香港中文大學
The Chinese University of Hong Kong

Efficient Video Analytics

Researcher: YAN Yuhang Supervisor: Prof. Eric Chi Lik LO
CUHK UG Summer Research Internship 2023

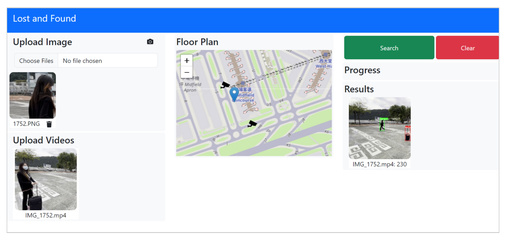
Introduction

Our ultimate goal was to create an accurate and efficient lost-and-found system for Hong Kong Airport. By inputting a linguistic or pictorial description of the owner and the lost item, our system can quickly find the moment the item was lost. Our preliminary expectations before embarking on the Summer Research were threefold:

- Design a user-friendly front-end web UI.
- Implement the fundamental functionality of the system.
- Enhance the performance of our system as much as possible.

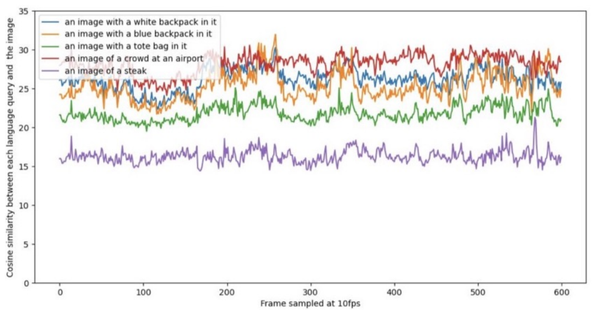
Basic Implementation

We started by building the user interface and using the "Center-Distance Method" to detect the "LOST" status of an item.



Detection Methods and Models

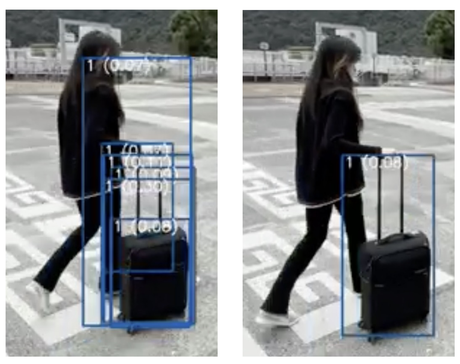
- **CV Models**
  YOLOv5 + Deep SORT
  YOLOv5 + CLIP
  OWL-ViT
- **"LOST" Detection**
  Center-Distance
  Detecting Co-occurrence



Further Approaching

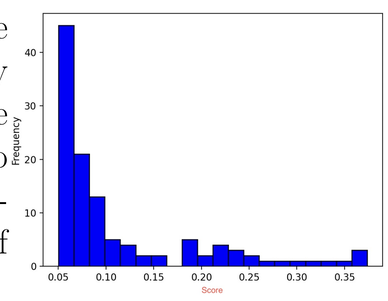
- **Non-Maximum Suppression**

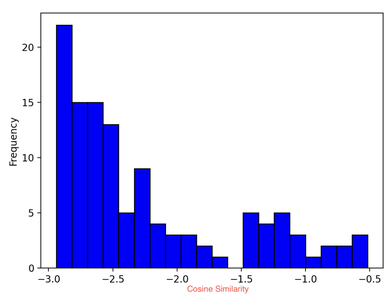
We found there was a large amount of overlap in the objects recognized. We integrated the Non-Maximum Suppression algorithm into our model, which simplified multiple recognition bounding boxes caused by the same object into one.



- **Scores and Cosine Similarity**

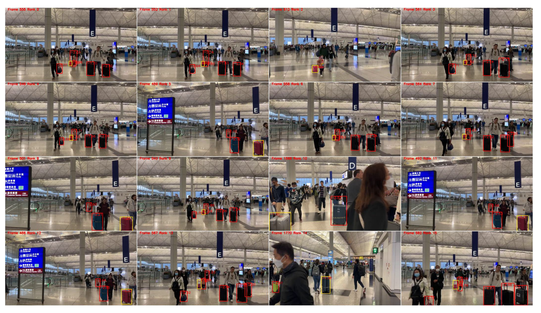
We use the Cosine Similarity directly as the score for the prediction result to widen the gap between the scores of different objects.





