

# Rich feature hierarchies for accurate object detection and semantic segmentation

## Paper Summary

Ross Girshick Jeff Donahue Trevor Darrell Jitendra Malik  
UC Berkeley

## 1 Summary

The paper proposes an object detection algorithm combining a pretrained CNN model and region proposals. Hence, the model is named Region Based Convolutional Neural Networks (R-CNN). The model basically puts bounding boxes on objects on the input images. The algorithm uses selective search to decide where to put the bounding box which is basically a rectangle with an object in it. These regions are fed to compute features and classify them using linear SVM classifier. A comparison between OverFeat (another then proposed sliding window detector-based method) and R-CNN has been shown. Where R-CNN outperforms OverFeat by a very significant margin.

## 2 Contribution

Selective search-based region proposals for feature extraction and classifying them with SVM classifier is the major contribution of the paper. Proposed system basically extracts a number of region proposals for input images and sends them to convolution. Which finally gets classified using linear SVM. Also, a supervised pre-trained CNN is proposed if the amount of data is not sufficient. ILSVR2013 (200 class) was used for training.

## 3 Limitations

The computation time is very high for R-CNN. Since, the model has to classify 2000 region proposals for each image, it takes a huge time for training. Also, it uses three different models altogether to do the task of object detection which are – CNN for feature extraction, Linear SVM classifier for detecting objects and Bounding box regressor. As a result, it takes longer times to predict every new image which lessens the opportunity of using such model for real life bigger datasets. Selective search algorithm doesn't learn anything while training. Hence, there is possibility of proposing bad region proposals.

## 4 Result

OverFeat's previous best result was 24.3 percent. Where R-CNN's mAP is 31.4 percent, a significant improvement.