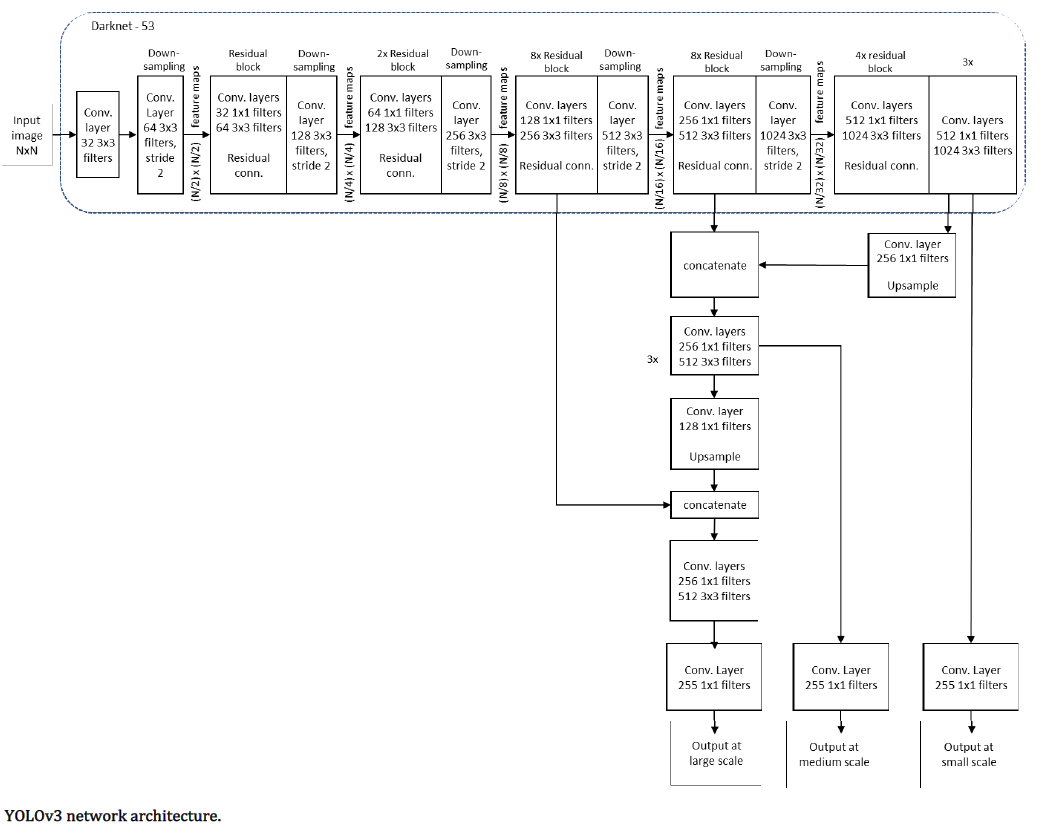
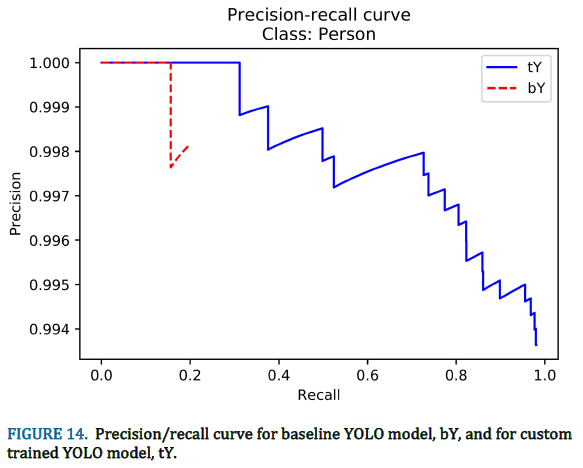
**IR detection:**

