HARRY POTTER MARAUDER'S MAP

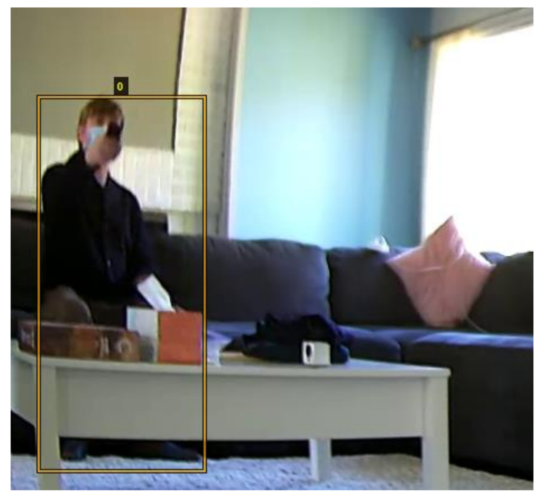
STUDENTS: MICHAEL MUSTER, WILLIAM WRIGHT, YANG XU

Motivation & Requirements

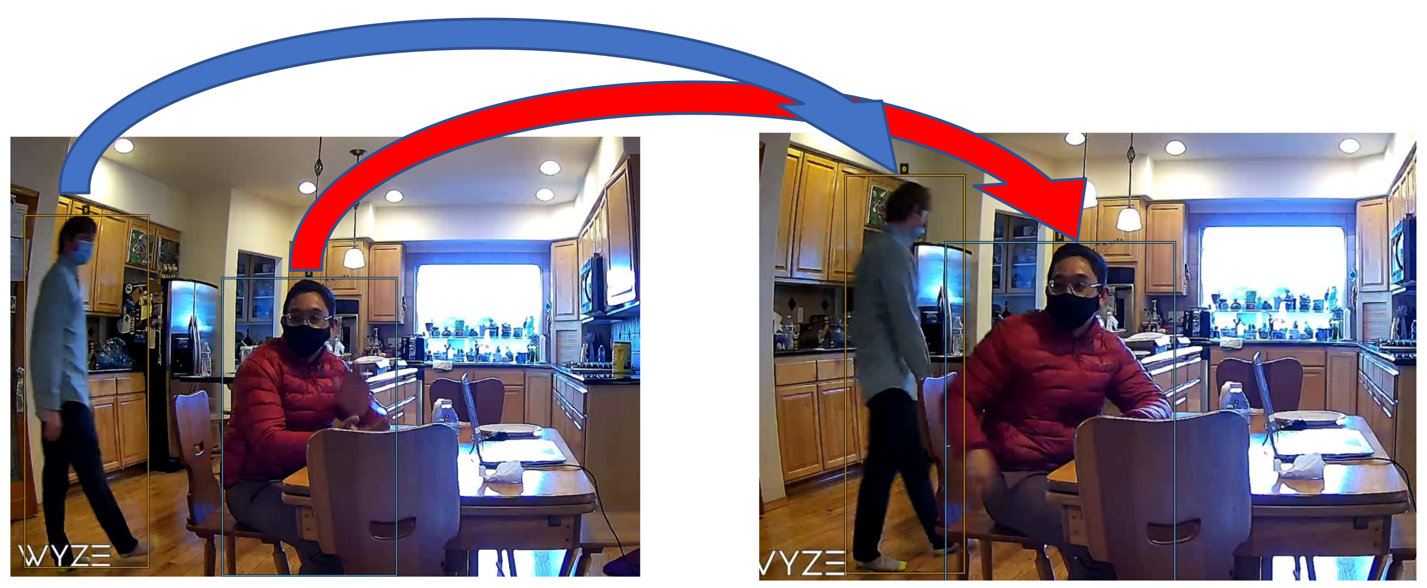
Just like the titular Harry Potter Marauder's map we want to provide users with the means to follow any individual as they wander around a mapped space. Our Android app and machine learning model work together with Wyze cameras to provide a Multi-Target Multi-Camera Tracking (MTMCT) system for keeping your private space safe and secure. Some example use cases include tracking at risk individuals in a long-term care home or detecting unauthorized intrusions. Our goal is to handle at least 4 different cameras at once with minimum 4 different people.

