Motion Detection and Tracking

Nicholas Mok, Defense Technology Agency

Sam Delaney, Jesselle Grata, Bruce Young

Summary

Commissioned by Nicholas Mok, a research scientist of NZ's Defence Technology Agency, our team's project was to research and analyse the performance of existing tracking algorithms. Initially, we had been given the task to analyse detection as well as tracking algorithms but due to a change in scope with our client, we focused on using the detection results of one detection algorithm to be used in our trackers.

The detection algorithm that we decided to use was "Mask R-CNN". Mask R-CNN is a neural network based algorithm that aims to solve the instance segmentation problem (Zhang 2018) in computer vision where objects of the same class are grouped together than separately.

Chosen Trackers

For our tracking algorithms we decided to use the OpenCV library which was mentioned to us by our client. Using OpenCV meant that multiple, different tracking algorithms were available to us and our team would be able to work with the same format for every algorithm and we would not have to struggle with converting different outputs. This also meant that we were working with a single language (Python) which would be simple for our client to read, understand and potentially use for himself. The trackers we used were:

- GOTURN (Generic Object Tracking Using Regression Networks)
- KCF (Kernelized Correlation Filter)
- CSRT (Channel and Spatial Reliability Tracking)
- MOSSE (Minimum Output Sum of Square Error)

Purpose, Objective and Goals

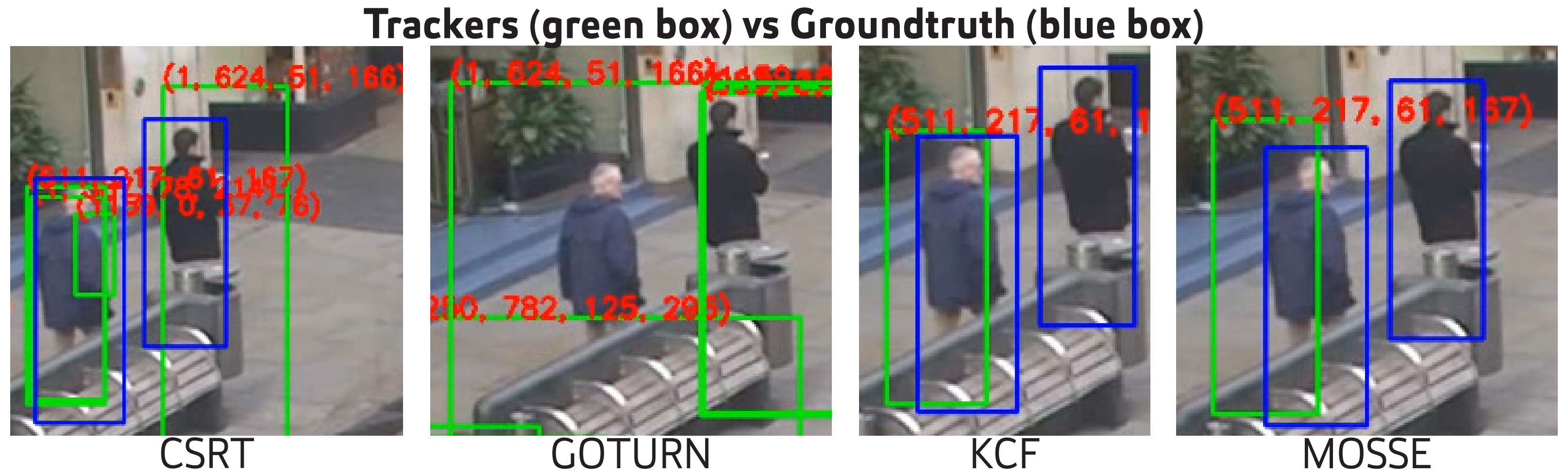
The purpose of this project was to evaluate modern computer vision techniques which consisted of analysing modern tracking algorithms in terms of their speed and accuracy. This was later changed where connected camera systems would be able to track an object of interest through multiple public cameras and that the tracker is able to retain an object's identity.

Methodology

The team methodology we chose was 'Agile Scrum'. In the planning phase of our project, team was unsure just how much research and development was needed to be done for this project. Therefore we concluded that it was appropriate to use agile scrum. Using scrum, this allowed the team to produce sprints where the focus was either on research or development and allows for unique actions for during the different phases.

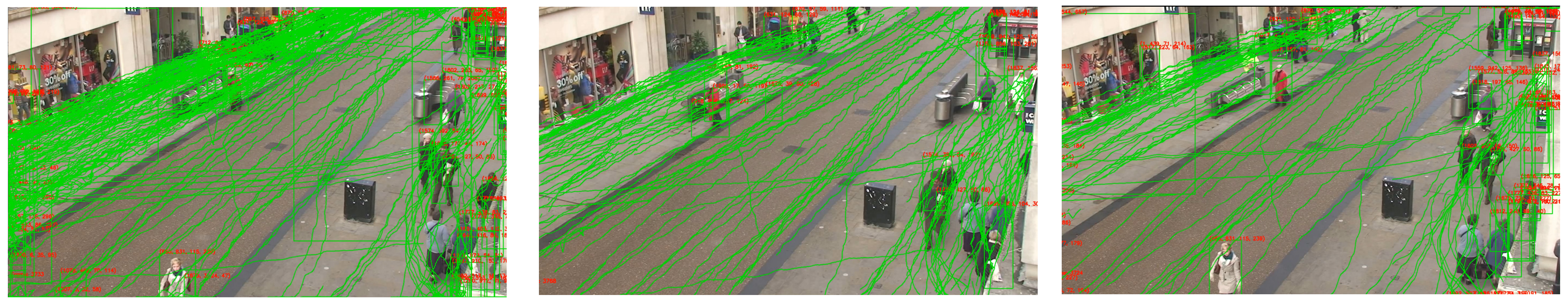
Results

Trackers (green box) vs Groundtruth (blue box)



