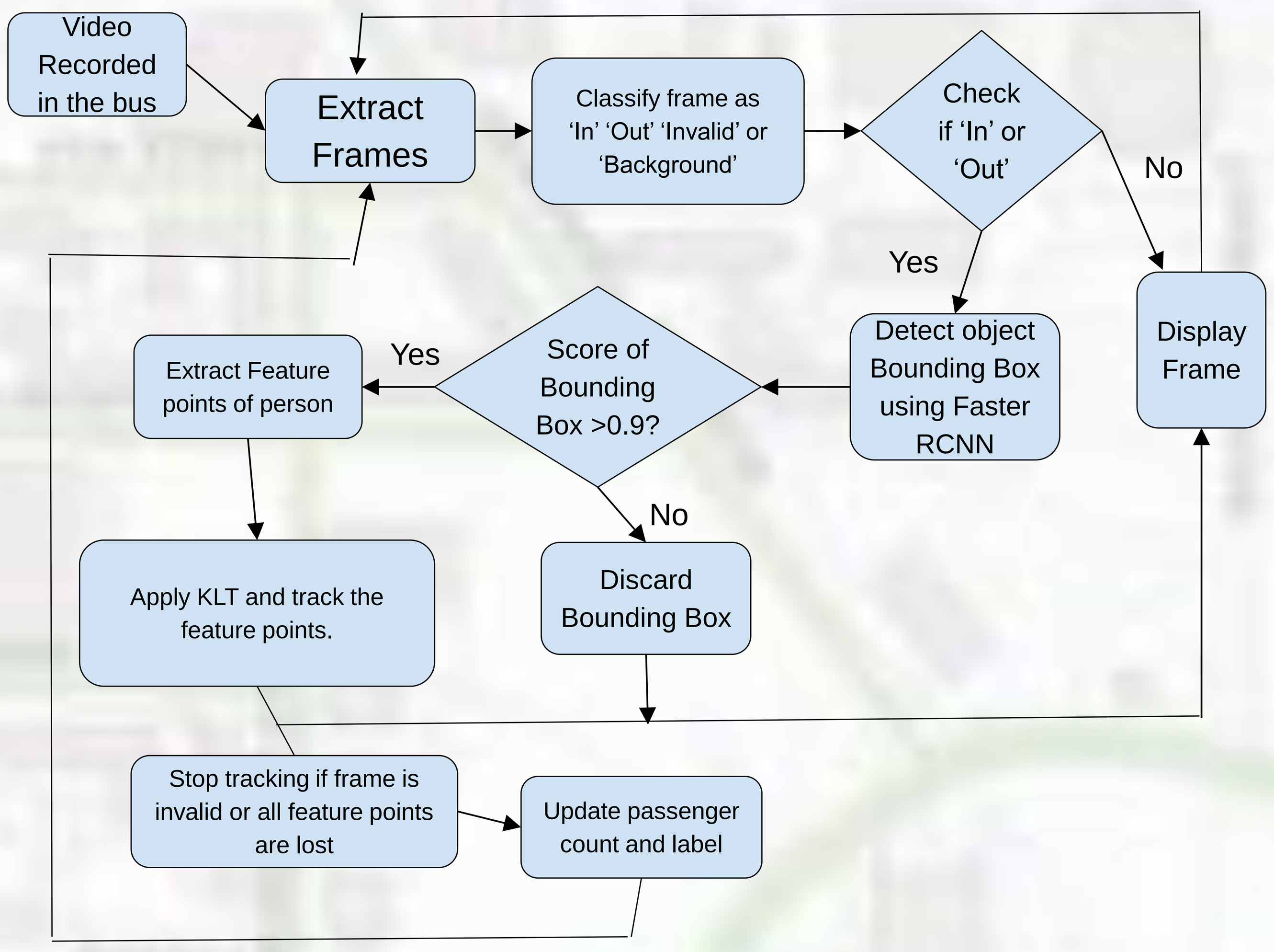
Automating Flux Calculation In A Georgia Tech Stinger

Anush Ananthakumar | Shubhankar Borse | Madhuri Phute

Motivation

- Automation of human counting and tracking mechanisms have been an important research area in the computer vision domain. Various researchers have developed algorithms utilizing Image segmentation and machine learning concepts to achieve automated counting.
- In this project, Deep Learning is utilized along with a tracking algorithm to automate a passenger counting system.

