Finding Defects on Glass Crystal

Students: Junchen ZHOU (City University of Hong Kong), Yuhan XUE (Chinese University of Hong Kong)
Mentors: Dr. Kwai Wong, Dr. Raymond Chan

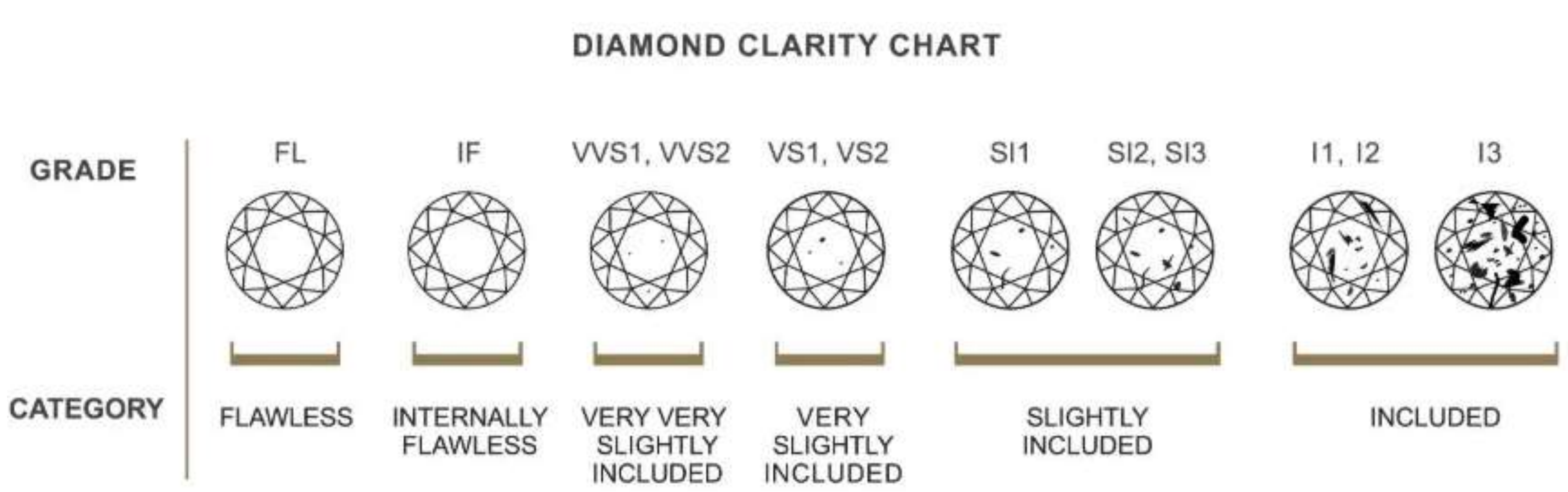
Introduction

Class crystals are widely used in various industries, including electronics, optics, and jewelry, which ensures the necessity to develop an image and video analysis system for identifying defects present on glass crystal surfaces.

Up to the present, significant progress has been made in the classification of defects in glass crystal images through the training of CNN models. However, when dealing with diamond video data, numerous challenges persist.

This project introduces an advanced image processing and video tracking system that utilizes machine learning algorithms, computer vision techniques, and deep neural networks to efficiently categorize defects on glass crystals throughout videos.

Figure 1. Glass crystal and clarity chart.



Datasets

There are 3 parts about the dataset in total.
Part1: Images
Images about the glass crystal with the defects
Part2: Videos
There are 400 frames in one video, which is used to record the track of the glass crystal.
Part3: XML files
The XML files include the detailed position information in the glass crystal images.

Aims

Label the defects on the video frames.
Track the defects in the video using Neural Network.

Methods

We utilize an adapted unsupervised deep tracking method to track inclusions in the video, employing the integration of unsupervised learning into the Siamese based correlation filter framework. This lightweight network architecture allows simultaneous learning of convolutional features and correlation tracking.

Discriminative Correlation Filters
DCF is a method that employs correlation between a target template and candidate patches to efficiently track objects in video sequences.

