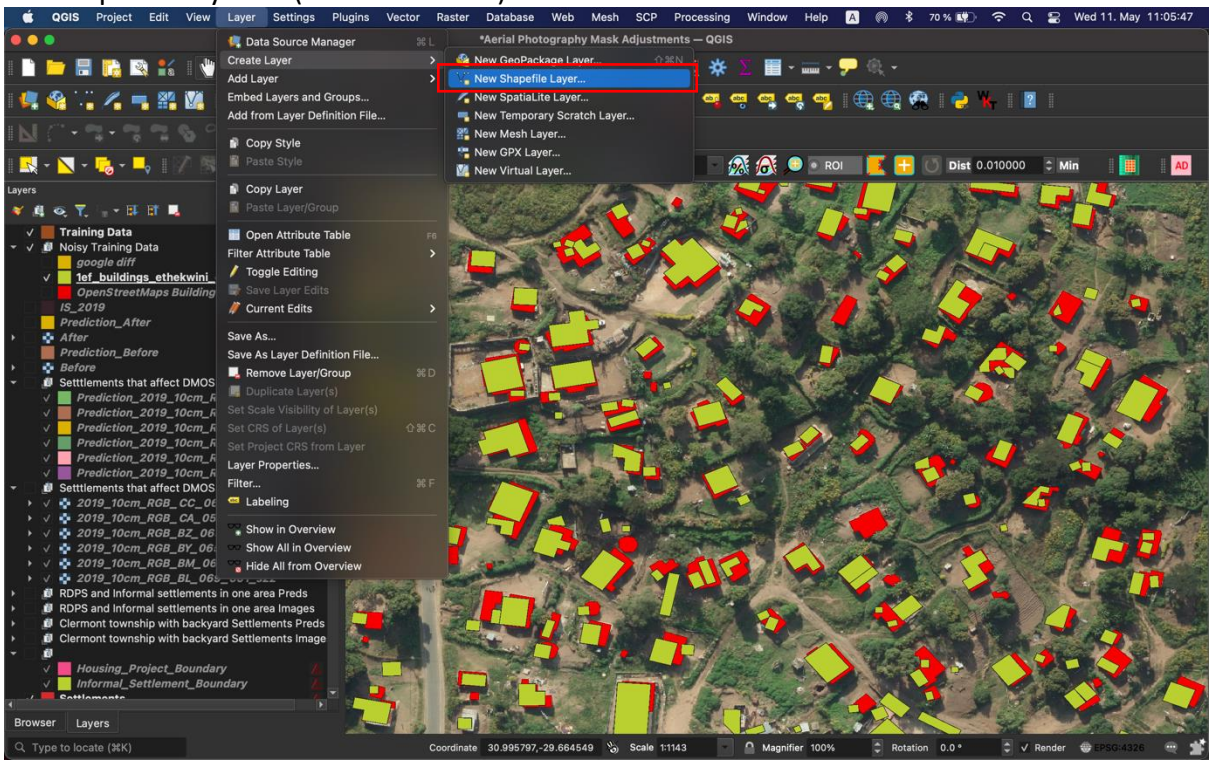
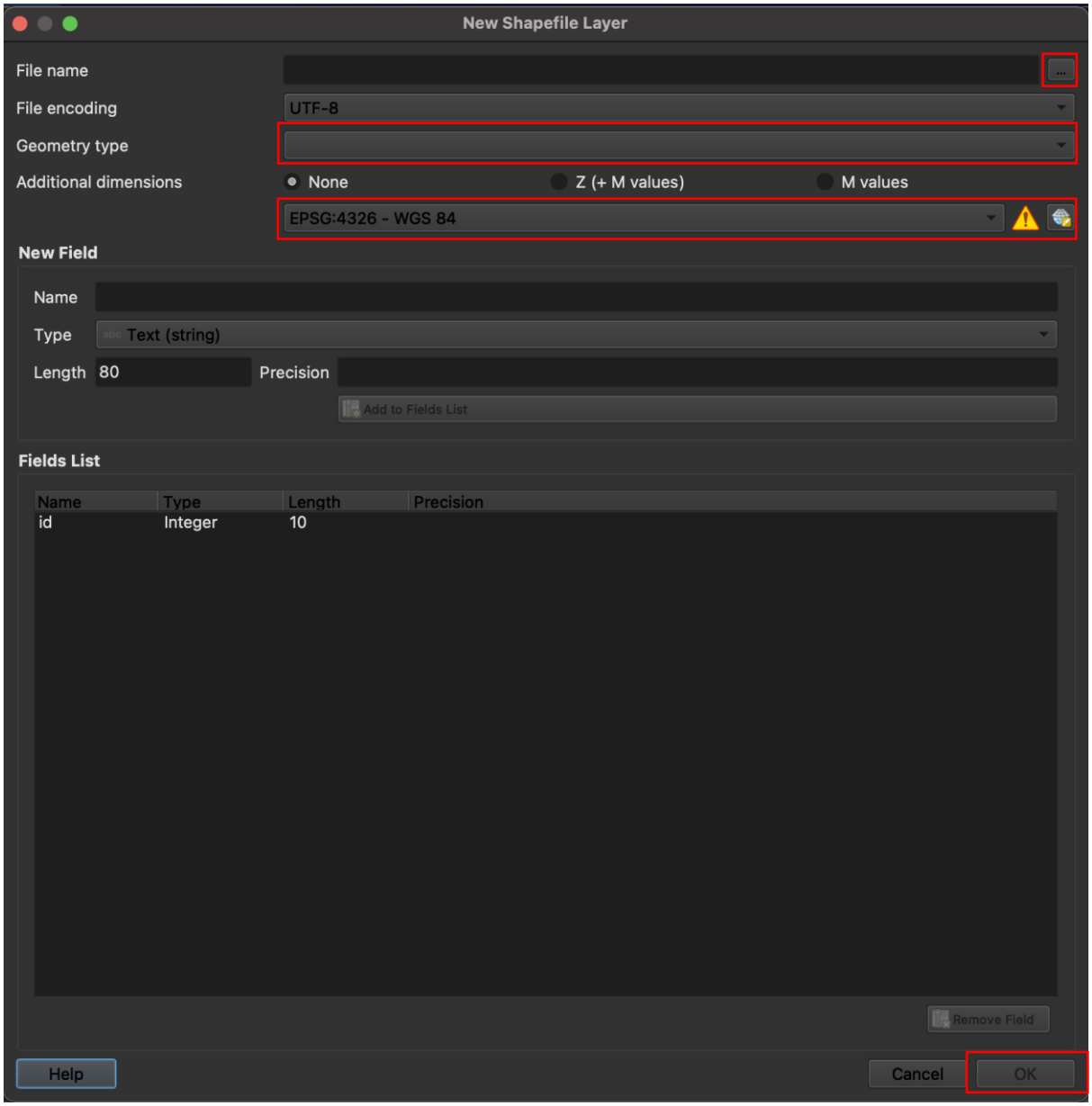
Documentation on how to create new training data for the UNITAC AI Building Tracker in QGIS

