Name: Angona Biswas

Task 2

Paper Summary:

Paper link, Author name, Publication date	Dataset	Limitations of research work	Short review of the model architecture and its working process			Major Findings or Contribution
			Preprocessing and post-processing	Feature Extraction	Model for Classification	
https://arxiv.org/abs/1311.2524 Ross Girshick el. al, 2014	PASCAL VOC 2007, 2010-12, ILSVRC2013 VOC has the 20 number of classes and ILSVRC2013 has 200 number of classes.	-Comparison is made with recent method but why only CNN is used for feature extraction, why other process is not used to claim is as efficient method is not explained. -Again, why linear SVM is utilized for detection, why not other machine learning algorithms- is not discussed. -Any image processing methods are not applied to RGB image datasets. - Time consumption analysis is not explained in this paper. - Feature reduction process is not applied but there is large number of images.	Generation of region proposal for tracing the set of objects for detection.	Feature vector is extracted by CNN. Feature vector dimension is 4096. Input image is RGB image and the dimension is 227×227. Network architecture consists of 5 convolution layers and two FC layers. SGD is used for fine tuning. Author named the overall method: "R-CNN" or Regions with CNN features.	Support Vector Machine (SVM) is used for training and classification with linear kernel. Hard negative mining method is adopted for large memory.	- Proposed an improved detection algorithm (R-CNN). This research work improves the mean average precision. - Authors explored that their proposed R-CNN outperforms the OverFeat by the large 200-class ILSVRC2013 dataset. - Grid search optimization is used to find out the negative region by using IoU overlap threshold. - Combined the deep learning and machine learning method.
https://arxiv.org/abs/1504.08083 Ross Girshick, 2015	PASCAL VOC 2012, VOC07, 2010, 2012	-For large dataset, no feature reduction process is used. - To claim the efficiency of the model, besides SVM, other ML models are not applied. - Exact training and test time are not mentioned. - Real time detection is not possible through this method.	Any kind of image processing steps are not used.	Author proposed an architecture of fast R-CNN. In this process, input image and multiple RoIs are input into convolutional network. Feature map of fixed-size is produced by pooling each RoI and FC	Output of network per-class bounding-box regression offsets and softmax probabilities.	-Fast R-CNN based model is applied for object detection where author contribute to speed up the performance and increase the accuracy. - Author utilized VGG16 for this experiment and compared it with SPPnet and observed that proposed

				layers produced feature vector.		method is enough faster in case of training, testing. - Fine-tuned VGG16 for obtaining higher mAP. - For coding, Python and C++ are used. - For experimental evaluation, 3 different pre-trained model is used. - FRCNN's softmax layer performs better than SVM in this case.
https://arxiv.org/abs/1506.01497 Shaoqing Ren et al., 2016	PASCAL VOC 2007, 2012, and MS COCO datasets are utilized.	-Though authors proposed an improved new method of feature sharing model but for more observation, other models are not included for verification of accuracy. - Another improved detection model is YOLO model which is induced in 2015. It's a real time object detector and vary fast system. This model can be implemented to observed the performance of datasets.	Any kind of image processing steps are not used.	-Proposed model is composed of 2 modules. One is deep convolution network for proposing the network. Second one works a RCNN detector using those proposed regions. - Sampling strategy used to train this network is "image-centric". - standard deviation is 0.01. - Use a learning rate is 0.001 for 60k mini-batches, and 0.0001 for the next 20k mini-batches on PASCAL VOC dataset.	- Proposed RPN is fully convolutional network that simultaneously predicts object bounds and objectness scores. - The output of the RPN is score of objectness. - Authors explained that how they slide the small network over the conv feature map to generate region proposals. - multiple scales are addressed by a novel scheme for designing anchors.	-Author of this research work proposed a real time object detection process with Region Proposal Network - Proposed system improved the detection accuracy. - Numerous datasets are used to explore the efficiency of proposed model. -Proper score threshold is applied for better performance. -Authors were able to detect multiple objects. -Authors explored the best result of 69.9% mAP using different anchors settings.

		- momentum= 0.9 - Experiment used CaffeNet. - Proposed unified is the combination of RPN and Fast R-CNN. Both shared the convolution layer to create shared feature.	