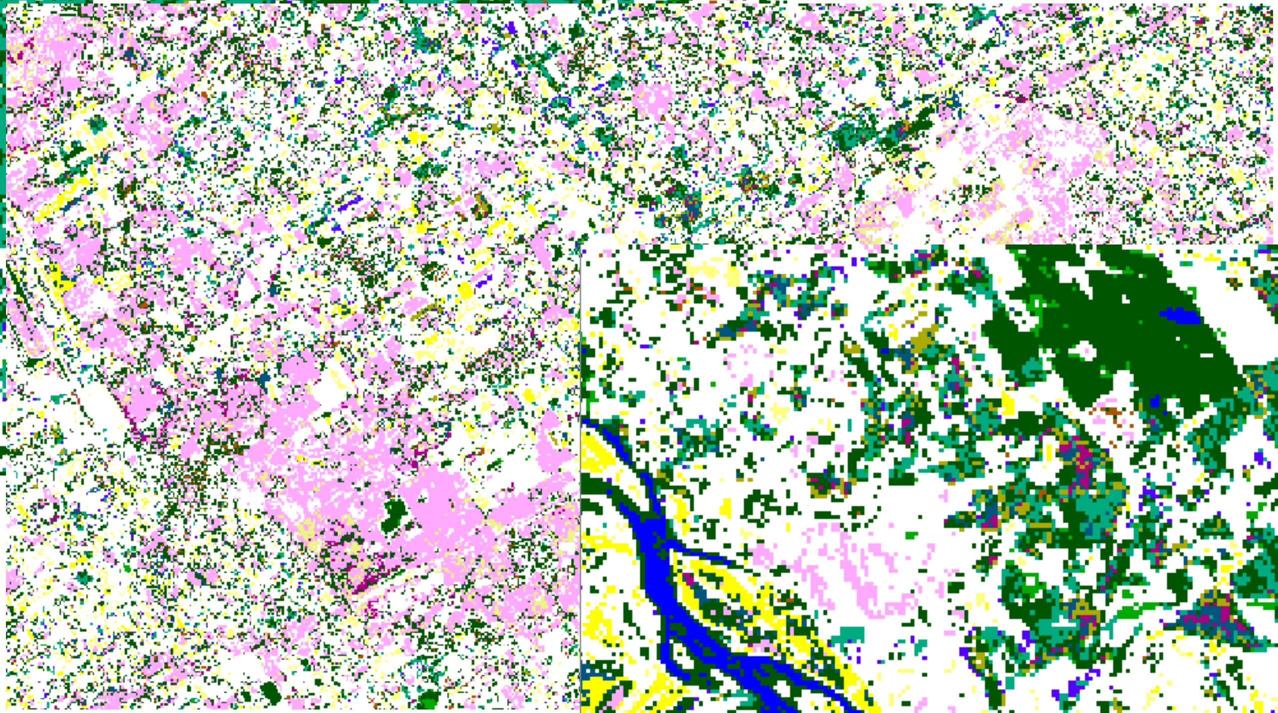




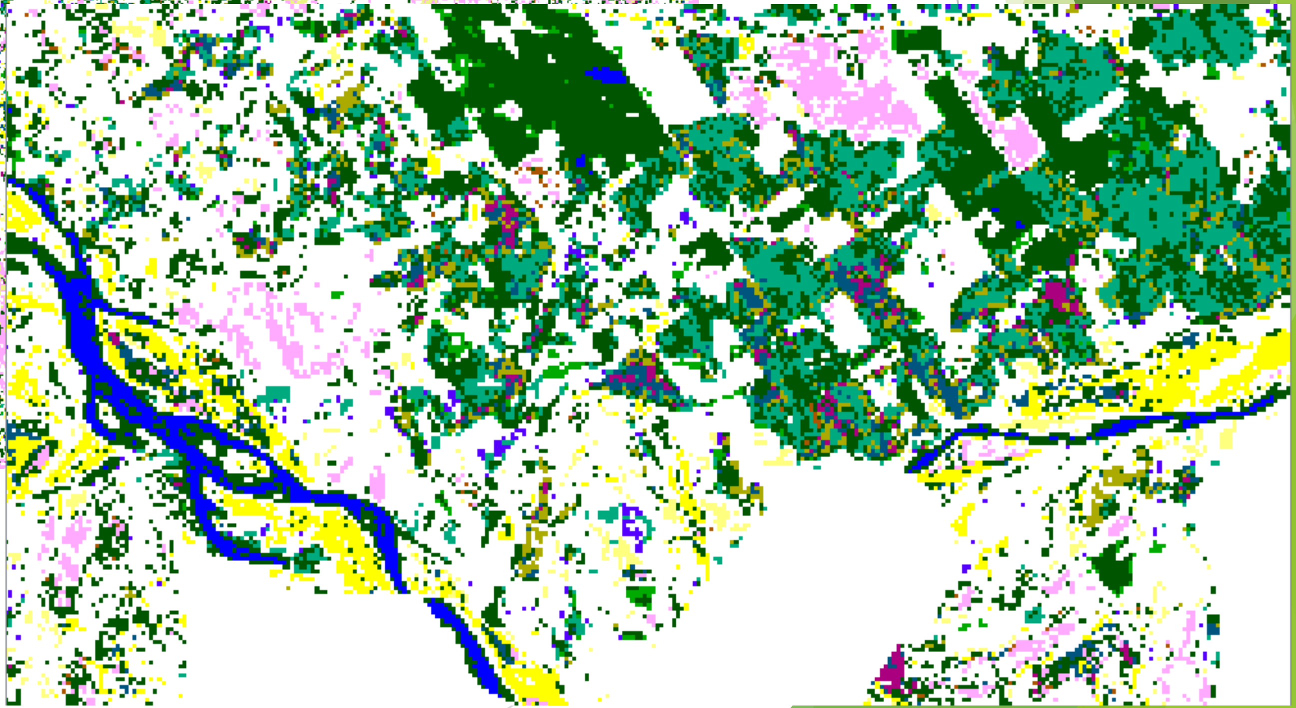
Agro Forest (Orchards - Light Green)



Barren and water logged land (Pink)



Water and soil around water(Blue and yellow)



Issues Resolved

- Instead of Available ML algorithms, we have designed a hybrid algorithm meeting our needs
- Have designed, a data collecting, storing and searching model using discrete global grid tile system.
- Using already available workforce in the agriculture department for collecting samples from field.
- A volunteer computing based structure to assist in making the computation time efficient.

Conclusion

- We are addressing food security by solving 3 out of four pillars of food security by FAO
  - **Availability**
    - Making food available by increasing produce
  - **Access**
    - Food is accessible and cheap if surplus
    - With sufficient availability of food, access to food would be easier
    - More food will reach to the mouths of undernourished and stunted.
  - Utilization
  - **Stability**
    - This methodology will stabilize agriculture based ecosystem
    - Agriculture ecosystem contributes in improving lifestyle of people by providing work opportunities

Conclusion

- Early warning system for food shortage.
- Improve crop yield predictions on the go with frequent availability of satellite data.
- Help in improving agriculture based economies
- Policy decisions around import and export of food could be assisted

Thank you