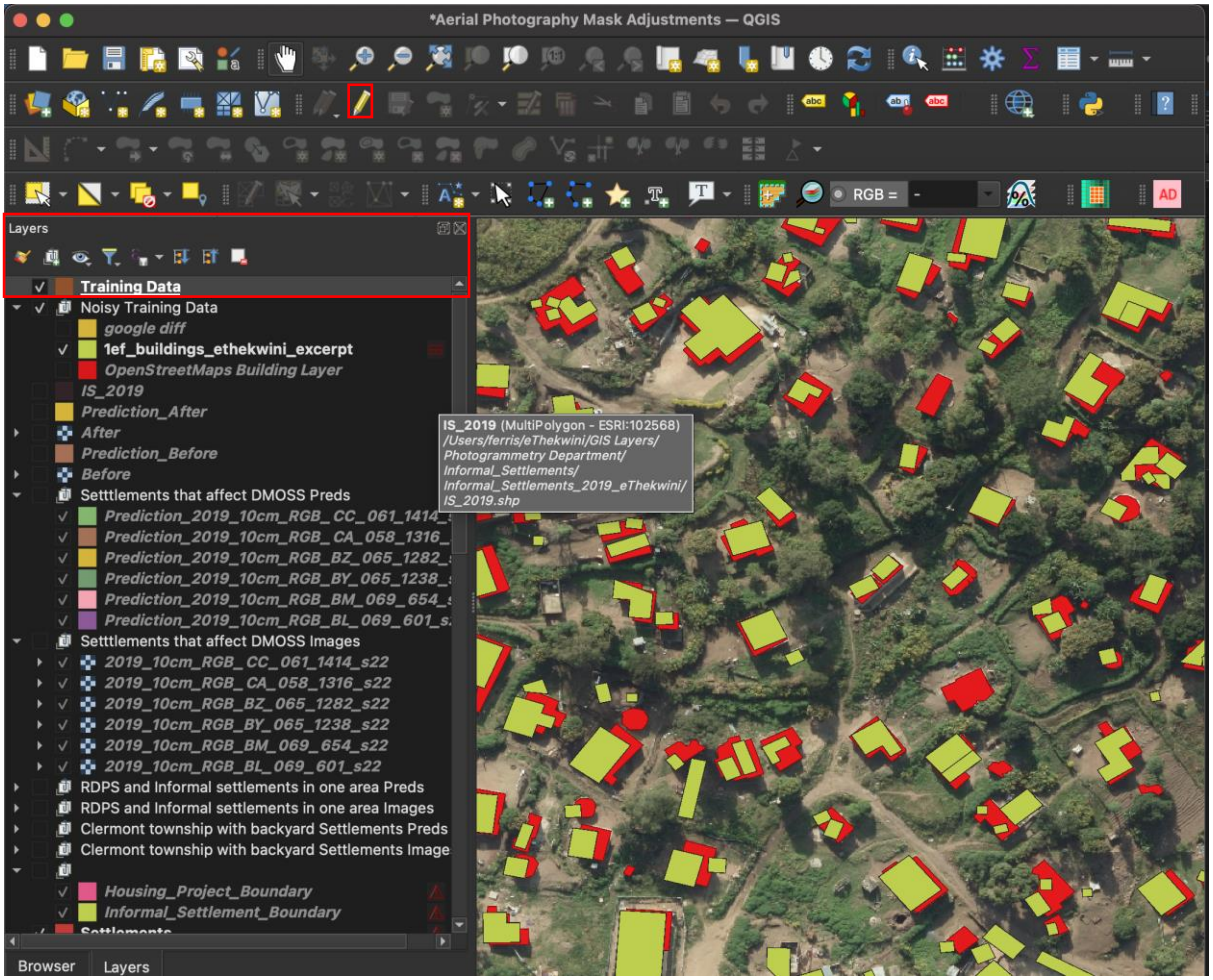
1. Open QGIS.
2. Load training images into the tool.
3. Select the “Layer” menu from the menu bar and go to “Create Layer” and then “New Shapefile Layer…” (see screenshot).



