MATH 6380P Project 3: Smartphones Glass Defects Detection

Jimmy WU, Zhaoyu SU

Dept. of Civil & Environmental Engineering, HKUST

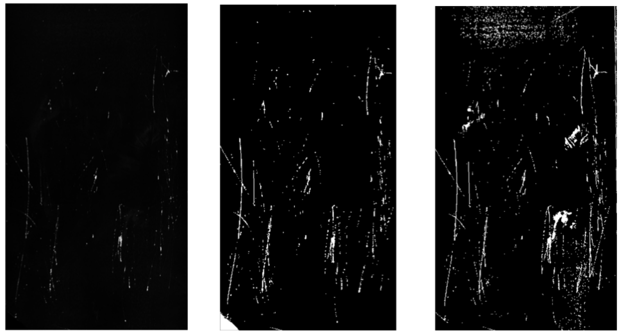
Objectives & Challenges

In this work, two deep-learning based methods for smart phones glass defects detection from ultra-high resolution scanning images are proposed and compared.

The main challenge for this work is that since no pre-made labels are available, the detection has to be executed in an unsupervised pattern. Due to the same reason, final results are compared by visualization interpretation.

Dataset Introduction

Thanks to Muhammad [1], we are able to get access to the dataset for this work. The dataset comprises of 60 scanning images of smartphones screens acquired from 16K line camera. Each image is in resolution of 24576 x 16384 with grey scale. Each pixel is corresponding to ~0.5 micrometer in reality, while the objective is to detect defects whose size is greater than 10 micrometer.



Raw Input | Result from DCNN | Result from WSN

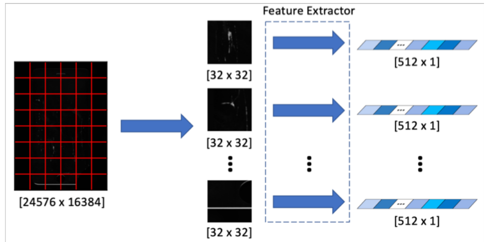
Result Visualization

Feature Extraction

The proposed methodology basically follows such an architecture: First, the patch-wise features are extracted via a feature extractor, which can either be a convolutional neural network (CNN) or wavelet scattering network. Next, the parameterized features for each patch are fed into a trainable dimension reduction method (U-Map), followed by a k-NN clustering. Finally, the classified patches are assembled back into the images, and a threshold is set for the final binary results.

Wavelet Scattering Net

The wavelet scattering net (WSN) [2] can be chosen as the feature exactor, just like the one used in project 1. The WSN takes 32x32x1 image as the input, and gives out 2x2x128 output features. The features are subsequently flatten into 1D.



DCNN Feature Extractor

As there is no suitable pre-trained DCNN models available for 32x32x1 grey scale image (although some DCNN models pre-trained on Cifar-10 may help, but similarity is a big issue for our project), therefore, we train a modified U-Net [3] (no more shortcut connections) based self-auto-encoder to be our feature extractor.

During the training phase, the modified U-Net is forced to learn the mapping from the raw input to itself, and the features at the lowest feature spatial space are reserved as the final feature representations. Similarly, the output is also flatten into 1-D

During the U-Net training phase, 10,000 32 x 32 patches are randomly cropped from the raw images, and then serve as both the training input and output. During the inference, the raw input is again cropped into 32 x 32 patched in row major, and the output feature for each path is a 1-D vector with size of 512.

Conclusion

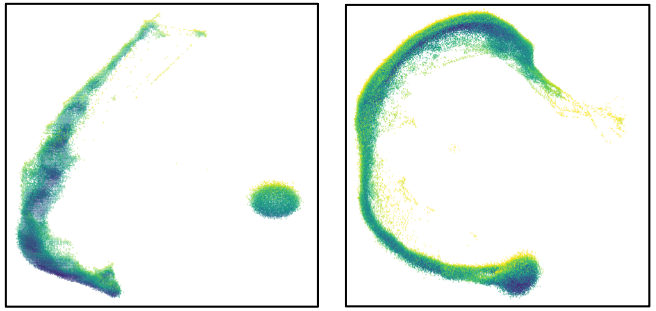
In this work, two screen defect detection methods are proposed and compared, whose major difference is the feature extractor. Through visual interpretation, the results from DCNN model is more better than the WSN's, especially for the false positive errors. However, the classification for different categories of detects are unable to be performed due to the lack of training examples and ground truth.

Dimension Reduction & Clustering

In order to perform a unsupervised classification, classical clustering method such as k-NN are applied here. As k-NN is no longer valid for high-dimension input data, the dimension reduction method is necessary to map the high-dimension features into low-dimension. Traditional t-SNE is not adopted due to its computation intensive and non-deterministic properties. Instead, a recently proposed method U-Map [4] is therefore applied.

The 512-D features are converted into 3-D representations, and a binary clustering method is used to group each patch into defects and non-defects areas.

U-Map Visualization



U-Map result from DCNN | U-Map result from WSN

Reference :

1. mumbhutta@connect.ust.hk
2. Oyallon, Edouard, Eugene Belilovsky, and Sergey Zagoruyko. "Scaling the scattering transform: Deep hybrid networks." In *International Conference on Computer Vision (ICCV)*. 2017.
3. McInnes, L., & Healy, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
4. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.