20-07-(Yolov3) Thermal obj. detection in difficult weather**:**

* Yolov3 is faster than R-CNN, SSD and Cascade R-CNN with comparable AP



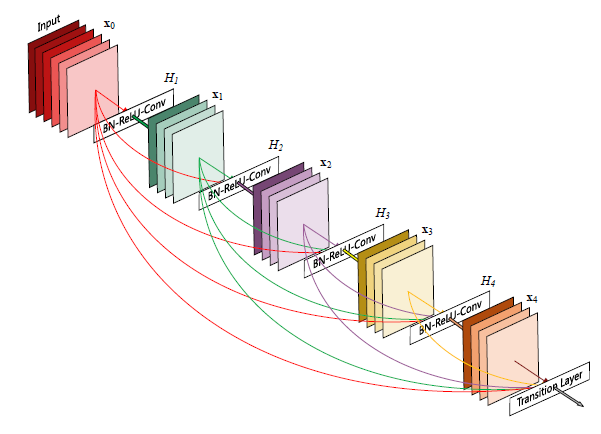
* Yolov3 pre-trained on COCO is used as baseline model → name it bY
* Baseline plus 4270 IR images training → name it tY
* For training details read page 9 of the paper
* Evaluation process is explained well, might be needed later
* Detection is compare to ground truth and is true positive if IoU > 50%

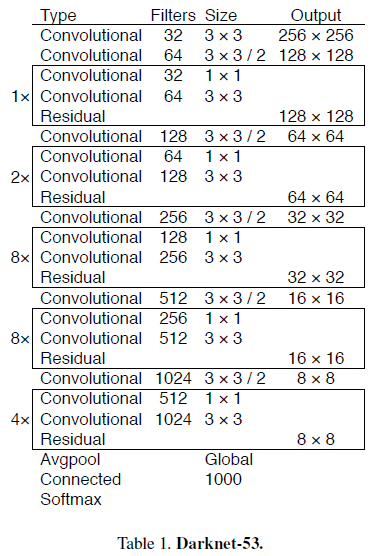


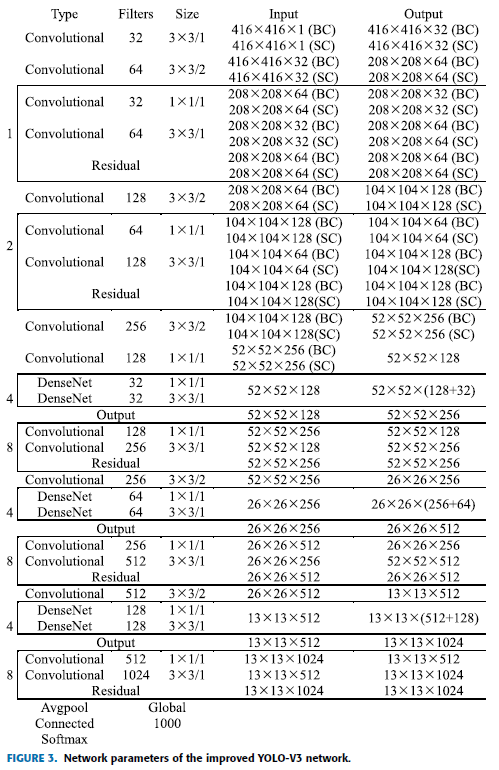
* Precision-recall curve for a class, provide the AP
* Original Yolo on person → AP 19.63% with 15.5% recall at 100% precision
* Yolo with 3k IR image trained on person → 97.93% AP
* IR performs better in rain since the temperature difference is higher
* They demonstrate that Yolov3 can learn with relatively small dataset (1k) and small number of iteration (1600)
* Trained Yolo shows good generalization properties w.r.t non seen images

20-04-(Yolov3)Using deep learning in infrared images to enable human gesture recognition for autonomous vehicles:

