The paper presents a architecture to combine Fast R-CNN and RPN for both object classification. The idea is to allocate the position of object with a bounding box around it. A small network will produce a score with sliding a window over a feature map. Each sliding window generates 9 anchors with different scale and aspect ration. The anchor is also used to feed the classification layer and regression layer. A pyramid of anchors can works with singe scale of image, feature map and filter, which is more cost-efficient. Classification layer and regression takes the anchor information to compute the loss function based on IoU with the ground-truth box. The important part of the architecture is to combine the information of proposal and feature maps. The paper provides a number of strategies to share the features fro RPN and fast R-CNN. The paper uses the alternative training which uses the trained RPN for fast R-CNN. Both of them are trained based on pre-trained ImageNet model with fine-tuned. Shared convolutional layers are fixed and unique layers are fine-tuned. But it should prove more explanation for this sharing features. Because we may assume that all the convolutional layers are somehow share some low level features. It can be possible that even the low level features are not learned the sharing can still work. More explanation should be made to justify why it works.