LSGI552 Project

Machine Learning for Tree Monitoring by Google Street View and Shown in 3D map

Michael LI

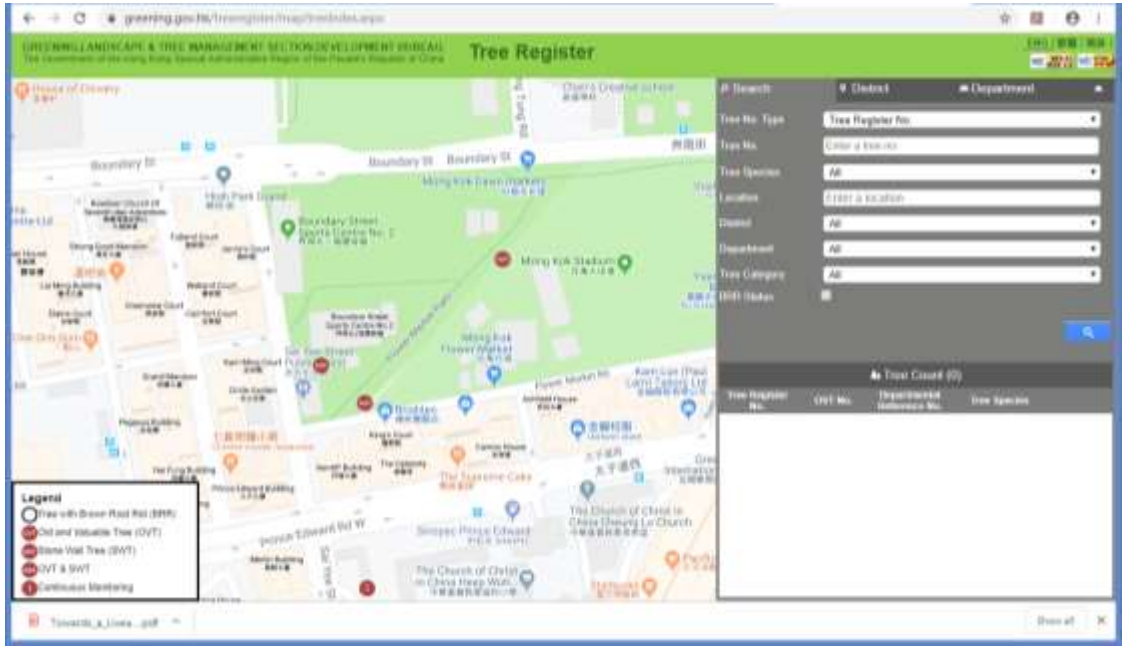
Supervisor

Dr.-Ing YAO Wei

Problem

- In Hong Kong, many trees are geographically distributing
- Trees grow in variety of occasion such in roads, slopes, building gap, grass area, etc
- Arial images provide overall coverage but hard to distinguish individual tree
- Difficult and time consuming to count and monitor the distribution of trees by manual work
- Fatal accidents sometime occur due to unmanaged tree

Current Hong Kong Tree Register



https://www.greening.gov.hk/treeregister/map/treeIndex.aspx

To promote community involvement in tree risk management to better protect public safety, the Tree Register was set up since July 2010 to provide information to the public. In September 2015, the Tree Register presents the latest information of trees on a geo-referenced map with search functions and statistical summary.

Situation

- Current Tree Count: 865
- Manually Input tree data and status
- Inefficiency to monitor and develop the tree database
- 2016,698,523 trees from 554 different species in Hong Kong as recorded in October 2018
- 2D Map only

Objective

- To develop an application by using machine learning (YOLOv3) to automatically extract the tree objects from Google street images and to calculate the tree location and distribution
- The tree distribution will be shown in 3D spatial and visualization platform for Hong Kong tree management

Related Work (1)

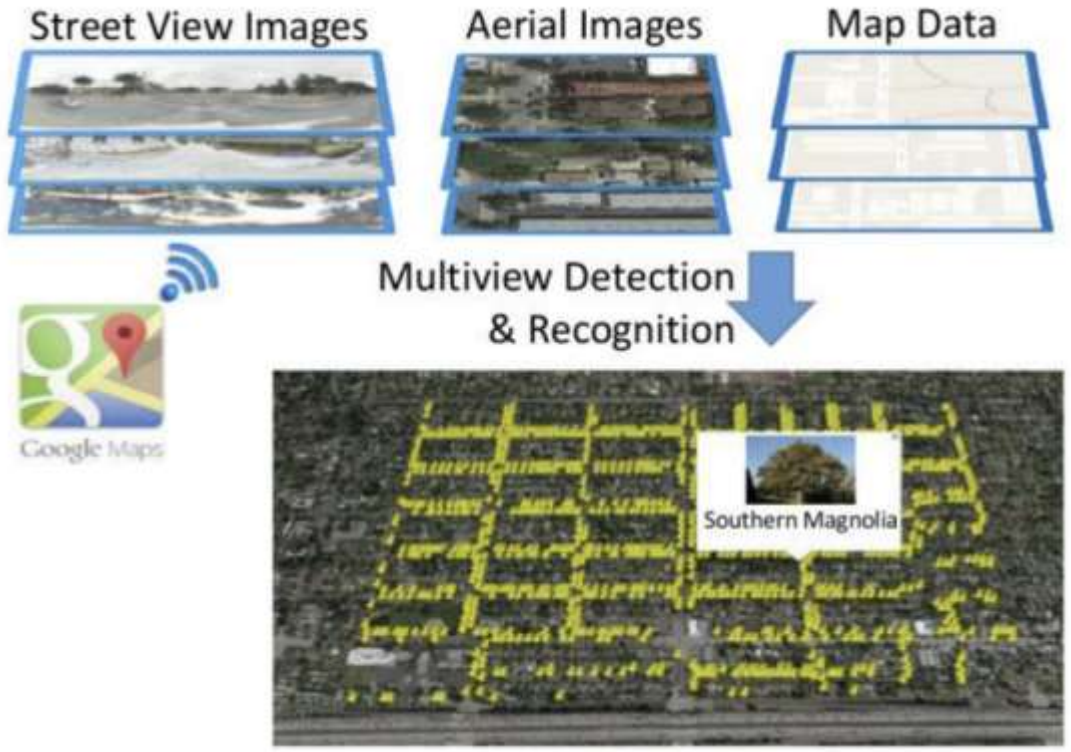
- Satellite images, aerial images, hyperspectral data, and point cloud from Light detection and ranging (LiDAR) have been commonly utilized for tree detection, tree health assessment and species recognition, etc.

- Google Street view imagery also provide sufficient amount of information for tree detection and classification.

Related Work (2)

RegisTree project - is a collaboration between ETH Zürich and the California Institute of Technology (Caltech) investigating the detection and classification of public objects in publicly available imagery from Google Maps

- Input: Street view images, Aerial images and Map data
- Target: Automatic mapping of stationary recurring objects from Street View.
- State-of-the-art: Faster-RCNN, Object recognition, Image geolocation, Object geolocation.

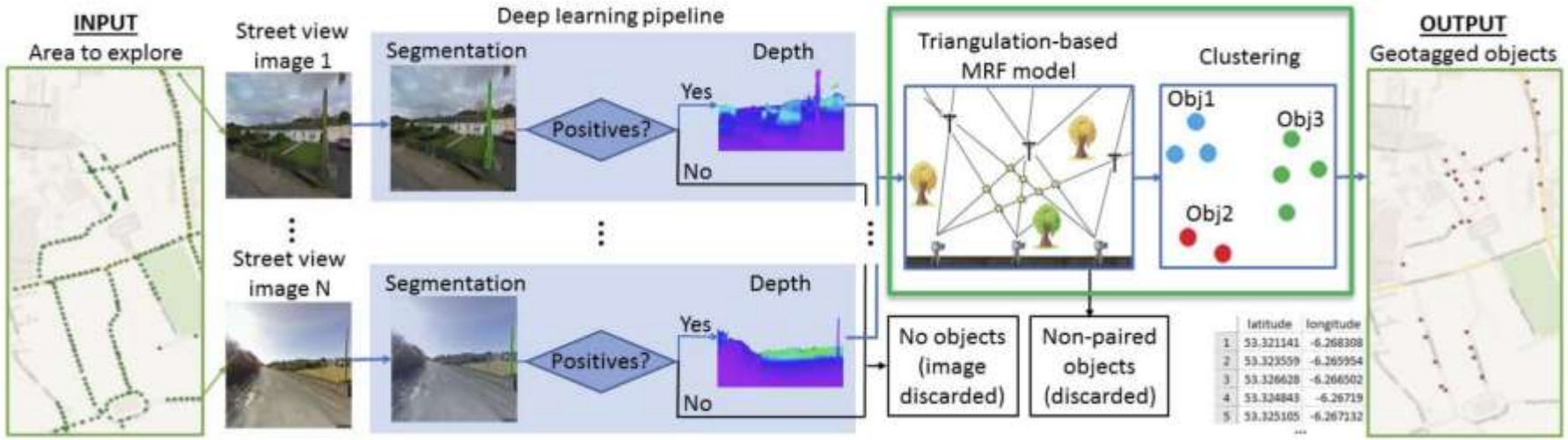


Wegner, J. et al., CVPR 2016

Related Work (3)

Krylov et al., (2018) proposed a special model by using Markov Random Field (MRF) to perform objects triangulation and geotagging:

- Input. Google street view imagery only
- Target. Automatic mapping of traffic lights and telegraph poles from Street View.
- State-of-the-art: Deep Learning, Object recognition, Triangulation, Spatial clustering, Object geolocation
- Accuracy: 2 meters



