Steel Defect Detection

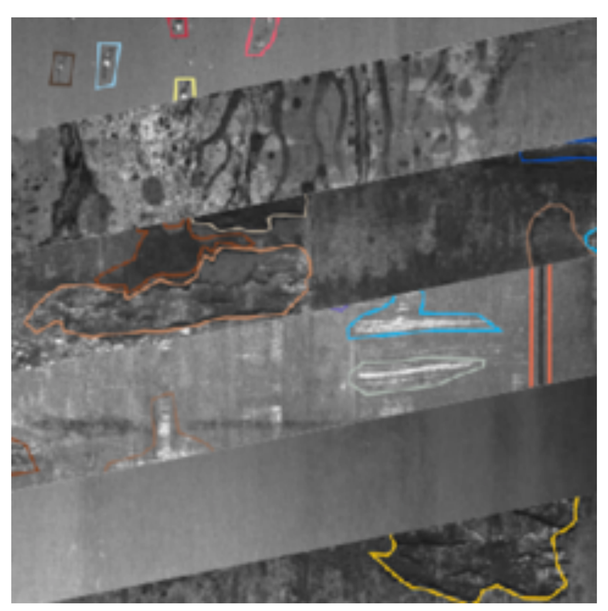
Yinan Tian @ STANFORD.EDU

Overview

Steel is one of the most important building materials of modern times with complicated production process which is known to be delicate. Early and effective defect detection is very important in terms of quality control and cost saving.

In this project, deep convolution neural network (CNN) is used to detect 4 specific types of steel defect. Furthermore, U-NET structure [1] consisting both down-sampling and up-sampling layers is deployed to localize and segment the defect region.

Data



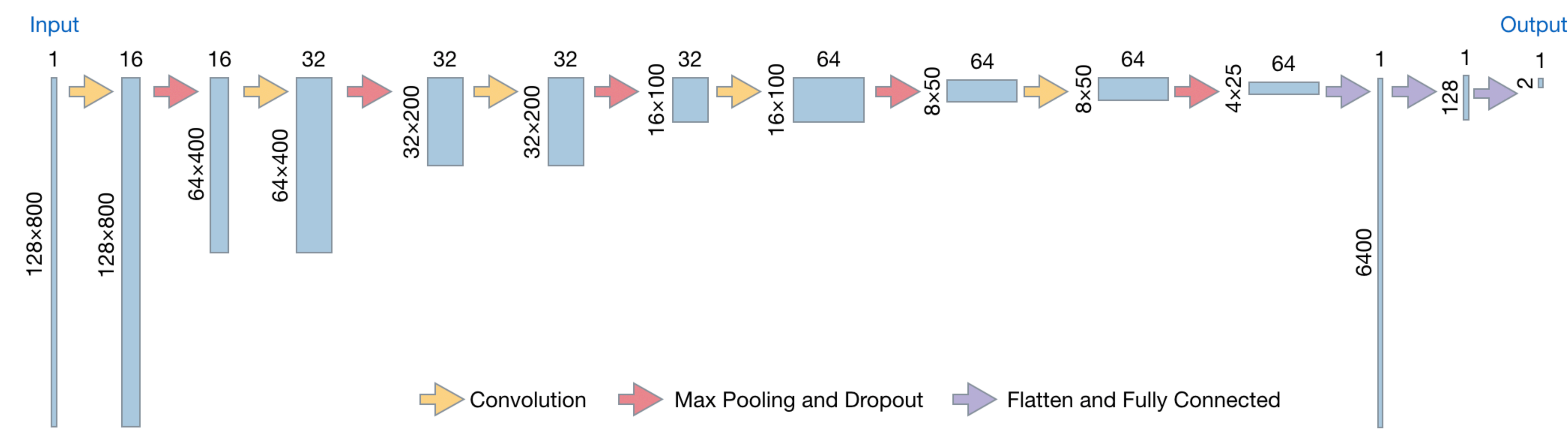
The image set is collected on thousands of steel samples. Example defects are shown as on the left. The raw images went through pre-processing and augmentation steps before being input to the CNN model.

- **Pre-processing**: Due to the fact that each image is subject to different lighting condition when the image is taken, each image is normalized by the global mean and std of itself.
- **Augmentation**: The biggest constraint of the data set is its imbalance. Out of the entire training and validation dataset, 41% of the samples are identified as class 3, while the percentages for class 1, 2, 4 are 7.1%, 2.0% and 6.4% respectively. To prevent overfitting on class 3 and also provide more data to other classes, augmentation is done on class 1, 2 and 4 through random flipping and shift so that each class has the same number of data points.

