Project Proposal

Mohamed Kamal Abd Elrahman	60
Mohamed Elmaghraby Mohamed	55

- **Moving Object Detection**
  - Object Detection is a common Computer Vision problem which deals with identifying and locating objects of certain classes.
- **Motivation**
  - Object detection in the stream of video is a core part of many applications in camera surveillance.
- **The current state of the art accuracy under real-time constraints**
  - In object detection accuracy on COCO(Common object in context) is 37.4% mAP(mean average precision) achieved by CenterNet on 52 fps.
- **A short survey of available models under real-time constraints**
  - CenterNet is a new approach to object detection problems where the network uses Hourglass blocks and detects the center point of the object.
  - CornerNet with its variation (original, squeeze) again uses hourglass blocks output top-left corner and bottom-right corner.
  - SSD detectors
  - Yolo
  - RetinaNet

- **A detailed description of the model to be used to build on.**
  - Use the tiny-Yolo model to build on it

Layer	Filter	Filter size	stride
Conv	16	3x3	1
Max pool		2x2	2
Conv	32	3x3	1
Max pool		2x2	2
Conv	64	3x3	1
Max pool		2x2	2
Conv	128	3x3	1
Max pool		2x2	2
Conv	256	3x3	1
Max pool		2x2	2
Conv	512	3x3	1
Max pool		2x2	2
Conv	1024	3x3	1
Conv	1024	3x3	1
Conv	1024	1x1	1
Fully connected			

- **Why you will use this particular model.**
  - Because Yolo considered the first model to achieve good performance near real-time.
  - Use the tiny version for speed of training and inference.
  - And has a simple architecture
- **The proposed updates to the model.**
  - We proposed since we are working on a stream of videos to use a background subtraction technique to use a foreground mask and integrated it as another channel to the image input.
  - make the model smaller
- **What benefit do you expect from this update**
  - Increase the accuracy of detecting moving objects without a significant increase in computational power.
- **Why do you think it is a good idea to try it?**
  - Because the foreground mask is sparse and expected to carry most of the information of the moving object
- **How will you evaluate your results**
  - Use the standard approach (average precision) and compare the result against other detector networks.
- **What kind of evaluation metric you will use to compare your results**
  - For every detected box compare it with ground truth using IOU(intersection over union) if IOU > 0.5 consider a detection.
  - And calculate precision accuracy.
- **What types of plots will be used to point out the comparison results**
  - Scatter plot Accuracy in (average precision) vs time of inference
- **A survey of available datasets for your course project problem**
  - MOT (multi-object tracking) video of labeled bounding boxes, standard benchmark competition for object tracking.
  - Video object tracking on Kaggle.

- **What datasets you will use**
  - We are going to use Mot(multiple object tracking) benchmark dataset composed of several videos and used to benchmark object tracking MOT site
- **A detailed description of the dataset to be used**
  - A sequence of image and for every image the bounding boxes of objects
  - And there exist sequences of videos under different density.
- **Why will you use this particular dataset**
  - It provides different videos and the labeling for each video
- **Your graduation project brief problem statement, even if it is not the same as the proposed course problem statement.**
  - Video Synopsis: A method to generate a time-bounded video by grouping the moving objects to generate the synopsis in a shorter time video. Grouping is done with respect to relative Spatio-temporal distance and chronological appearance of the objects.
  - The main part of our project focuses on scheduling the activity tubes (array of frames describe each object from its appearance till it disappears) to reduce the collisions between the objects, maintain the chronological appearance and decrease collisions.
- **If your proposal is related to your graduation project, point out the differences that will be made between this problem statement and your graduation project problem statement.**
  - In our graduation project, we focus mainly on the scheduling problem also we use other techniques like
    - Object detection model.
    - Object tracking Model.
    - Segmentation.
    - Stitching.

- **resources/papers**
  - Redmon, J., and Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection
  - Redmon, J., and Farhadi, A. YOLO9000: better, faster, stronger. Computer Vision and Pattern Recognition(CVPR) (2017).
  - More to be added while working