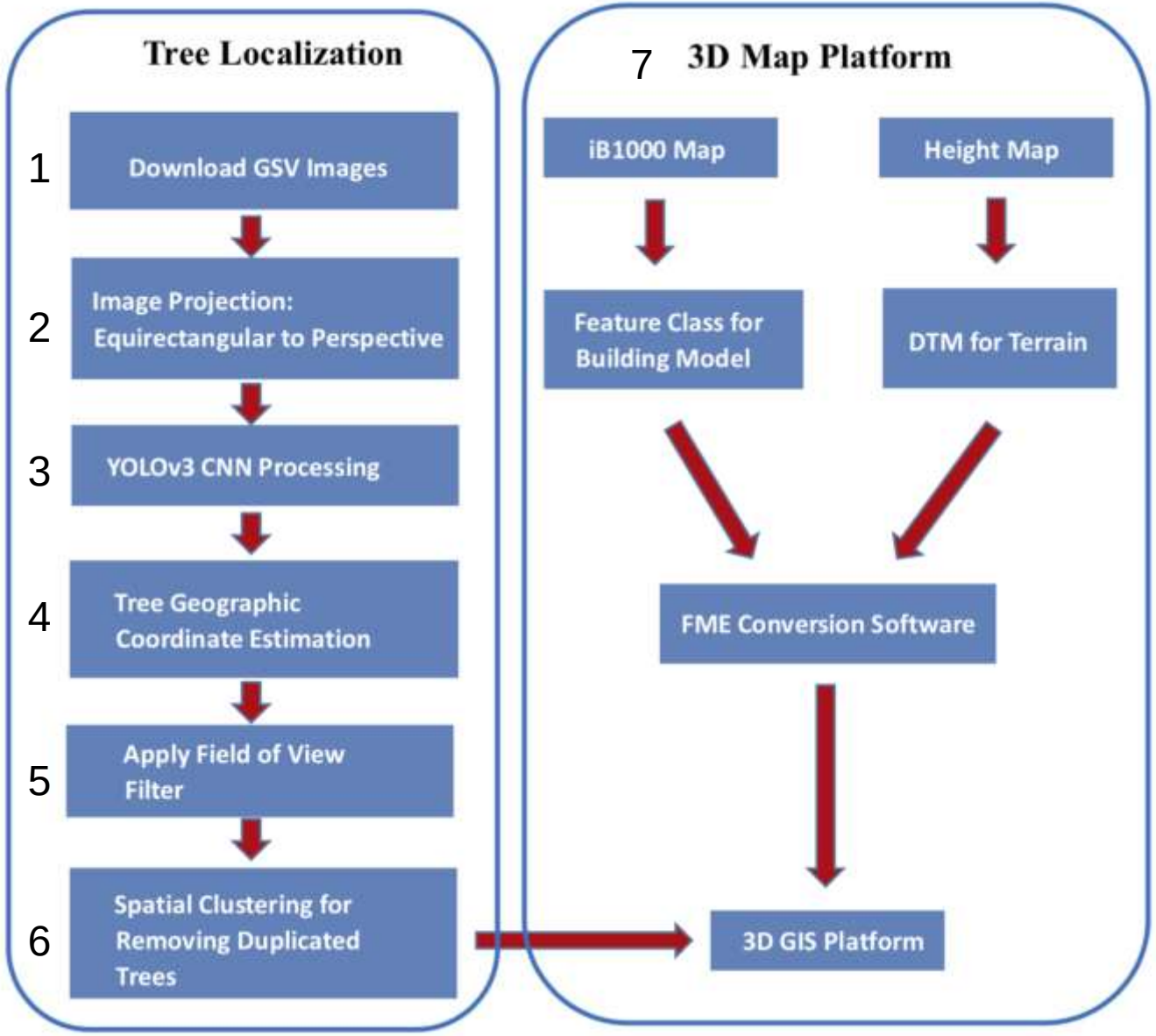
Krylov et al., (2018)

Methodology

I proposed a new method to perform objects detection and geotagging:

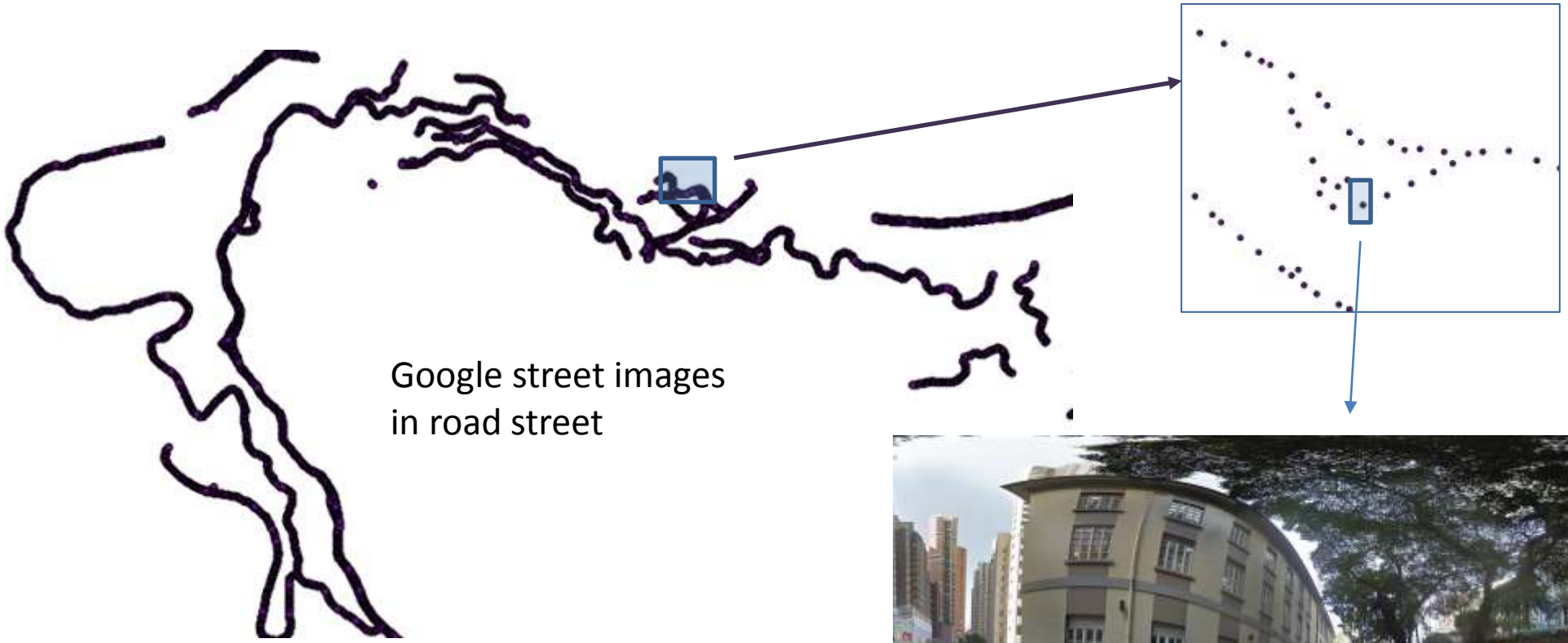
- Input. Google Street view imagery only (RGB and Depth images)
- Target. Automatic mapping of tree objects from Street View imagery and visualize in 3D Map
- State-of-the-art: CNNs, Object detection, Spatial clustering, “field of view” filter, Object geolocation
- Accuracy: 3 meters for 2 selected roads

Workflow for development of automated tree detection in 3D Map platform



1. Download GSV Images (1)

Tree Localization



Google street images
in road street

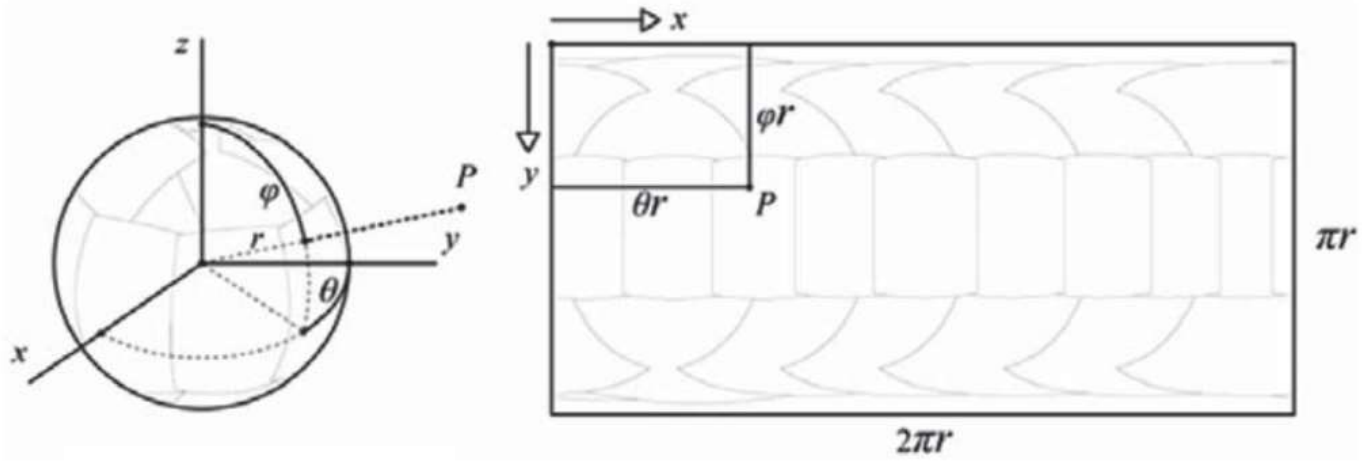
26 Roads in Hong Kong Island are selected.

