Technology Selection

First of all, you use the labelImg to create their own label, the which is good, but I am not sure that they should include the whole utility pole or not. For now you just include the top of the utility pole, but the rest of the poles are not included, for what I saw in the examplecapture.py, the coordinate of top and bottom are needed, so maybe you should include the whole part of the utility pole when labeling your images?

Second, you choose YOLO v2 as your main model, I don't know whether you want to make your program to be real-time detection or not, and because I haven't seen your training model and result, maybe after you get a better GPU and get the result, you could consider change other models, and make comparison with your current YOLO, so that you could find the model suits the target most.

Third, the Deep learning framework you choose, darknet which is written by C? I am not sure it has a visual interface like Tensorboard, which could help you get the information about your training process. If it does not have one, maybe you could consider changing another framework.

Fourth, I don't see the code about the Google Street view API and the ArcGIS, in your next step maybe you should add them in separate python files.