4. Now click on the three dots at the top-right of the window (see red box in screenshot below) and navigate to the folder where you would like to store the training data and give the file a name that you will later recognise. Click on the “Save” button.
   Now choose “Polygon” in the “Geometry type” dropdown menu and the coordinate reference system (CRS) that your aerial photography is in from the “Additional dimensions” dropdown menu. Finally, confirm your choices by clicking on the “OK” button at the bottom right.

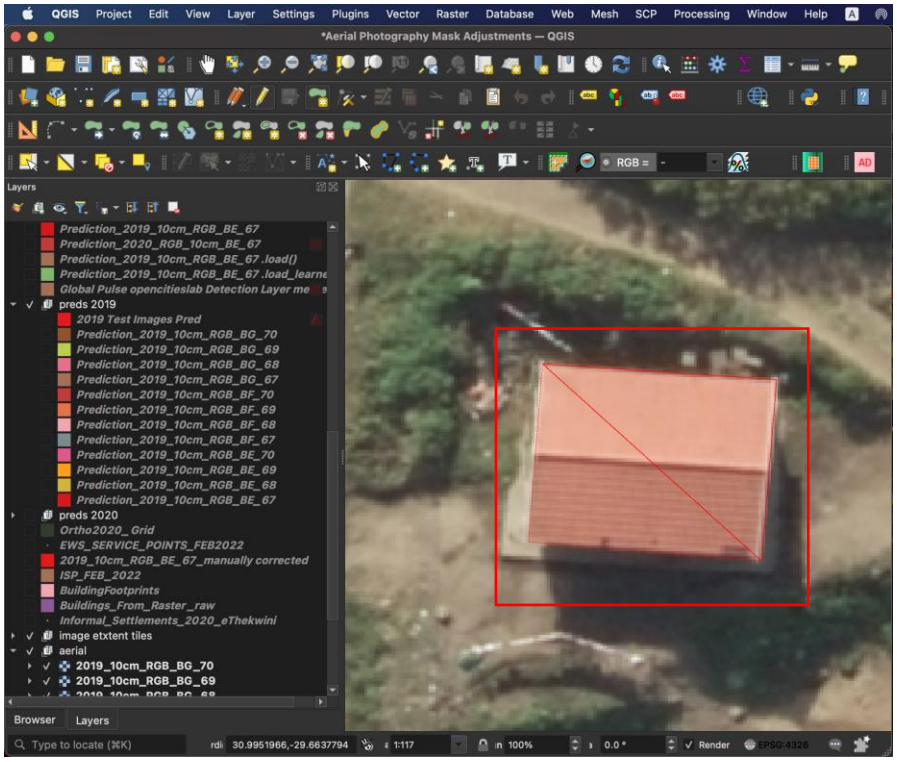


5. Now you will see a new layer with the name you chose in the “Layers” menu on the left hand side of QGIS (in this example “Training Data”. Click on the layer and click on the “Toggle Editing” button in the menu above (marked in red in the screenshot below). Now you should see more buttons from the QGIS menu being useable.

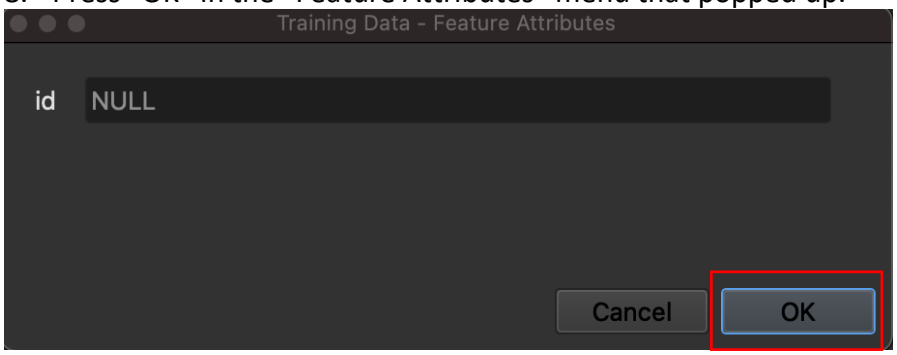


6. Click on the “Add Polygon Feature” button marked in red below to be able to start marking building pixels. Now click on the corner of a building you would like to mark and keep clicking into the other corners until the entire building is marked in red

7. Now click on the corner of a building you would like to mark and keep clicking into the other corners until the entire building is marked in red. To confirm your selection, right-click once.



8. Press “OK” in the “Feature Attributes” menu that popped up.



9. Continue marking all buildings in the images you would like to use to create training data (for eThekwini, one image of 10,000 x 10,000 pixels corresponding to one square kilometre on the ground was used).
10. Once you are done, click on the “Save Layer Edits” button in the menu (marked in red below).

11. You can now close QGIS and open the Python script used for tiling images and labels.

12. The script is set up to run with the data being stored on Google Drive. Now you will have to make changes under “3. Set Paths for Images and Masks”, where you need to adapt the directory for the “path”, “images_list”, and “shp_path” variables to match where you stored these on your computer. Lastly, you will also have to adapt the third parameter of the “save_masks” function under “Create Masks Corresponding to Selected Images” to match the file location on your computer.

13. You can now execute the script. The small image and label tiles will be saved in your chosen output folder and you can now use them to train the deep learning models.