equirectangular image

1. Download GSV Images (2)

Tree Localization

Spherical cameras used by Google for spherical panorama acquisition



$$x = R(\lambda - \lambda_0)\cos\varphi_1$$
$$y = R(\varphi - \varphi_1)$$

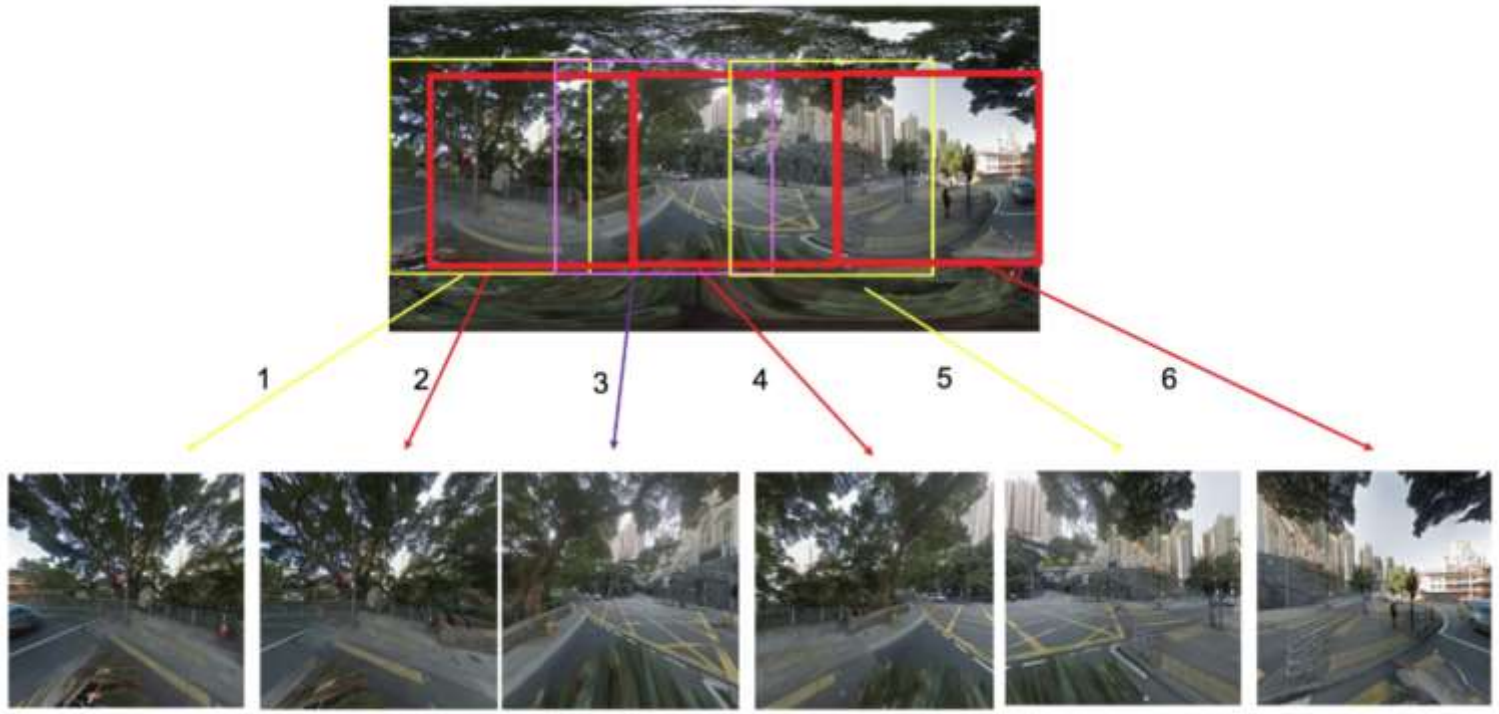
https://en.wikipedia.org/wiki/Equirectangular_projection

The forward projection transforms spherical coordinates into planar coordinates. Relationship between spherical coordinates and equirectangular coordinates.

2. Image Projection

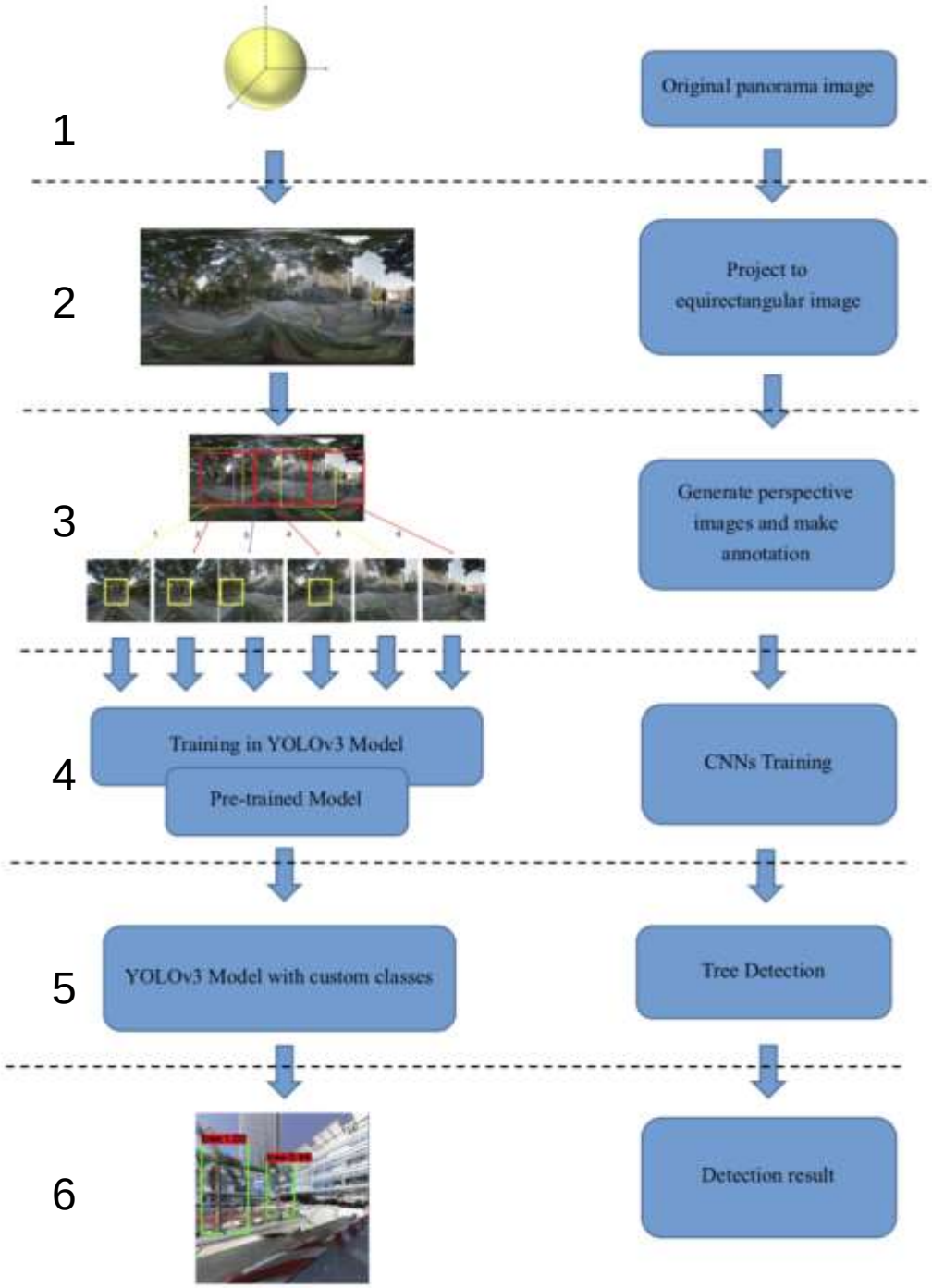
Tree Localization

Equirectangular to perspective projection at different field of view (FOV). Yellow box is an area to carry out the projection. (Bottom)



The equirectangular image is exported into 6 perspective images (FOV 114°) for CNNs training