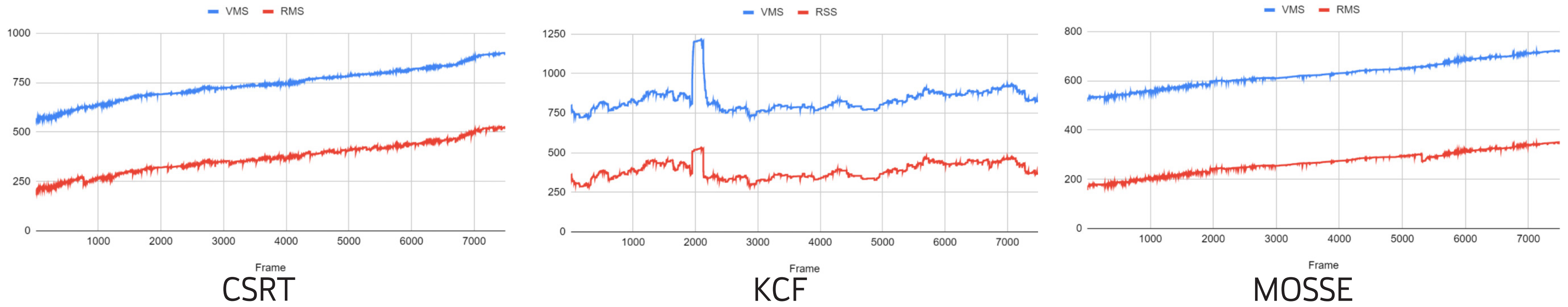
CSRT GOTURN KCF MOSSE

Generated Heatmap



CSRT KCF MOSSE

Memory of Usage (MB) over Frame - VMS (blue) & RSS (red)



CSRT KCF MOSSE

Conclusion

We found that GOTURN was significantly worse than the others and our system could not handle it.. This is because it needs to be trained and it works of movement meaning that it was basically unusable. CSRT had very good accuracy however it seems unable to remove old bounding boxes and since it is not very fast to begin with it made it far too slow to be the best. MOSSE was the fastest out of all of the algorithms however it frequently lost track of objects meaning that it had very bad accuracy because of this we found that KCF was the best. With good speed and accuracy we conclude that for our project KCF is the best tracker.

Over time, the computational strain on the system's memory rose as more trackers were added that were not removed. This could be fixed by setting a limit on the amount of trackers that can be present in one scene at a time, removing either the earliest or the tracker sees that its target has not moved for the past few seconds.

Another area of improvement would be to use detection algorithms to find each object in a frame and use trackers to link consecutive frames in order to keep a target's identity. This could potentially solve the problem where trackers would lose reference to their targets if they are obstructed and intuitively create new trackers when new targets enter the frame.

Development Process

1. Using Mask RCNN to process the video
   a. Produced a text file of boundary boxes
   b. Produced a video with masks and boundary box
2. Converted data to be in the correct format for our program
   a. Produced converted text file
   b. Produced text file of new boundary boxes for each frame
3. Wrote a program to enable each trackers to:
   a. Display boundary boxes.
   b. Display a heatmap
   c. Output a video of the tracker process
   d. Output a tracker text file of the memory that is used.

Challenges and Learning

When testing the algorithm mask RCNN, the problem the team had was with tensorflow. The current version of tensorflow was not capable with the version of mask R-CNN. Therefore the team had to downgrade the version. Also some of the other team mates were unable to get the algorithm working due to technical difficulties.

A key learning point was how the team learnt the importance of peer programming. The team had been having some problems with part of the code for testing the algorithms. With the use of peer programing the team was able to solves these problems three times faster than if they had done it alone.

The team learnt about the importance of client feedback at one stage of the project a team member was unable to setup a environment on their pc to do MASK RCNN. After discussing this with the client they suggested they use a computer the process the data and give the result of that to the team member who could not setup up the environment since after it was processed they could work on it.

References

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017, December 25). Mask Rcnn. IEEE Xplore. https://ieeexplore.ieee.org/document/8237584

Matterport/Mask_RCNN. (n.d.). GitHub. https://github.com/matterport/Mask_RCNN.git

Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2012). Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. In European Conference on Computer Vision (pp. 702 - 715). Springer. https://doi.org/10.1007/978-3-642-33765-9_50

Rosebrock, A. (2018, July 30). OpenCV object tracking. PyImageSearch. https://www.pyimagesearch.com/2018/07/30/opencv-object-tracking/

Supervisor: Professor Edmund Lai

Semester 2, 2020