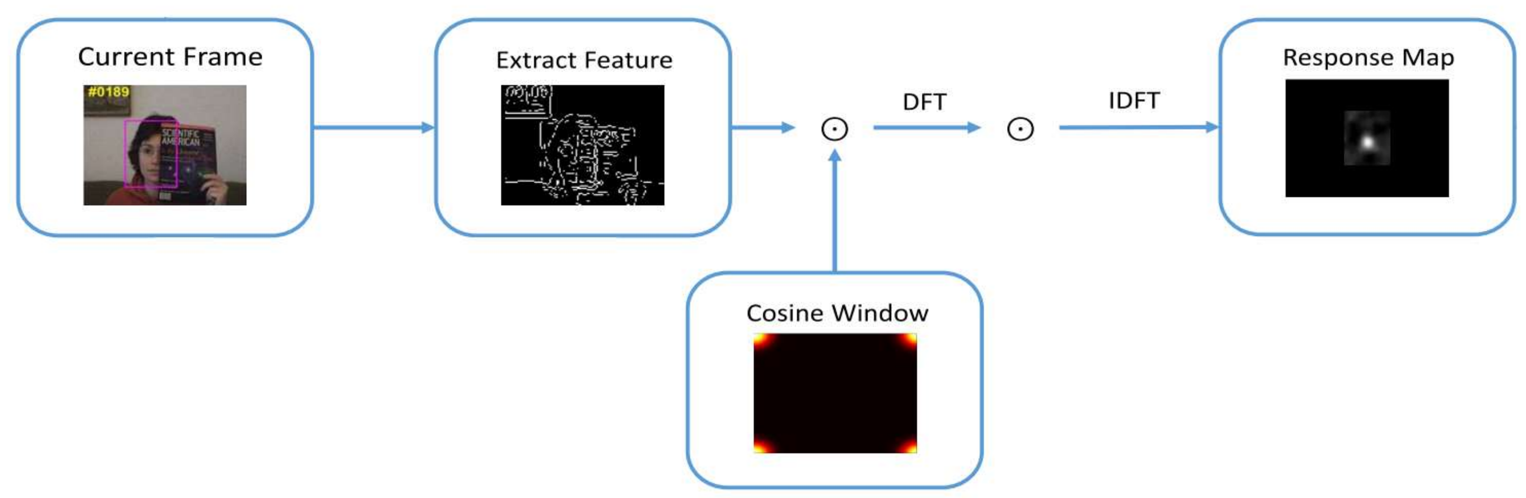


Figure 1. General framework for correlation-filter-based object tracking.

Unsupervised Learning Prototype
The unsupervised tracking method utilizes a Siamese correlation filter backbone learned through forward and backward tracking. An overview of Unsupervised Learning Motivation is shown below:

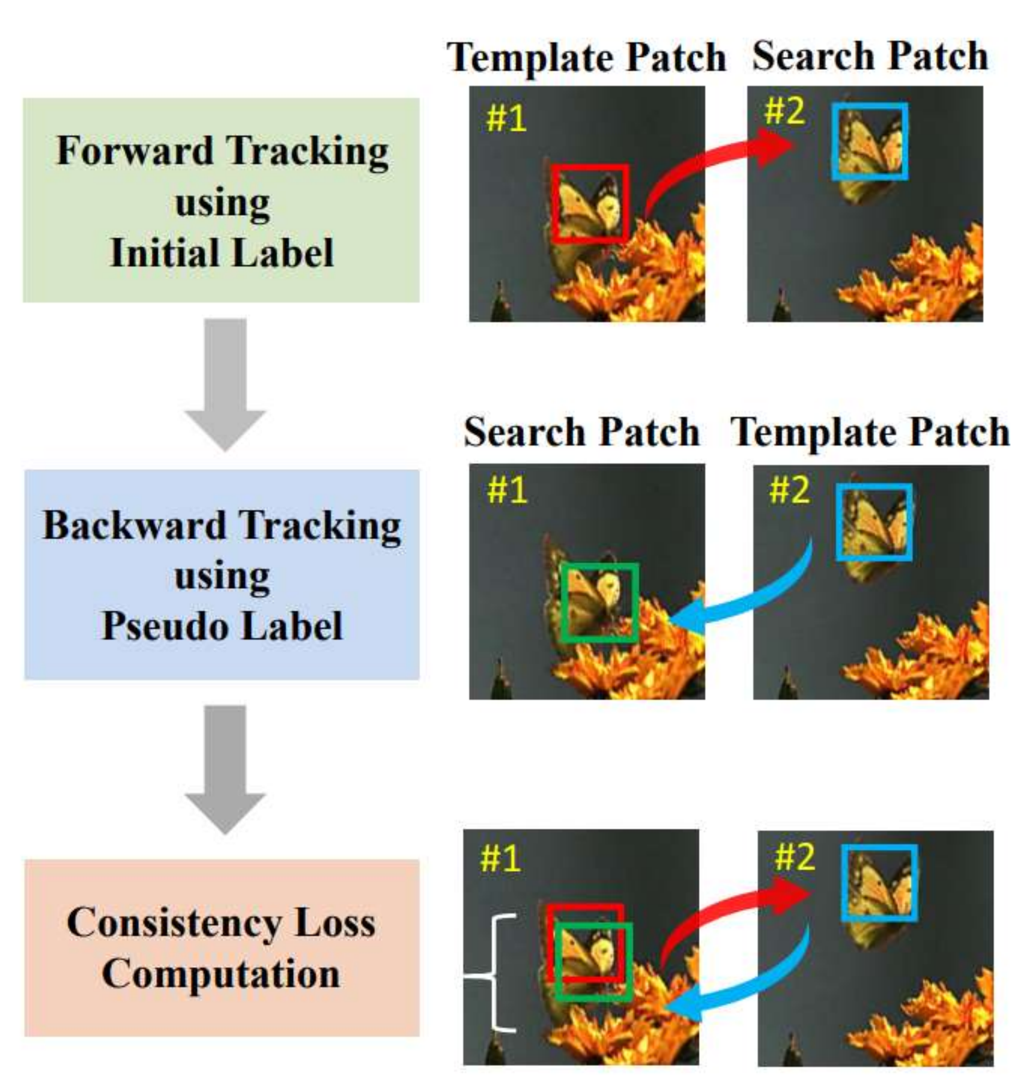


Figure 2. An overview of unsupervised deep tracking.