3. CNN Processing



Tree Localization

4. Tree Geographic Coordinates Estimation

$If\ x <= W/2, x' = W/2 - x$

$If\ x > W/2, x' = W - (x - W/2)$

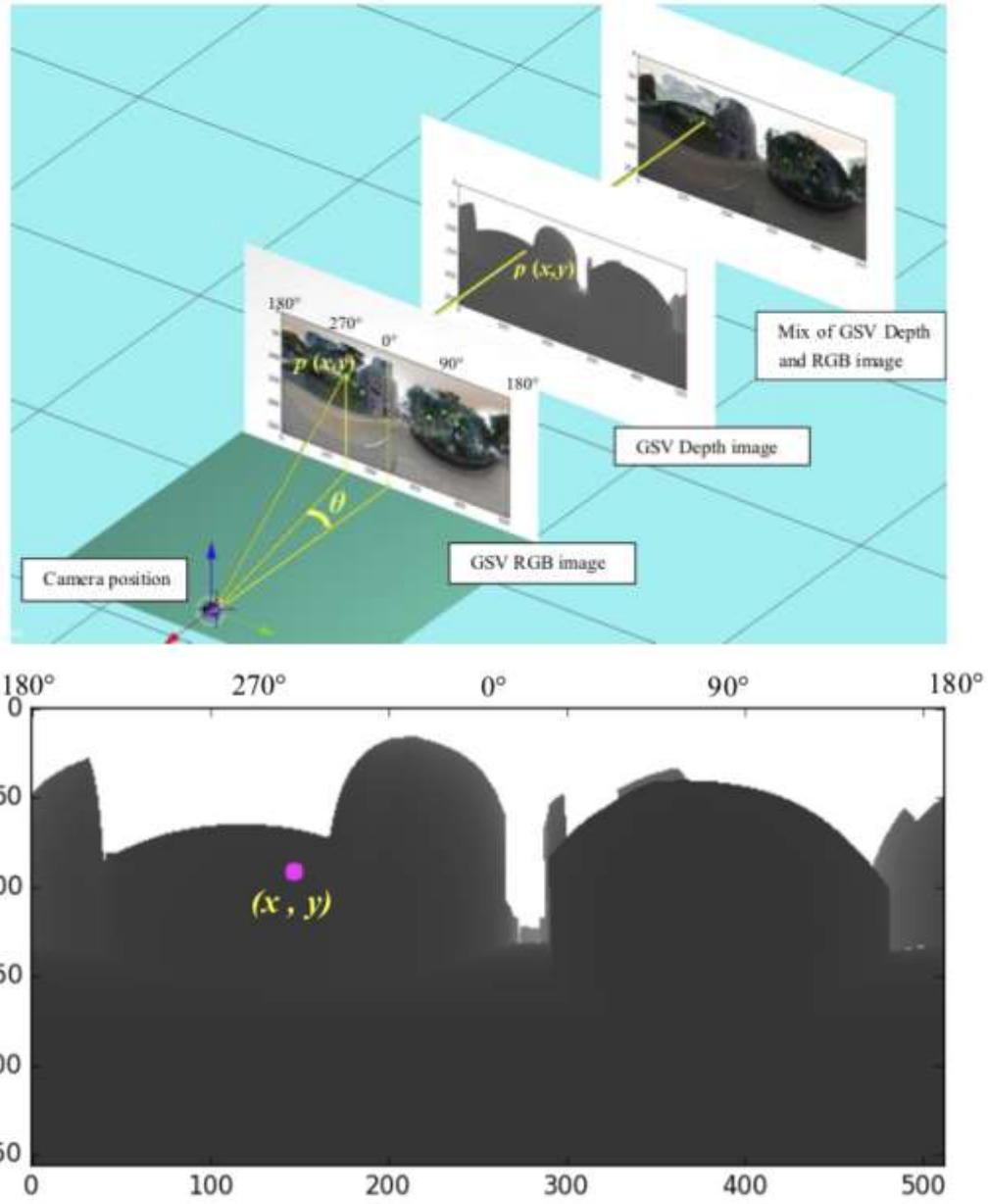
$\theta = 360/W * x'$

$\theta' = \theta^0 + \theta$

P' (lat, lon) = GetDistance(P (lat, lon), D, θ')

Where

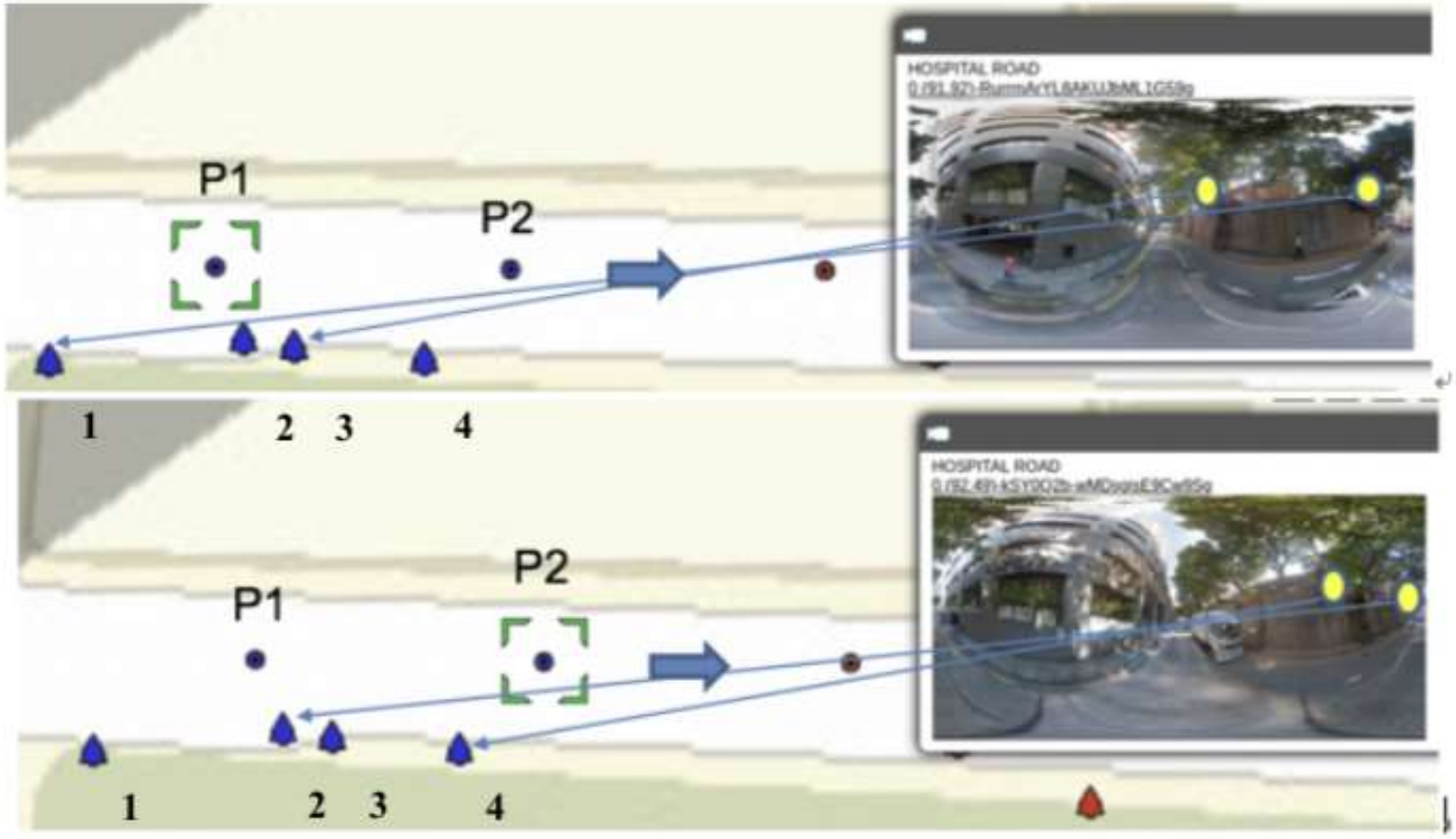
H - the height of the depth image (256px)
W - the width of the depth image (512px)
x' - the remapped x coordinate in depth map, the center is defined as zero
p (x,y) - the coordinate of the detected point in depth map image space
P (lat, lon) - geographic coordinate of the camera center
P' (lat, lon) - geographic coordinate of the detected tree
D - the distance from the camera center to the detected point
θ - the yaw angle in depth image space
θ^0 - the default GSV tile yaw angle
θ' - the final yaw angle
GetDistance - the function for calculating geographic coordinate *P' (lat, lon)* of the detected tree from camera center point *P (lat0, lon0)* having yaw angle *θ'* and distance *D*



(Top) Fusion of GSV Depth and RGB color images. The detected beam passes through a same pixel for GSV RGB and Depth images. (Bottom) Depth image

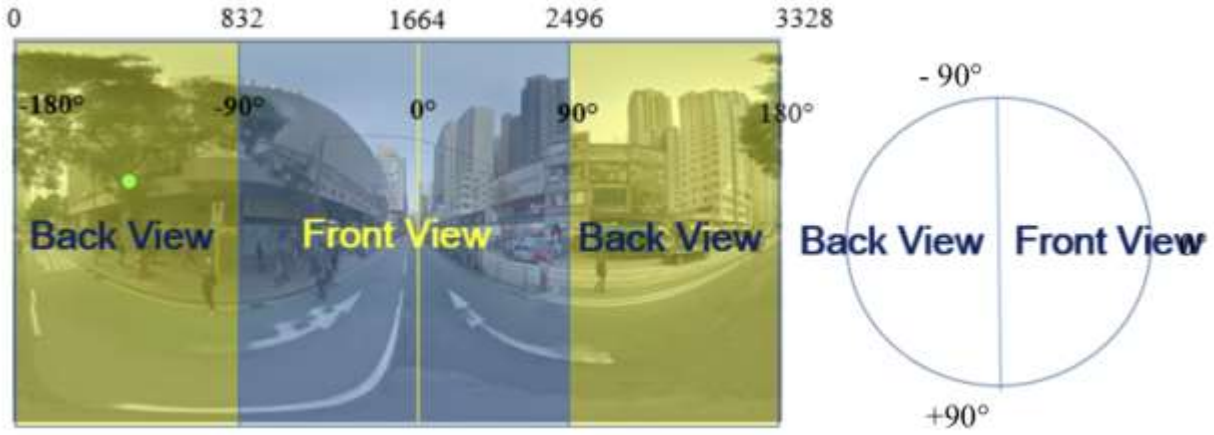
5. Field of View Filter (1)

