Shuojia Fu
Magnus Freyer
Evan Hymanson

Exploring Multifaceted Data Labeling Techniques for YOLO V5 Training in Wind Turbine Detection

Introduction and Problem Statement:

---Problem Description: Global Wind Energy Council claimed that by 2025 wind may provide 25-30% of the global electricity. As more wind turbines are built across the world, maintaining a complete record of wind turbine locations becomes a challenging task and add difficulties to land management..

---Proposed Solution: The study aims to identify wind turbine in the satellite images by computer vision model. Two model approaches are given 1. combination of ResNet and Saliency map, with YOLO V5. and 2. Grounding Dino with Yolo V5

Literature Review:

Past Approaches include saliency detection in high resolution remote sensing images, computer vision algorithms including Scale-Invariant Feature Transform, Speeded Up Robust Features, ResNet, MobileNet.

Dataset:

data source: Geolocation data from the USGS's wind turbine database was merged with imagery obtained from NAIP via Google's Earth Engine API ('USDA/NAIP/DOQQ')

zoomed_in_images: around 6 000 true(with wind turbines) and 6 000 false images((without wind turbines) with size 130x130 meters.

zoomed_out_images:around 20 000 true and 20 000 false images with size 290x290 meters

Methods:

Approach 1: Applied modified ResNet50 with 3 additional layers: on both two dataset, combined with saliency map, determine bounding box and coordinates of the wind turbines, and input them into the YOLO V5.

Approach 2: Applied Grounding Dino model on 290*290 meter dataset, determined bounding box and the coordinates, input into the YOLO V5.

Conclusions and Future work:

Approach 2 provided significantly better results than approach 1. However, the model did show minor limitations, such as occasional false positives, particularly with images taken from the edge or those containing shadows. This can be attributed to the model interpreting shadows as turbines that are partially outside the frame.

Future work should focus on expanding the input images to include different terrains and countries to make it more generalizable. Also, images that cover a wider area in the physical terrain should be used to limit the incidence of turbine shadows without the turbine in training data and false positives in testing data.

Experiments and Results:

Approach 1: 130*130 meter dataset: The modified ResNet model achieves a 96.67 percent accuracy with 2 epochs, and a 98.4 percent accuracy with 8 epochs in the validation set. Most of the images containing single wind turbines display a good saliency mapping on the location of the wind turbine. However, for some images, saliency map includes the shadow of the wind turbine. And for images containing multiple wind turbines, saliency map tends to show a single saliency area containing all the wind turbines.

290*290 meter dataset: decrease in accuracy in ResNet and more noises in the saliency map

Approach 2: 290*290 meter dataset:for some images, multiple bounding boxes are drawn around the same wind turbine or a big bounding box is drawn containing all the wind turbines. To solve such a issue, we determined the overlapping of the bounding boxes and eliminate the largest ones.

The Yolo V5 model was trained for 20 epochs, demonstrating a consistent drop in loss over each training period. The same occurred for training on the bounding box locations of each turbine image. Additionally, mean average precision of the model increased over training epochs. The same trend occurred for the validation set as well.

–West Coast Testing:Out of 56 generated images with 20 instances of wind turbines. It identified all 20 with confidence ranging from 38- 87 percent. It had 3 false positives.

–East Coast Testing: It correctly labeled 15 out of 16 classes of wind turbines across 56 images.The correct labels had confidences of 36-79 percent. It incorrectly labeled 5 objects or shadows as wind turbines

Stanford University