Background & Models

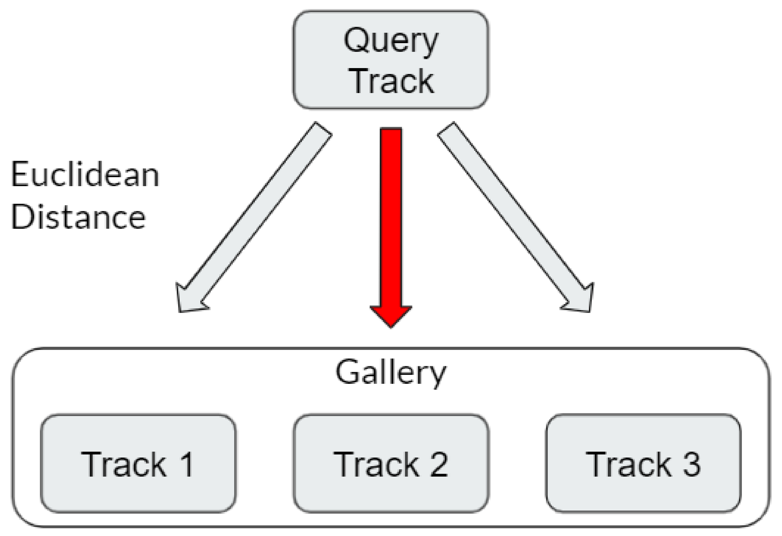
The first step towards achieving MTMCT is by checking each camera frame for persons. When a person is identified a bounding box can be drawn around the detected person/object. In our case we are using Detectron2 which is a Faster RCNN created by Facebook and pretrained on the COCO dataset. Detectron2 outputs the set of bounding boxes for each detection.



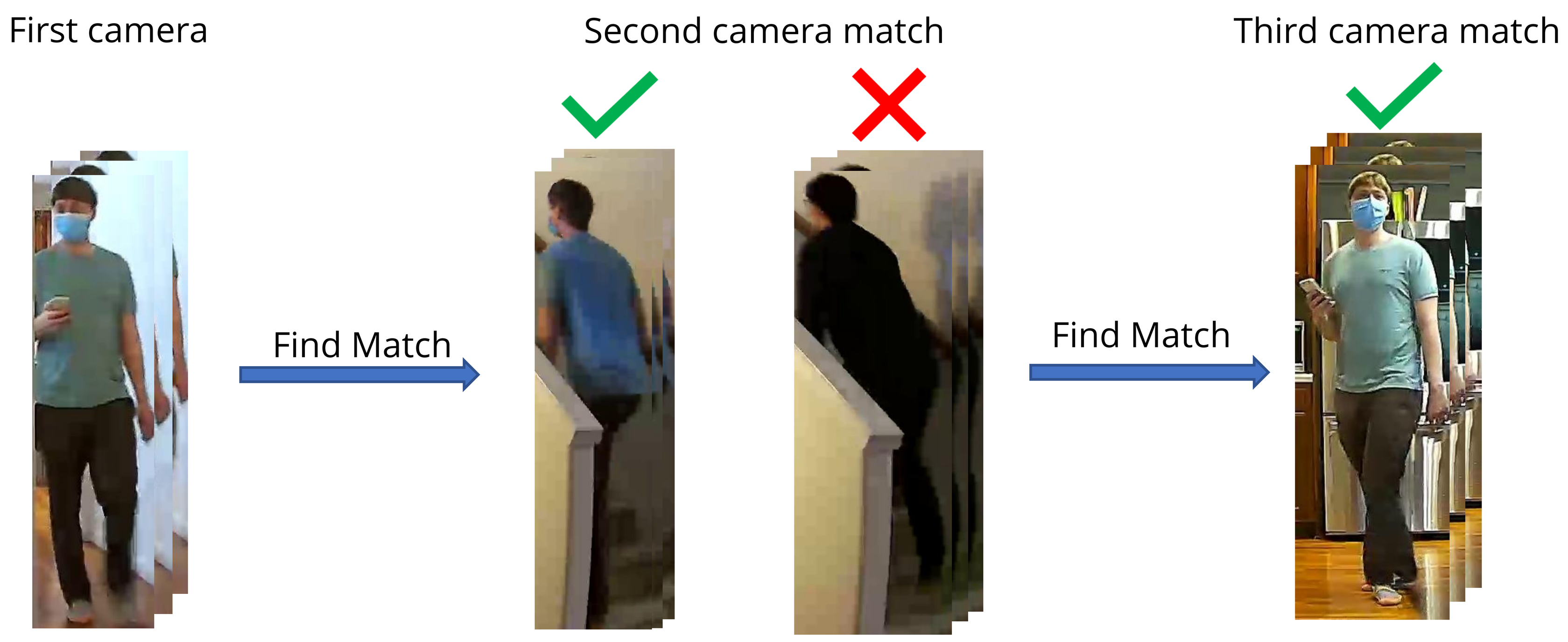
The next step to achieve MTMCT is through single camera tracking. We are using DeepSort which uses a combination of feature embedding and a Kalman filter to generate a predicted trajectory for each objects next video frame position. If a close match between the next frame bounding box and the DeepSort predicted position then they are associated together as the same individual.