Tree Localization

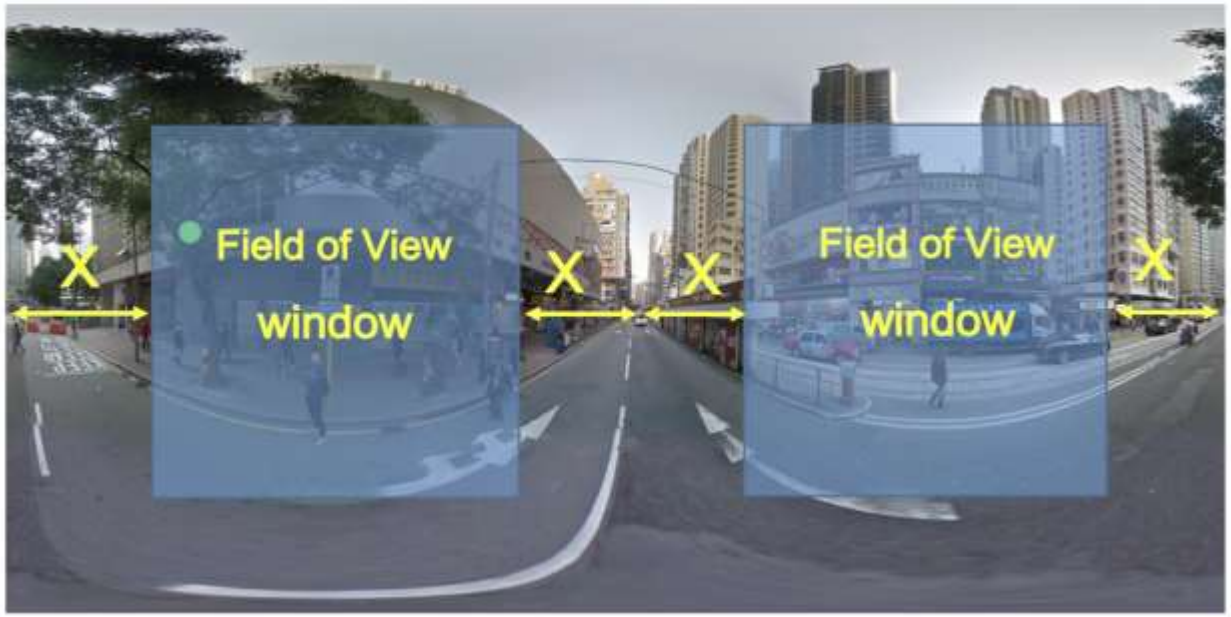


Duplicate tree detected for two street view images P1 and P2 are the panorama points. The arrow shows the direction of driving.

5. Field of View Filter (2)



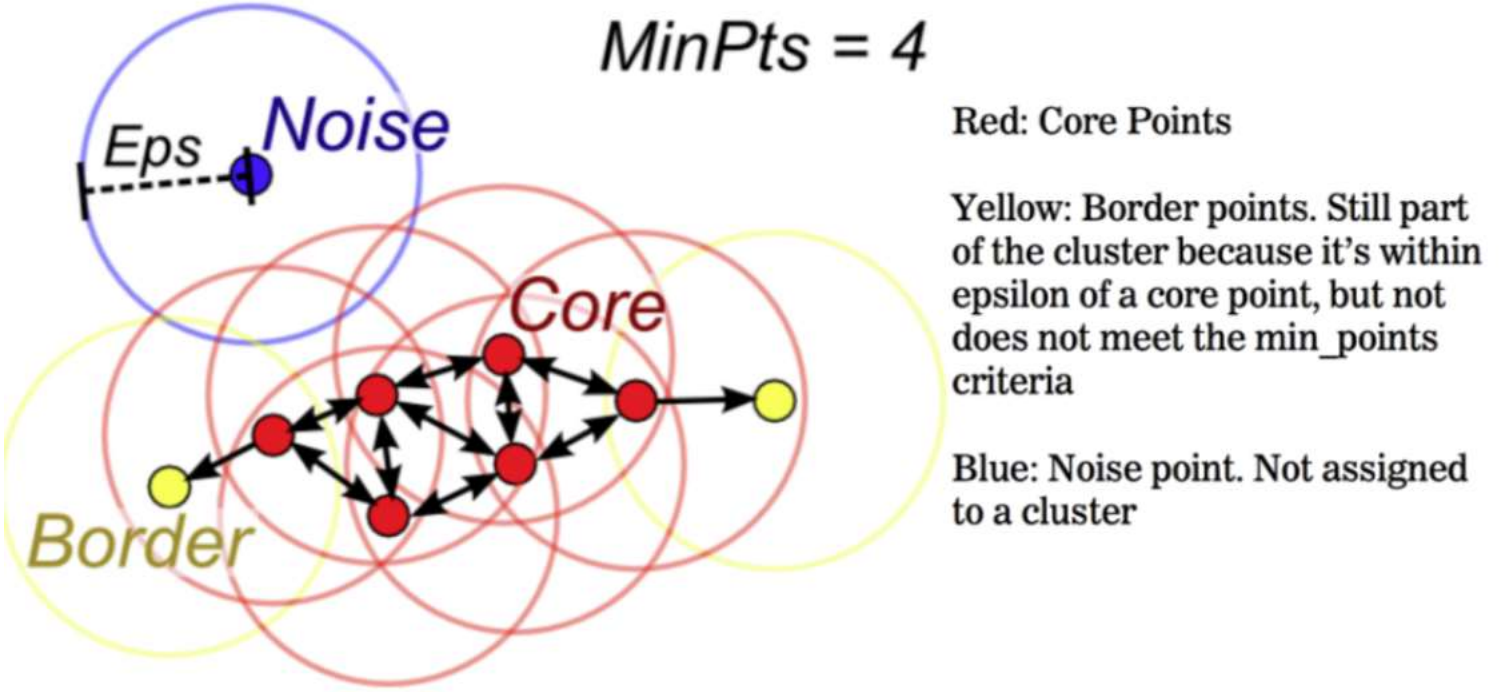
Front view and back view in equirectangular image (left) and panorama sphere (right)



Window filter with offset setting (X) is applied to reduce duplicated tree detection.

6. Spatial Clustering for Removing Duplicated Trees

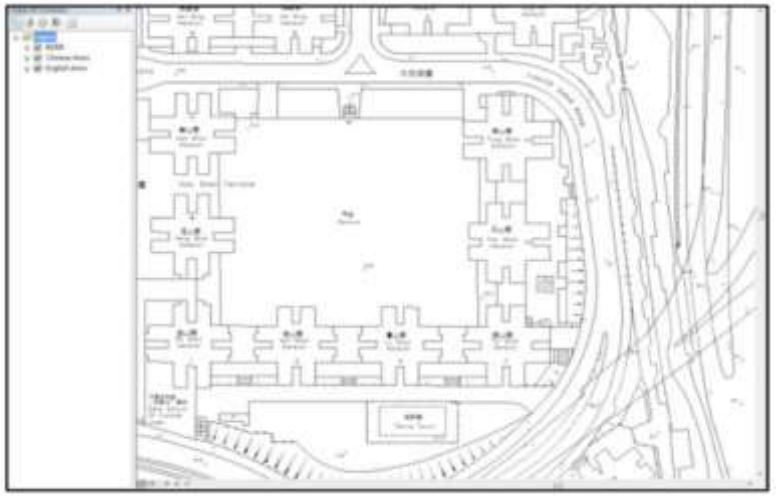
Tree Localization



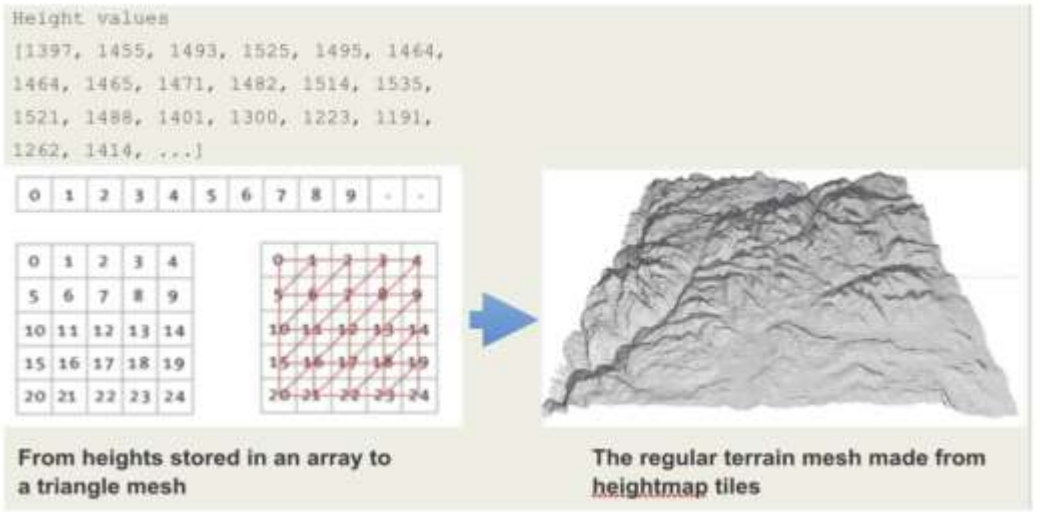
Demonstrating density-based spatial clustering with noise (DBSCAN) algorithm for spatial clustering (Wikipedia)

7. 3D Map – Data Source

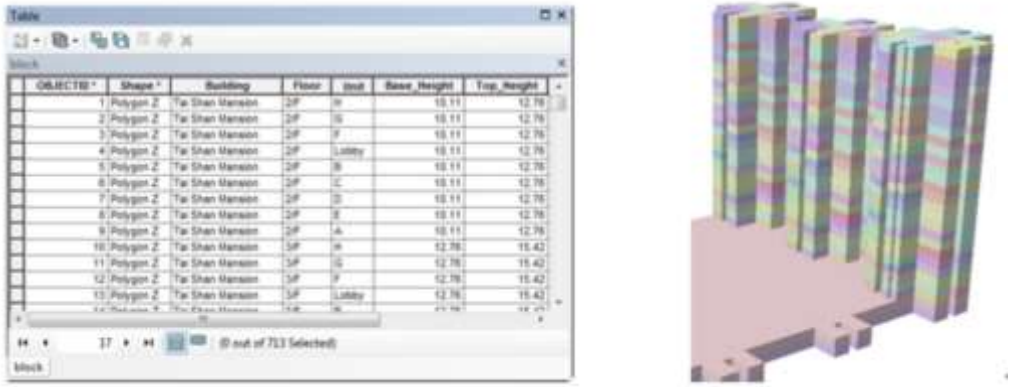
3D Map Platform



FGDB format iB1000 (Lands Department. (2017))