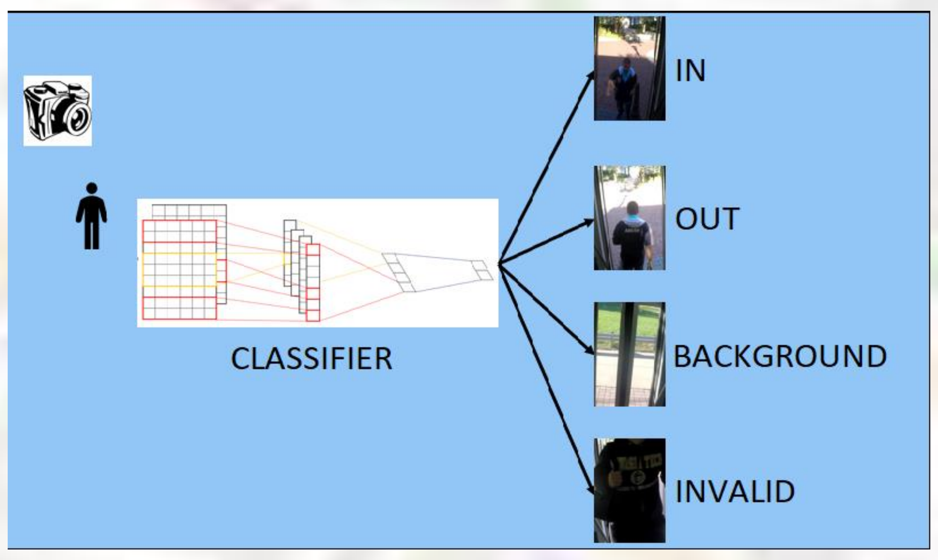


Project Workflow

Classification

We implemented a 3 stage algorithm; a CNN Classifier, a Faster R-CNN to detect the person, and KLT tracking to track the person's features in consecutive frames until the person exits.

- CNN Classifier – Classifies frames into 4 classes based on the presence or absence of people.



Labels given by the Classifier to frames

- The accuracy of the classifier can be judged by it's confusion matrix, which is as follows

Categories	Background	In	Invalid	Out
Background	0.9667	0	0.0280	0.0053
In	0	0.9643	0	0.0357
Invalid	0.0206	0.0270	0.9524	0
Out	0	0.0320	0.0400	0.9280

Confusion Matrix of Classifier

Detection

- Faster R-CNN Detector – Detects and produces a bounding box around the person.
- Faster R-CNNs use Region Proposal Networks to reduce the computation cost to a marginally small amount.

Bounding Boxes as detected by Faster R-CNN

- The intuition here is that the computed region proposals depend on features calculated in the forward propagation stage of the CNN, and thus, during real time testing, the results of the convolution layers can be shared

Tracking

- The features of the detected object are extracted and passed to the KLT tracking algorithm.
- The Kanade Lucas Tomasi (KLT) feature tracker was developed to perform local search using an approximation to the second derivative of an image.

Detected Features of Passengers Entering and Leaving the bus

```
Number of people ENTERING the bus:
     4

The labels for each person Entering/Leaving, in sequential order:
 OUT IN IN IN IN
```

Storing the Count of Passengers and Sequential Order in Which They Enter/Leave

Comparison of Detection Algorithms

- Table below gives the detection rate and time required for detecting a human in the frame using 4 different algorithms.
- A detection is considered as correct if the detected and ground truth bounding box overlap more than 60%.
- The Faster R-CNN is approximately 40 times faster than detecting a person in the frame with an R-CNN detector.
- Hence, although the detection reduces slightly, Faster R-CNN gives almost real-time detection.

Algorithm	Detection rate	Time required
R-CNN	96.86%	20.52 sec
HoG + SVM	89.22%	38.2 sec
Frame Differencing	82.5%	4.14 sec
Faster R-CNN	93.92%	0.419 sec

Comparison of Detection Algorithms

Conclusion

- This project is an attempt to present a novel approach to automate the passenger count and distinguish between entering and leaving passengers on a bus.
- The algorithm worked with high accuracy and successfully detected the flux in the bus.
- It was observed that Faster R-CNN works more efficiently for real time human detection and tracking in comparison to other contemporary algorithms.
- Future work in passenger counting can be focused around using semantic segmentation for extracting pixels corresponding to human and separating it from background for enhancing the feature extraction and tracking algorithms.

Acknowledgement

We would like to express our gratitude to Dr. Ghassan AlRegib for providing us an opportunity to pursue this project and provide continuous support.

Further we also thank the Georgia Tech Parking and Transportation Services for permitting us to capture video from the bus.