After the bounding boxes have been associated together into tracks we can extract features from these tracks and compare them to a new track in a different camera. Person Re-Identification (ReID) can identify same person matches between cameras using Euclidean distance. We use OSNet, which uses a lite 3x3 CNN layer and aggregation gate, to extract features for comparison.



Data Pipeline



Our pipeline takes raw video taken by WYZE cameras, detects people, and then creates a global trajectory for each detected person as they travel between cameras. The global trajectories and camera streams are provided to the user via our Android app. The above example shows how one camera ID is matched to other cameras to produce a global travel path.

Dataset & Training

We created our own custom dataset using 3 separate locations with 21 different cameras and 15 unique IDs. This allowed us to focus our dataset on areas we needed to improve our model. To improve the realism we have also developed different scenarios such as entering the front door, taking off your shoes and then going to sit on the couch. The 3 separate locations are summarized below.

Location	IDs	Cameras	Footage/camera	Scenarios
William's	0,1,2,3,4,5	0,1,2,3,4,5,6	366 seconds	Multi-cam, clothes switch
Kevin's	0,3,4,5,6,7,8	7,8,9,10,11,12,13	714 seconds	Single-cam, leave and return
Michael's	0,3,9,10,11,12,13,14	14,15,16,17,18,19,20	600 seconds	Multi-cam same clothing, real skits

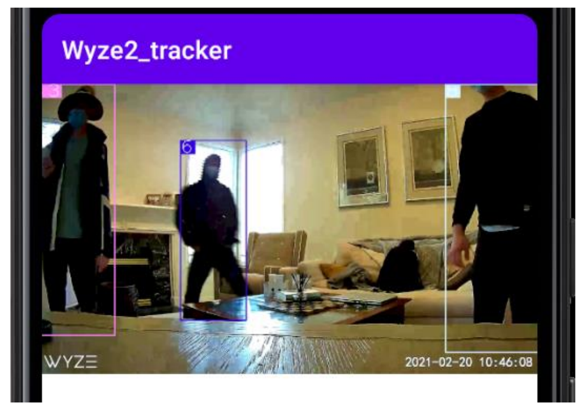
Our ReID model was first pretrained on ImageNet but we performed retraining on our own dataset to improve the accuracy. The model was trained for 50 epochs using this dataset. An example training query and the matches is shown to the right.