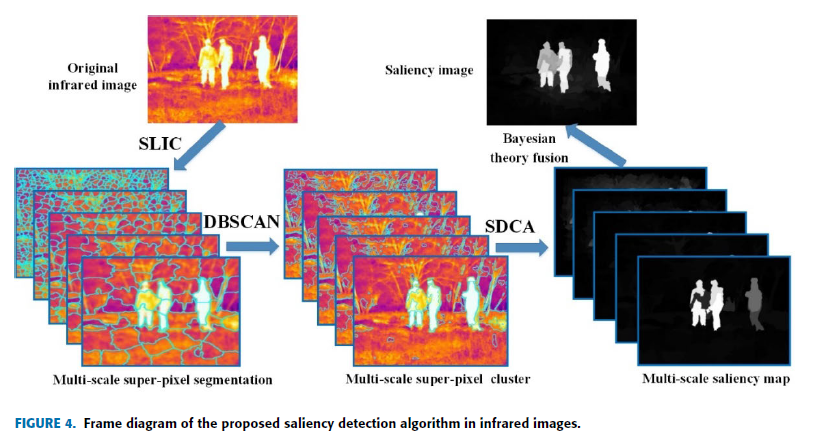
* The network take two inputs. IR frame and its corresponding saliency map to enhance the reuse of features
* Three DenseNet blocks are added before the residual components in Yolov3 to enhance the convolution feature propagation (i.e. feature reuse)
* The feature map from two inputs are concatenated followed by a 1\*1 convolution to linearly merge the features.
* In the process of Yolov3 a large number of feature information is lost and since a target object in IR provide relatively less number of features, it is unfavorable
* To address this issue DenseNet blocks is used before the residual blocks of Darknet-53





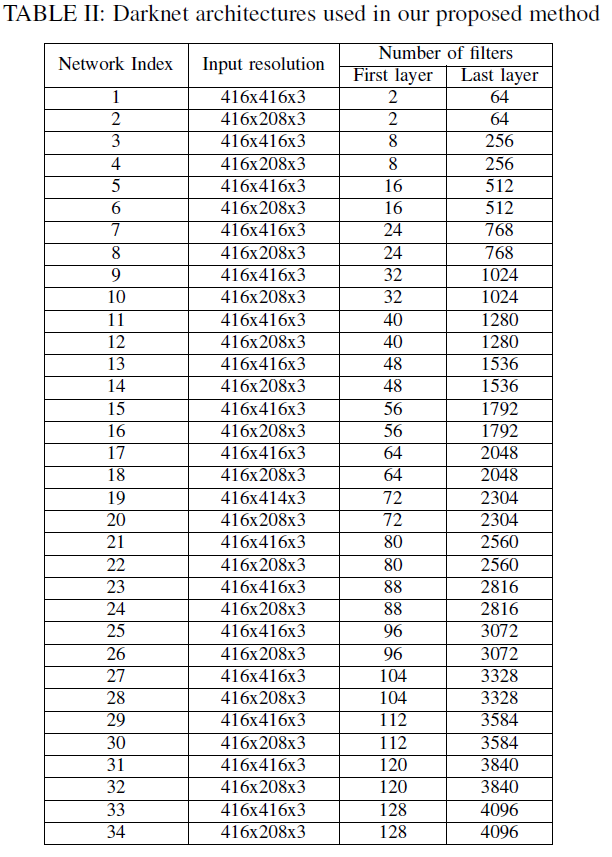


* Due to lack of information in IR frame as well as its high sensitivity to temperature changes, saliency maps is used as the second input
* Read the paper for exact process of saliency map generation

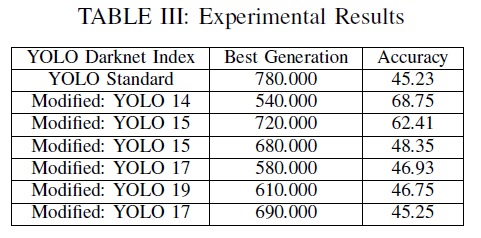


18-12-(Yolov2)Obj. recognition on long range thermal image using state of the art dnn:

* They have tried 34 Networks with different input output size and etc.



* Dataset is one million big and is owned for security purposes
* In total 2720 different setting led to 2720 different results. The best is shown below for 8 classes such as Human, boat, vehicle and animal in 2 size of tiny and normal



* Generation is referred to epochs, I suppose