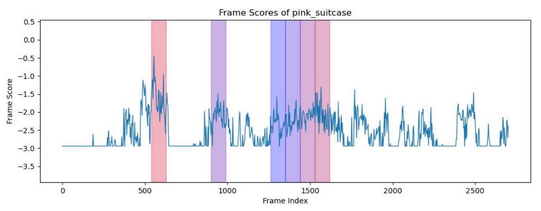
- **Top-K Selection**

We integrated the Top-K Selection algorithm into the model to solve the problem that the False Positive rate of our model is to high. The yellow boxes are the Top-16 of the image detecting result of "pink box".



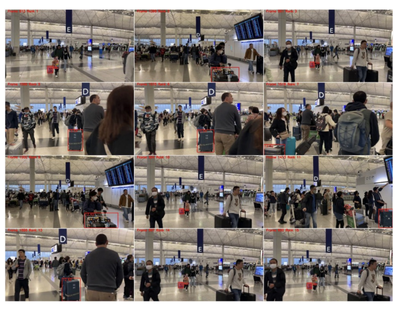
- **Chunk Extraction**

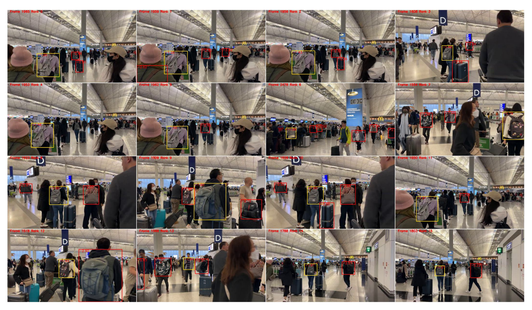
We developed new algorithms that rank and output chunks instead of frames. The redder the colour, the higher the score of the chunk.



Result

Now, we can run either an image or language query on a video and pick out frames or chunks containing objects with Top-K scores from the result.
These are two examples: image query for the pink luggage and language query for "black and white striped backpack".





Future Work

- Utilize the "Detecting Co-occurrence" method.
- Incorporate Everest, a video analytics system for Top-K queries, into our model.

References

[1] A. Bansal, K. Sikka, G. Sharma, R. Chellappa, and A. Divakaran. Zero-shot object detection. In *Proceedings of the European conference on computer vision (ECCV)*, pages 384–400, 2018.

[2] Facebook Inc. Create react app. `https://github.com/facebook/create-react-app`. Accessed: 2023-06-15.

[3] G. Jocher and contributors. Yolov5. `https://github.com/ultralytics/yolov5`, 2020.

[4] Z. Lai, C. Han, C. Liu, P. Zhang, E. Lo, and B. Kao. Top-k deep video analytics: A probabilistic approach. In *Proceedings of the 2021 International Conference on Management of Data*, SIGMOD '21, page 1037–1050, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383431. doi: 10.1145/3448016.3452786. URL `https://doi.org/10.1145/3448016.3452786`.

[5] M. Minderer, A. Gritsenko, A. Stone, M. Neumann, D. Weissenborn, A. Dosovitskiy, A. Mahendran, A. Arnab, M. Dehghani, Z. Shen, X. Wang, X. Zhai, T. Kipf, and N. Houlsby. Simple open-vocabulary object detection with vision transformers, 2022.

[6] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever. Learning transferable visual models from natural language supervision. *CVPR*, abs/2103.00020, 2021. URL `https://arxiv.org/abs/2103.00020`.

[7] F. Romero, C. Winston, J. Hauswald, M. Zaharia, and C. Kozyrakis. Zelda: Video analytics using vision-language models, 2023.

[8] N. Wojke, A. Bewley, and D. Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3645–3649. IEEE, 2017.

Contact Information

- YAN Yuhang
- yhyan2@cse.cuhk.edu.hk
- Dept. Computer Science & Engineering
- Supervisor: Prof. Eric Chi Lik LO
- Dept. Computer Science & Engineering