Digital Terrain Model



Building records of attribute (left) and extrusion results of selected buildings (right)

3D building modelling

Experiment

Experiment

1. 26 Roads in Hong Kong Island are selected for geo-tagging
2. 2 selected roads (Hospital Road and Johnston Road) for ground truth checking
3. Total street view images: 6618

Experiment

26 Roads in Hong Kong Island are selected

Road Name: Gloucester Road, Hennessy Road, Man Cheung Street, Lockhart Road, Johnston Road, Lung Wo Road, Upper Albert Road, Lower Albert Road, Garden Road, Kennedy Road, Macdonnell Road, Hornsey Road, Robinson Road, Conduit Road, Lyttelton Road, Lee Nam Road, Shek Pai Wan Road, Cyberport Road, Victoria Road, Pok Fu Lam Road, Shing Sai Road, Bonham Road, Hospital Road, Po Shan Road, Stubbs Road, Mount Butler Drive

Total street view images: 6618

Training with YOLOv3

Experiment

Training Dataset: Depending on the amount of data, 1500 and 2500 images are randomly selected for two sets of training. We have to input example of desired trees which only planted in local environment. Annotation with tree bounding box on training data has to be completed before. I used an opensource labelling tool to complete the task.

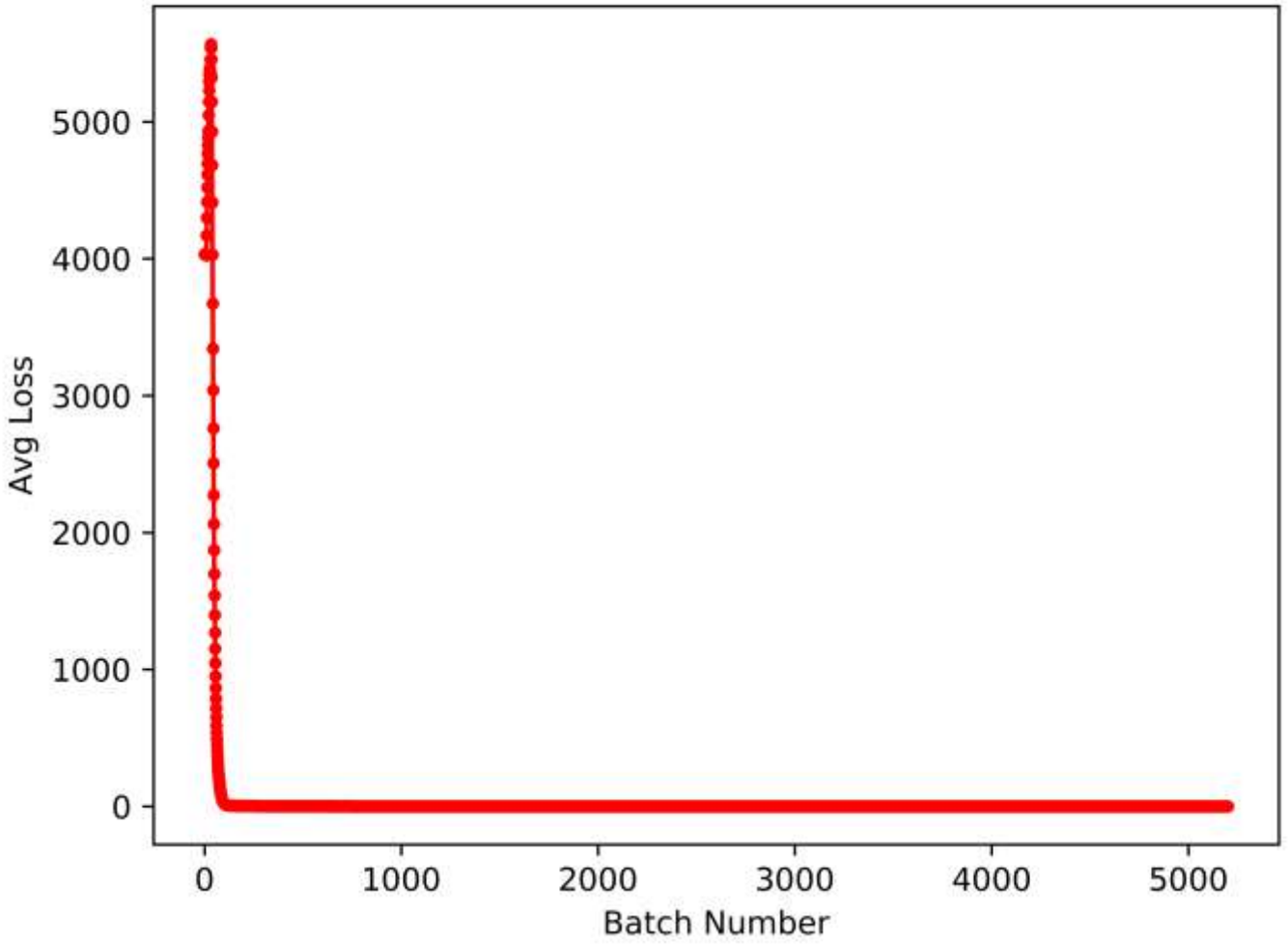
Test Dataset: This is the data set on which the trained model is tested. No image should be part of the both the training and the test set. 200 and 300 images are selected for two sets of testing.

```
1          2              3                    4                    5                    6
1: 8790.451172, 8790.451172 avg, 0.000000 rate, 6.481886 seconds, 64 images
2: 8790.721680, 8790.478516 avg, 0.000000 rate, 6.413727 seconds, 128 images
3: 8809.923828, 8792.422852 avg, 0.000000 rate, 6.710170 seconds, 192 images
4: 8788.920898, 8792.072266 avg, 0.000000 rate, 6.519844 seconds, 256 images
5: 8820.444336, 8794.909180 avg, 0.000000 rate, 6.427377 seconds, 320 images
```

Training log output from YOLOv3 network

Training with YOLOv3

Experiment

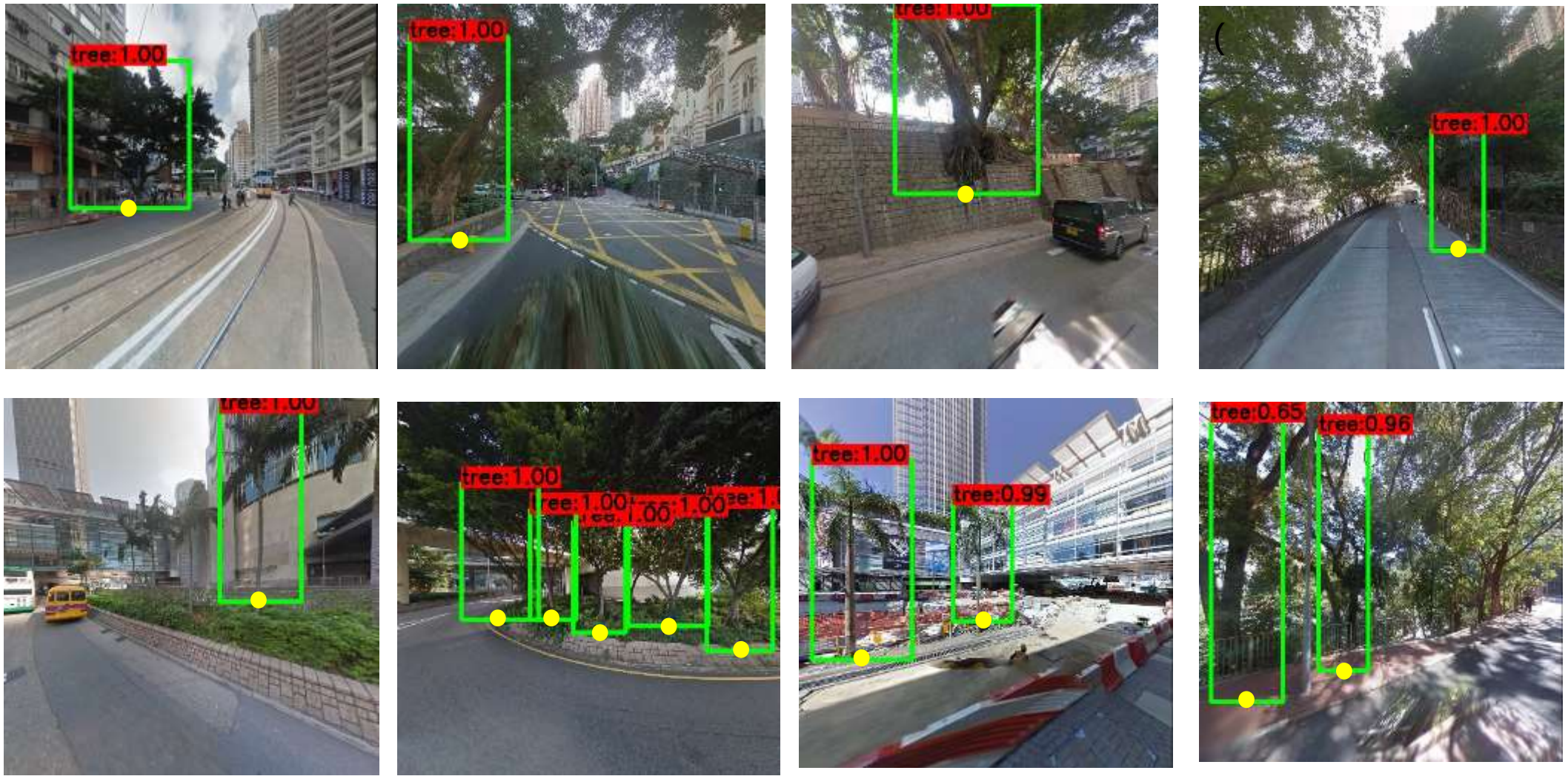


The training plots. Average loss vs batch number

Images for training: 1500
Images for testing: 200
Training Time: 12 hours

Results (Perspective Image)

