Identify Defects In Diamond

Students: PAK Ka Yee (City University of Hong Kong) , JIANG Chenxin (Chinese University of Hong Kong)
Mentor: Dr. Kwai Wong

Introduction

Diamond crystals are in lattice crystal structure, and they are often found with defects or inclusions, such as pinpoint, feather, internal graining needle, which will affect the commercial value of the crystals. Many efforts have been made to identify defect categories and pixel-level defection positions. However, the video segmentation on crystals has not to be studied to identify the defects on crystals yet.

Background study: An image segmentation has been developed to identify the defects on a crystal. Nevertheless, some tiny defects and reflections of the crystal became another two significant challenges to be solved.

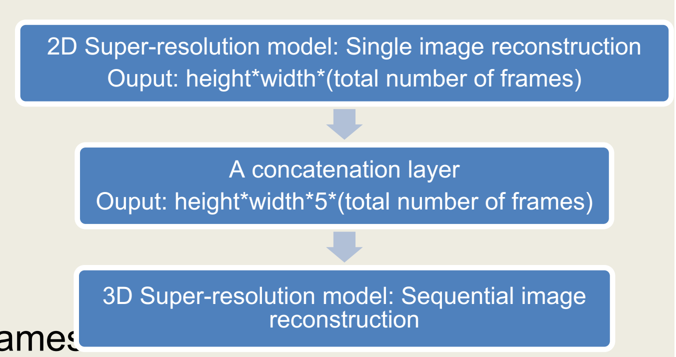
Objectives: When the crystal is rotating, the position of the defect will move along the rotation, and the reflection will change its reflect directions by the light. Therefore, this research project produces two approaches to solved the problems by video segmentation.

- For tiny defects: Super-resolution (The first model)
- For reflections: Trajectory Classification & Video Segmentation (The second model)

Methods and Experiment and 1st model

All 400 frames have been extracted and blurred before processing.

Super-resolution method: A concatenation layer is added between a 2D super-resolution model and a 3D super-resolution model to process the data from single frames into packages consisting of 5 consecutive frames



