Steel Defect Detection

Problem Description

"Severstal is leading the charge on efficient steel mining and production. They believe the future of metallurgy requires development across the economic, ecological, and social aspects of the industry—and they take corporate responsibility seriously. The company recently created the country's largest industrial data lake, with petabytes of data that were previously discarded. Severstal is now looking to machine learning to improve automation, increase efficiency, and maintain high quality in their production."[1]

How can I build a model that detects the defects in steel images? Therefore Severstal can improve its product quality and efficiency.

Target Client

Severstal

Data

The data is provided by the company itself as a kaggle competition. It consists of test images, train images, submission sample and labels. A total of 1.58 GB of data can be found on the competition site.

Approach

To solve this problem I will be using neural networks. So far CNN's are most advised for computer vision tasks. However, I'll be researching for some other architectures if there is a competitor to CNN. Additionally, I'll be trying to work on explainability of the model as well.

Deliverables

The final submission of the project will contain:

1. This proposal paper
2. A presentation of the project
3. Source codes for data wrangling, EDA, and final model build.
4. Exported model
5. Report of the model

[1] Severstal kaggle competition overview, Severstal: Steel Defect Detection, 04/04/2021