Object Detection



Tree detection with bounding box. Yellow dots represent the position of the detected trees in perspective image

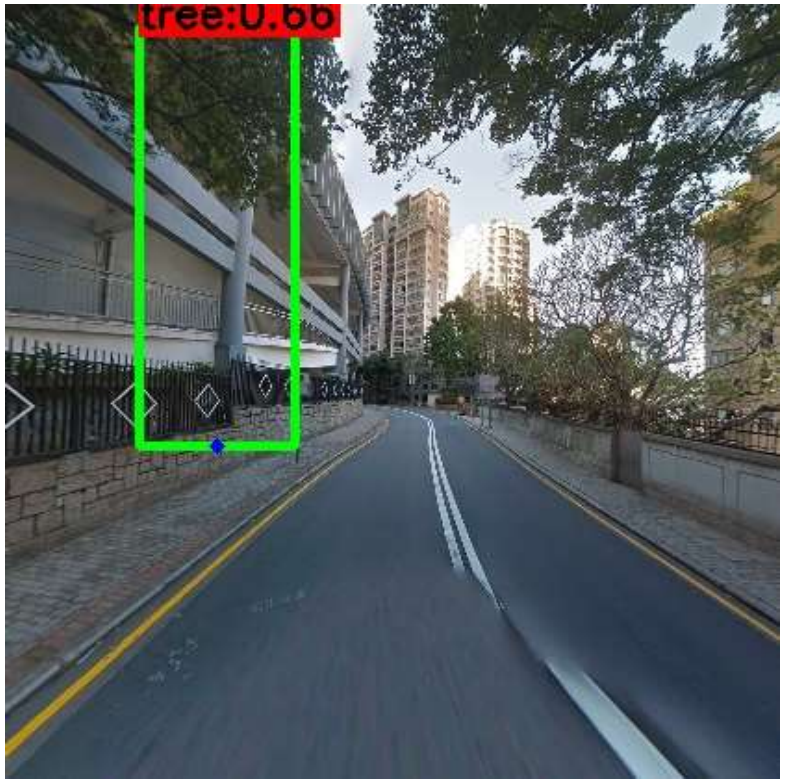
Results (Equirectangular Image)

Object Detection

Green dot represents position of the tree in equirectangular image

Results (wrong detection)

Object Detection



Common mistake is found for lamp posts, water diversion channel and building gap

Object Detection

Results (miss-detection)

Miss-detection. The mistake is highlighted in yellow circle

CNNs model cannot distinguish trees with variety of appearance

Accuracy in Tree Object Detection

Object Detection

	(a) Train image: 1500 Test image: 200	(b) Train images: 2500 Test images: 300
Number of roads	26	26
Number of 360 streetview images	6618	6618
Number of correct detections	10076	11026
Number of wrong detections	1035	1005
Number of fail detection	1801	1505
Accuracy rate	0.78	0.81

Summary of YOLOv3 network training and detection with different training set

Accuracy in Tree Geographic Position (1)

Tree Geographic Position

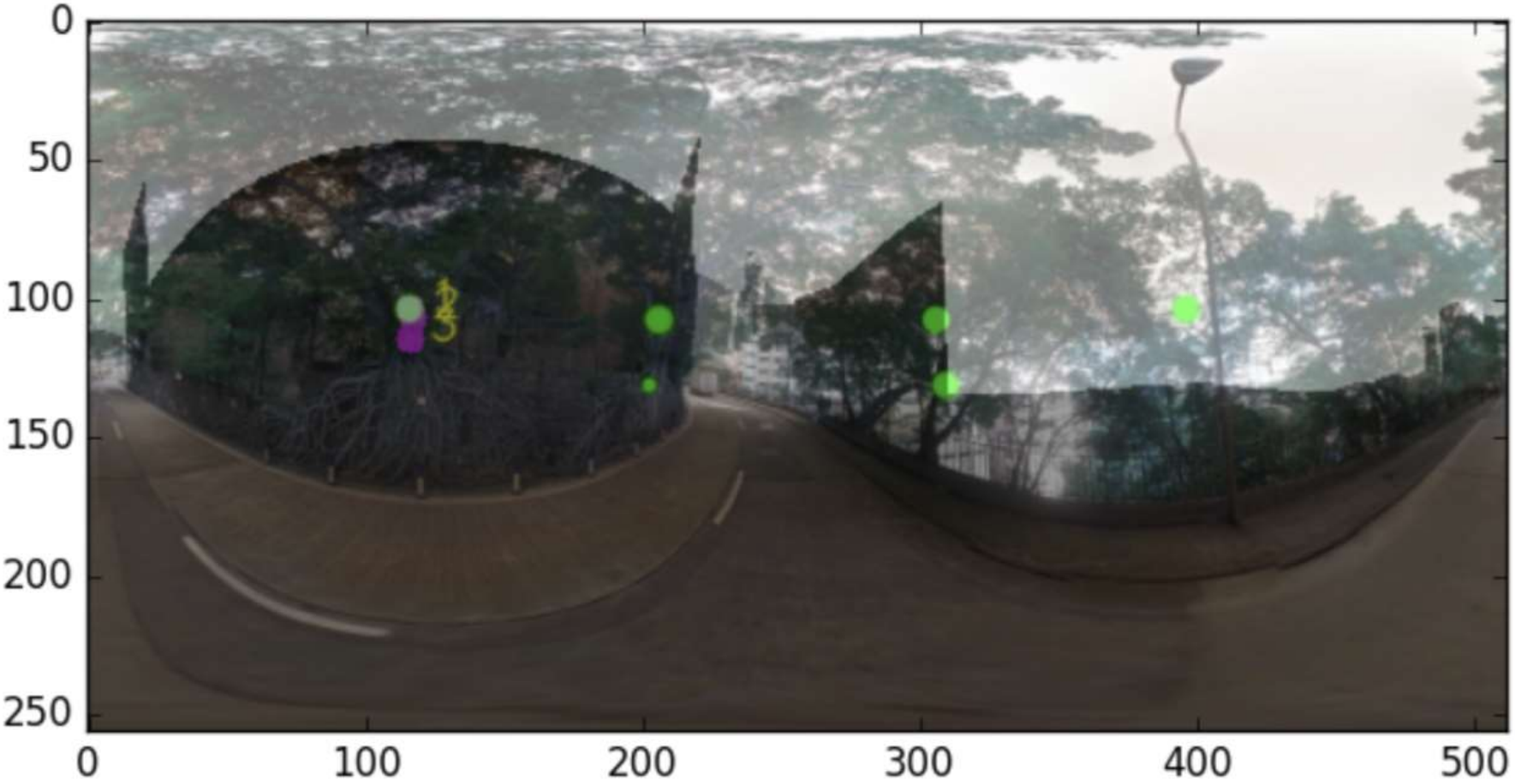
To compare the difference in tree's geographic position between estimation and real scene, two roads have been selected for onsite checking to determine the ground truth data: they are *Hospital Road* and *Johnston Road* having following properties.

1. Hospital Road: this road is extended along the hilly and mountainous terrain with steep slopes

2. Johnston Road: this road is extended along relatively flat terrain.

Accuracy in Tree Geographic Position (2)

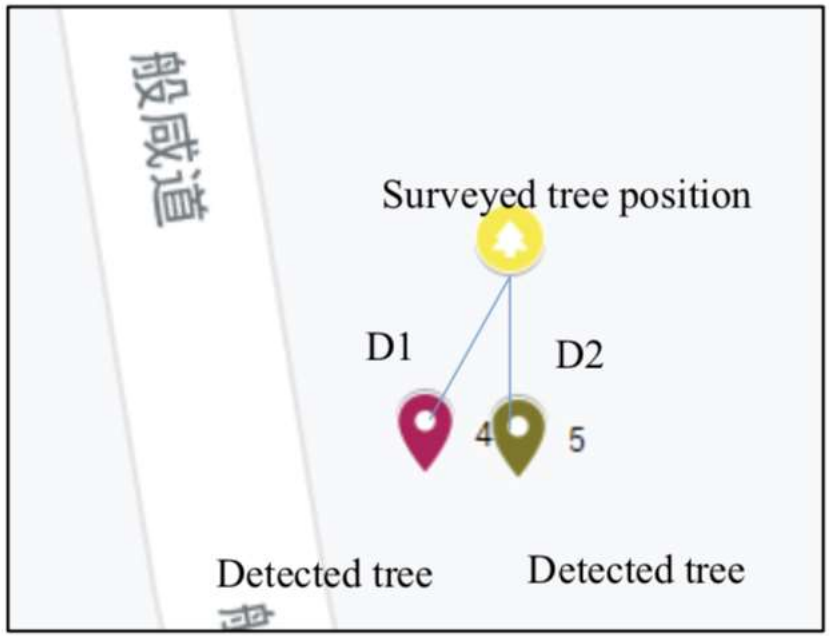
Tree Geographic Position



Result for a processed equirectangular image with detected points. This image is generated by fusing RGB with Depth images. The detected points (green dots) are marked inside the image.

