After running some simple tests to preprocess the images, to build a proper model the images that were captured needed to be recognized by a model, the images needed to be annotated. In this case taking the images using the 'labelmg' mg library in order to create a XML file which would help in the program detecting the position and class of the objects being detected. After setting up a script to automate the images for lablemg, it turns out the library was outdated and wasn't very accurate when it came to determining the size and shape of objects that were held in hand, as it kept annotating my hand instead. I have now decided to switch to an incredibly simplistic form of model training in the form of google's 'Teachable Machine'.