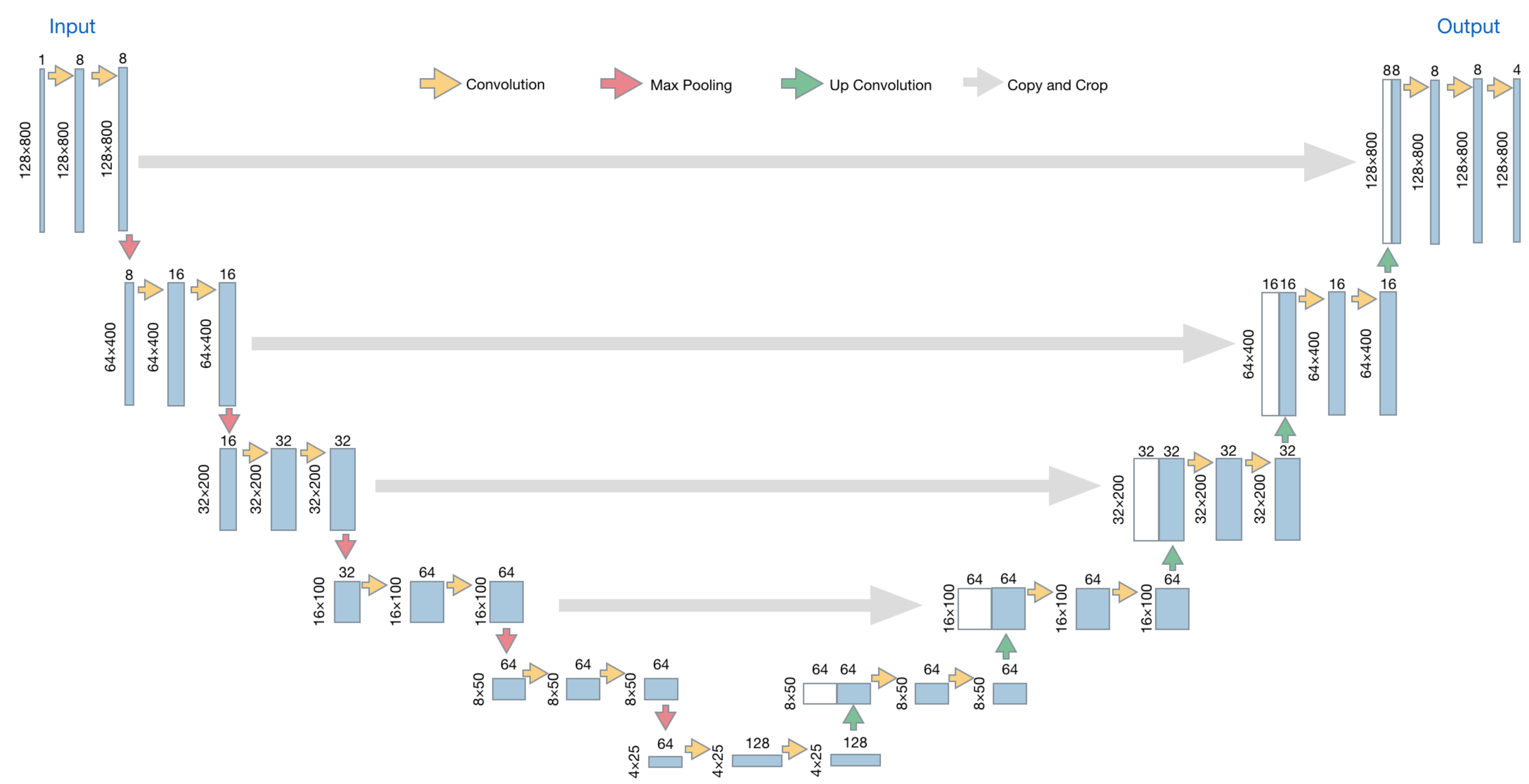
Model

Due to the complexity of this problem, it is further broken down into two steps:

- **Binary Classification**: A traditional CNN as below with convolution layer, max pooling layer, dropout layer and finally fully connected layer is used. This network will differentiate non-defective images from defective ones.



- **Segmentation:** U-NET supplements the contracting network by successive up-sampling layers. Within those layers, features calculated through contracting network is combined to re-construct the current image to the original image size. Detailed architecture is illustrated as below. The output of the network 4x 2D masks for 4x defect classes respectively.



- **Loss function:** Different loss functions including binary cross entropy, dice loss and Tversky loss are compared based on validation result. At the end, binary cross entropy is used in the binary classification step while U-NET uses the combination of binary cross entropy and dice loss as the target.

Video link: https://drive.google.com/file/d/1llLGHKNWWowwTBt4v-VFJ1Y3nTpU5O2n/view?usp=sharing

Result & Analysis

- Dice coefficient is used as the evaluation metric in the problem. Combining both steps, the final accuracy on the test set is 81%.
- Confusion matrix based on the defective images only is calculated as below.

Truth \ Prediction	Class 1	Class 2	Class 3	Class 4
Class 1	0.18%	0.03%	0.70%	1.50%
Class 2	0.18%	10.58%	0.35%	1.39%
Class 3	0.05%	0.63%	58.06%	1.16%
Class 4	0.03%	2.37%	0.55%	42.47%

- Compared to the detection performance in class 3 and class 4, the model is not as capable in predicting class 1 and class 2. However, as the number of images for class 1 and 2 is small in both train set and test set, the inaccuracy does not impact too much on the final accuracy.
- In addition, the table also indicates slight overfitting on class 4.

Future Work

- Explore the possibility of transfer learning
- Fine tune to resolve the overfitting on class 4
- Thorough case study on the miss detection on class 1 and class 2 and fine tuning

Reference

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: International Conference on Medical image computing and computer-assisted intervention. Springer. 2015, pp. 234–241.