Accuracy in Tree Geographic Position (3)

Tree Geographic Position



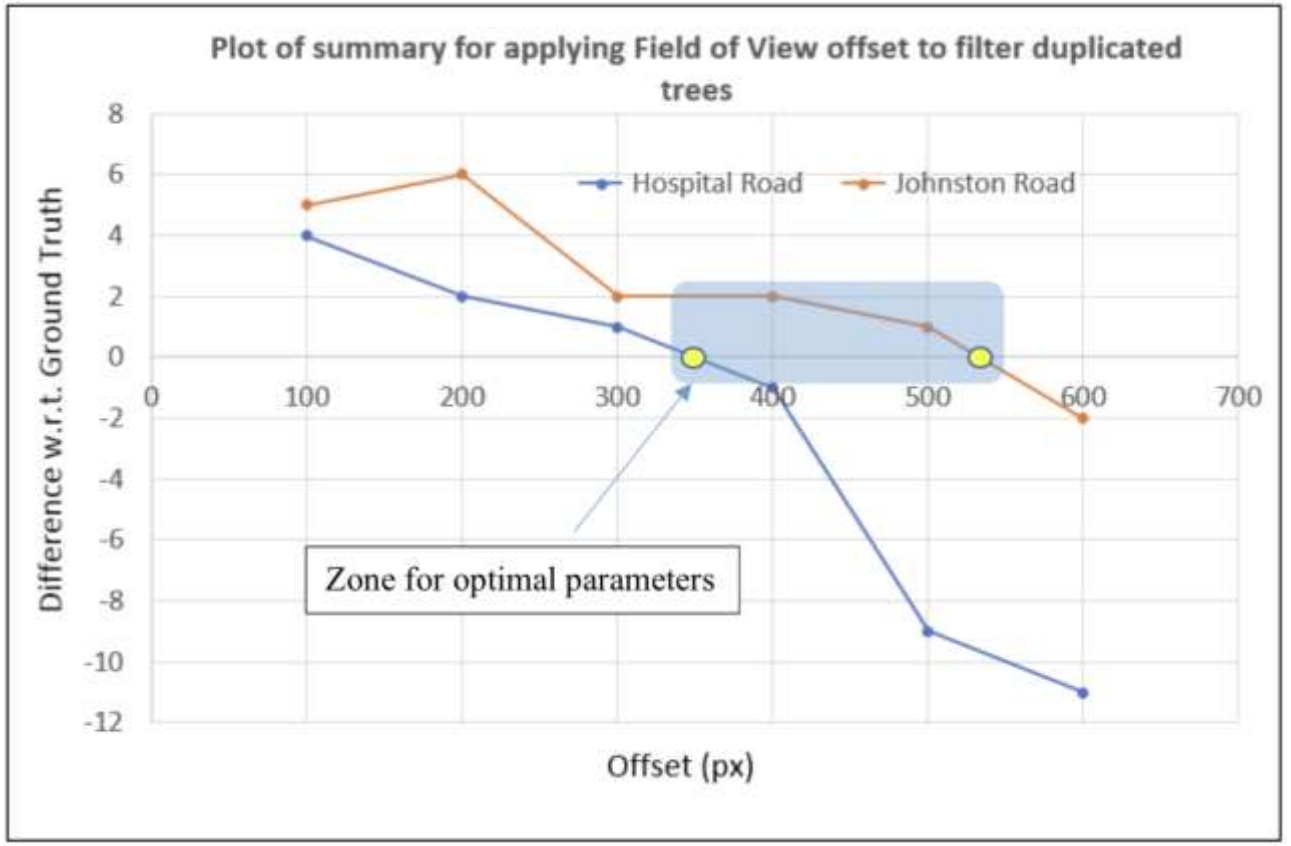
Calculation of mean planimetric deviation w.r.t. ground truth data. Detected trees labeled 4 and 5 belong to same surveyed tree

Accuracy in Tree Geographic Position (4)

Tree Geographic Position

	Hospital Road		Johnston Road	
Number of trees recorded in onsite survey (Ground Truth)	28		11	
Average distance (m) of clustered trees w.r.t. ground truth data with offset X (field of view)				
	Difference	Average distance (m)	Difference	Average distance (m)
X = 100px	4	2.1	5	2.9
X = 200px	2	2.1	6	2.9
X = 300px	1	2.2	2	2.8
X = 400px	-1	2.4	2	2.9
X = 500px	-9	2.2	1	2.5
X = 600px	-11	2.3	-2	2.6

Summary of results for number of clustered trees in different offset value X for the “field of view” filter.



Plot of summary for the result

Accuracy in Tree Geographic Position (5)

Tree Geographic Position

Result for Hospital Road in map view. Offset value for field of view is 400px. Spatial cluster (epsilon = 4m) has applied to remove duplicated trees. Average distance is 2.4 meters.

Result for Johnston Road in map view. Offset value for field of view is 400px. Spatial cluster (epsilon = 4m) has applied to remove duplicated trees. Average distance is 2.9 meters

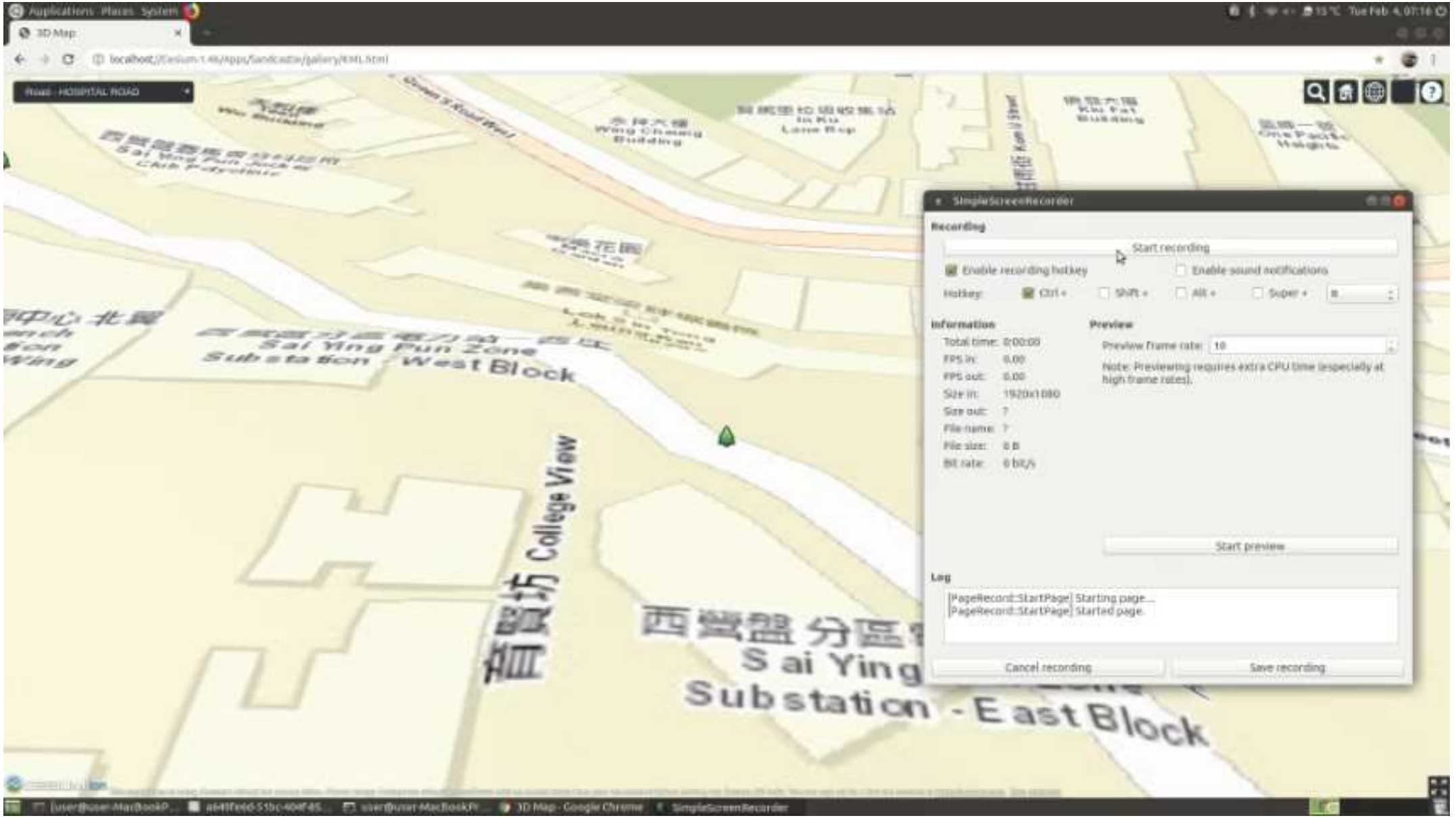
3D Map Visualization

3D

The distribution of detected trees along the Hospital road in 3D Map viewer

3D Map Demo

3D



Limitation

- Challenge in increasing the accuracy for estimation of geographic position of detected trees
- The effectiveness of the “field of view” depends on the distribution of trees
- Require more roads sample with ground truth data, the average offset value should more reflect the optimal value
- This study is only focussed on tree detection along roadside, which does not provide any information for tree species

Conclusion

- By applying CNNs using YOLOv3 network model is an efficient method for automatic tree detection and accurate geographic localization.
- Processed the 26 roads in Hong Kong Island to validate the approach it achieved mean distance of 3 meters with respect to ground truth data for 2 selected roads.
- Once the complete dataset for GSV images is downloaded, the pipeline is easy to be applicable to whole Hong Kong territory.
- Once the tree database is completed, on top of the inventory of the detected trees, it will be extended to tree species classification and health monitoring
- 3D GIS porotype system would complement the Hong Kong tree register database by reducing the traditional work of manual inspection.

Thank you