Step 1: Classification and Segmentation Model
We separate the video into each frame first. And then we convert processing video into processing image. The detailed workflow is as below:

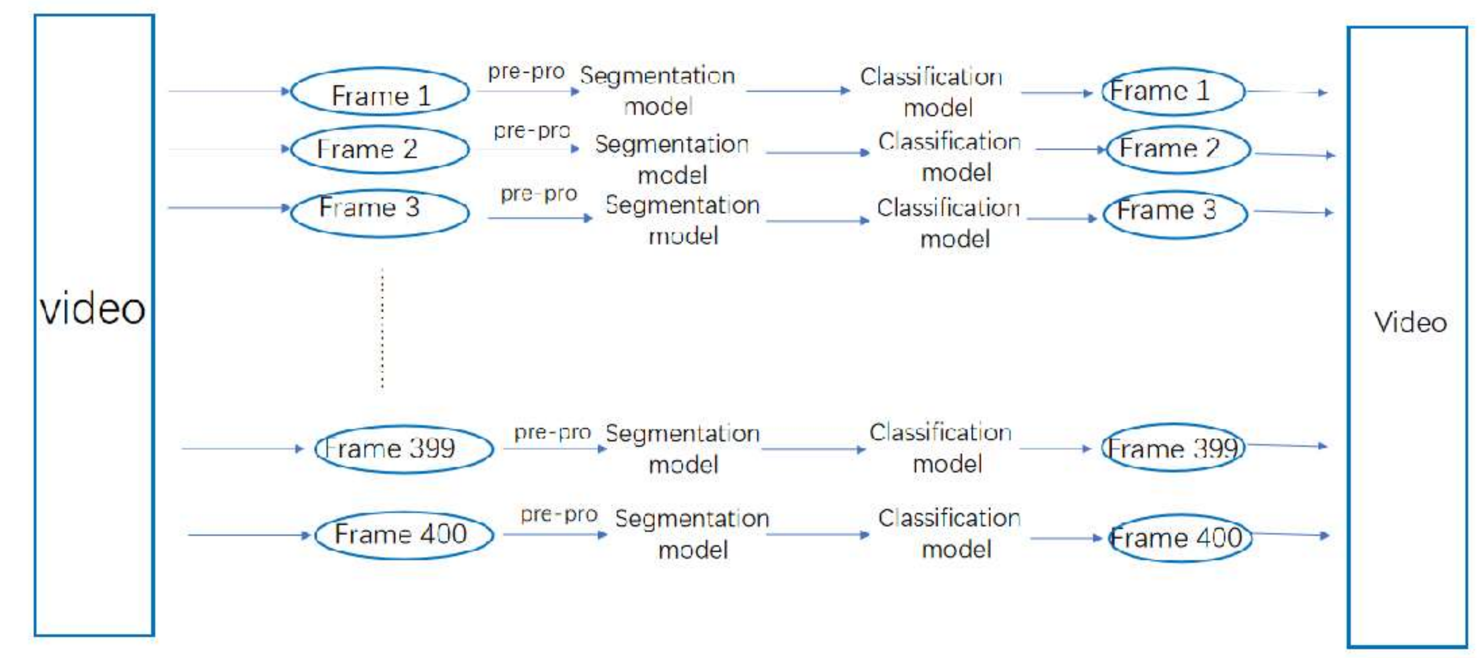


Table 1. Video processing workflow.

Step 2: Use YOLO to track the video
We feed the images and txt files with the accurate position information of the tracks into the model for training. After training the model, we put the video as the input for tacking the video.

Future Work

- Improve classification accuracy
- Conduct batch tracking processing for videos

Acknowledgements

This project was sponsored by the National Science Foundation through Research Experience for Undergraduates (REU) award, with additional support from the Joint Institute of Computational Sciences at University of Tennessee Knoxville. We would like to express our gratitude to Prof. Raymond Chan, Dr. Kwai Wong, Eddie and DAI Lingjia for their invaluable support and guidance throughout the project.

References

Wang, N., Song, Y., Ma, C., Zhou, W., Liu, W., & Li, H. (2019). Unsupervised deep tracking. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
Du, S., & Wang, S. (2022). An overview of correlation-filter-based object tracking. IEEE Transactions on Computational Social Systems, 9(1), 18–31.
Ultralytics. (n.d.). Home. Ultralytics YOLOv8 Docs.
Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-net: Going deeper with nested U-structure for salient object detection. Pattern Recognition, 106, 107404.