YOLO

YOLO use a pre-trained model, repurpose classifiers to perform detection. They use darknet frame work to conduct their training. This model is pre-trained by images from image.net. The advantages to use a pre-trained model is that it will have better accuracy, shorter training process and a faster training time.

They label images before they train the model. Then this network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.

Label Images

They build an application to label images. This application can read in images, then you can draw rectangular to mark object and tag them. Then the app will create a document file associate with the image, write this information, coordinate of four angels of rectangular and tag message inside. After that images and their corresponding text file will feed into the model to conduct training.