Model: the structure of 2D super-resolution

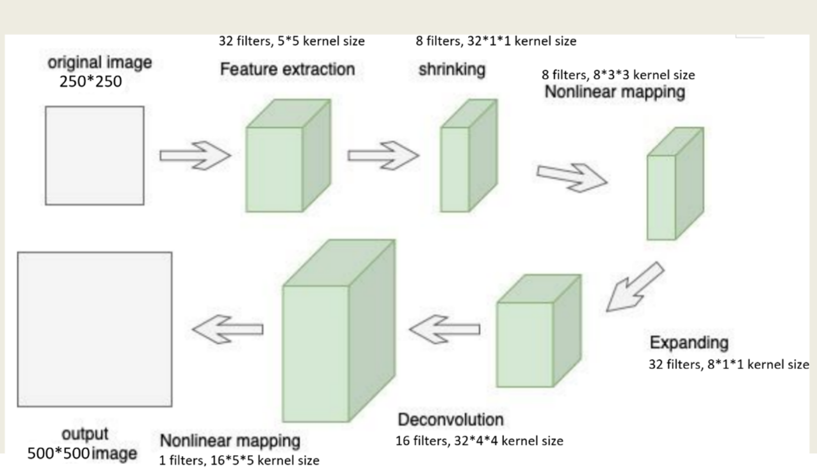


Figure 6: The structure of 2D super-resolution model

Model: 3D super-resolution

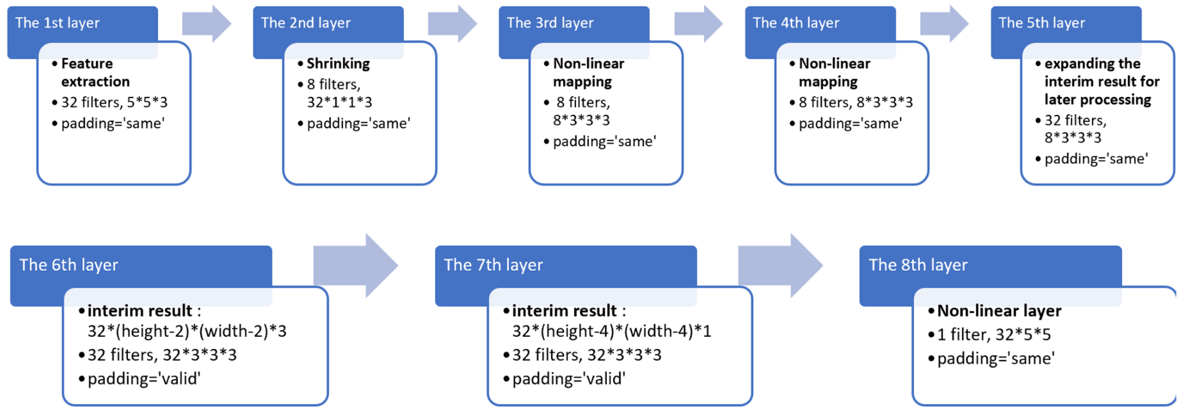


Figure 7: The structure of 3D super-resolution model

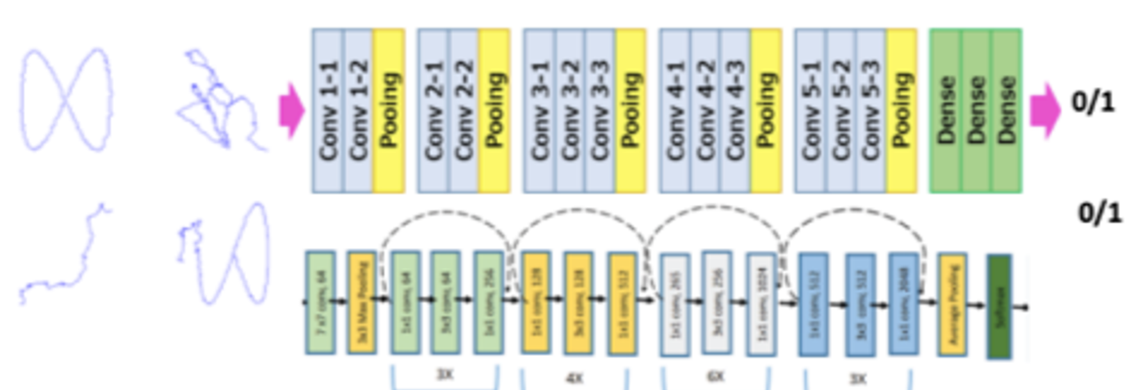
Methods and Experiment and 2nd model

Dealing with the Reflections

To distinguish the reflections against true defects, we tried to track the objects of our interests throughout the 400 frames.

Trajectory classification

First we analyze the motion of one defect or reflection. Tracking the object by CSRT tracker, and record the movement trajectory path, we discovered that the movement path of the true defect is usually an "infinity symbol" throughout the whole video, because true defects always move along with the diamond, but the trajectory path of artificial defect presents no regular pattern. With this regard, it's possible to classify those two based on different moving patterns with the help of CNN.



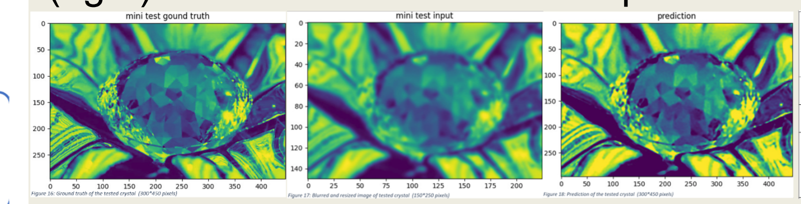
Inherently, the trajectory classification problem is a binary image classification problem, in which we can utilize mature models (VGG-16/ResNet-50) with transfer learning.

Video Segmentation

Another technique to approach the shape of the moving defects is video segmentation. We applied One-Shot Video Object Segmentation, a semi-supervised segmentation method. Three networks are integrated, in which the base network is a pre-trained VGG for image labeling, the parent network on the training set is to separate the background and foreground, and the test network is to segment particular object after fine-tuning on one specific example. For our multi-instance segmentation problem aiming to divide the true defects and reflections, we perform instance isolation and train them separately.

Performance (1st & 2nd model)

The below three crystal images are the ground truth (left), input (middle) and the prediction after passing through the model (right). Total trained number of epoch is 100 .



Epochs	Train Set (80% of all train images)			Validation Set (20% of all train pictures)		
	MSE	SSIM	PSNR	MSE	SSIM	PSNR
1	3509.5798	0.4599	14.9566	921.6595	0.5441	18.4867
100	293.2773	0.7732	24.1640	215.0823	0.7919	24.9263
Performance Compares the 1 epoch with the 100 epoch	-0.916435	0.681235	0.615607825	-0.76664	0.455431	0.348337

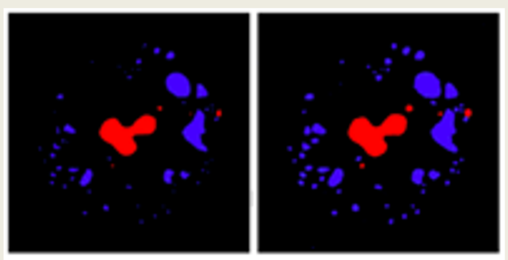
From the table above:the (training, validation) sets: MSE improves (91.64%, 76.66%), SSIM improves (68.12%, 45.54%), PSNR improves (61.56%, 34.83%).

Another approach: Results from a software Topaz Gigapixel AI.

Original crystal(450*300 pixels) After super-resolution(2700*1800 pixels) Original crystal(450*300). After super-resolution(2700*1800)

For trajectory classification, after 30 epochs, the accuracy on training and validation set converges to 77~78%, and the test accuracy is 75.78125%.

For video segmentation, since we don't frame, it is impossible to calculate the loss. But we can compare the ground truth and the prediction. of the 1st frame. Generally, the predicted region is bigger by observation.



Comparison of the ground truth (left) and the prediction (right) for the 1st frame.

Future work

1. The super-resolution model can consider standard upsampling to avoid any damages of the resolution.
2. Integrate the video segmentation and object detection method together.

Acknowledgements

An acknowledgement for Cheuk Hin HO provides the modified super-resolution model. 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 the University of Tennessee Knoxville. This project used allocations from the Extreme Science and Engineering Discovery Environment (XSEDE), which the National Science Foundation supports.

References

Crystal video: https://www.youtube.com/watch?v=VVEgcZO-zVY&ab_channel=RigidGems
Caelles, S., Maninis, K.-K., Pont-Tuset, J., Leal-Taixe, L., Cremers, D., & Van Gool, L. (2017). One-Shot video object segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)