Set	IDs	Cameras	Tracklets	Cropped Frames	Time at 15fps
Training	0,1,2,3,4,5	0,1,2	133	14,175	945 seconds
Testing	0,1,2,3,4,5	3,4,5,6	120	13,228	882 seconds

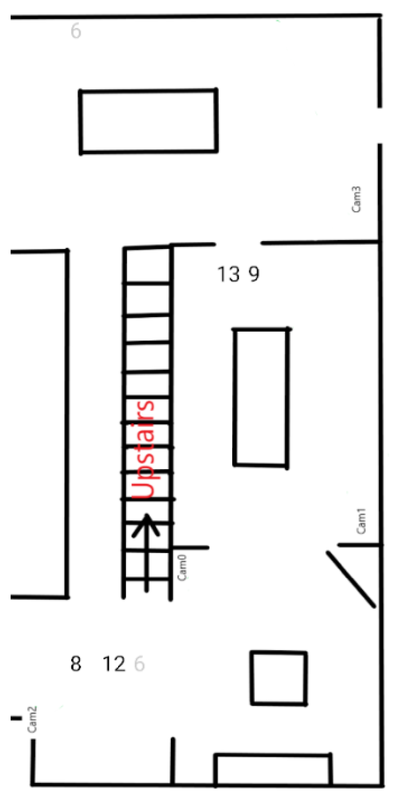
App

Our app is made up of 5 main sections. The list of cameras is clickable and takes the user to the live stream. The live stream shows the camera view and the identities. The last part is the live map which shows where each identity is.

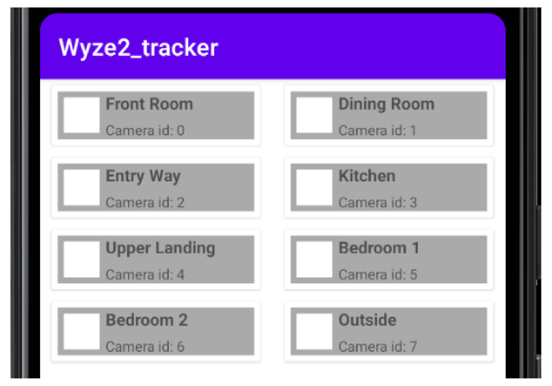
Live Video Stream



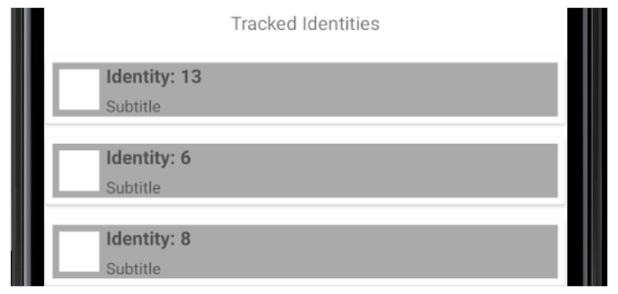
Live Marauder's Map



List of Cameras



Live Identity List



Results & Conclusions

The results from retraining our ReID model with our data for 50 epochs can be found below. The retraining was successful and resulted in a large improvement for our testing set.

Training	Mean Average Precision	Rank-1	Rank-5
Pre-trained	24.9%	9.2%	51.7%
Re-trained	88.3%	90.8%	96.7%

The results from retraining our tracking model with our data for 30 epochs can be found below. The retraining was successful and resulted in a large improvement for our testing set.

Training	IDF1	MOTA
Pre-trained	58.8%	64.4%
Re-trained	61.3%	75.4%

MTMCT Results for full pipeline using ReID in combination with pre-trained DeepSort tracking.

Training	IDF1	MOTA
Pre-trained	65.3%	73.4%

Future Work, References, and Acknowledgments

- Collect and label more data with more IDs/Locations.
- Host data pipeline online.
- Live data streaming to the app.
- Examine Joint Detection Embedding with FairMOT.

Faculty: Prof. Payman Arabshahi, Dr. Hung-Min Hsu
TAs: Haobo Zhang, Shruti Misra, Daniel King

[1] Wu, 2019 Detectron2 [Source Code]. https://github.com/facebookresearch/detectron2
[2] Wojke, Nicolai, Bewley, Alex, and Paulus, Dietrich. 2017. "Simple Online and Realtime Tracking with a Deep Association Metric.", arXiv:1703.07402
[3] Zhou, K, Yang, Y, Cavallaro, A, & Xiang, T, 2019, 'Omni-Scale Feature Learning for Person Re-Identification', *arXiv*, 1905.00953

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISOR: DR. HUNG-MIN HSU
SPONSOR: Wyze Labs, Inc.