Shankar Kumar

Proposed Architecture

Data Preprocessing and Augmentation

[N, 2048, 3]
Image vectors

Resnet101 Feature Extractor

[N, 2048]
feature vectors

Simple Classification CNN

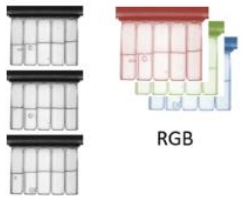
Positive Class

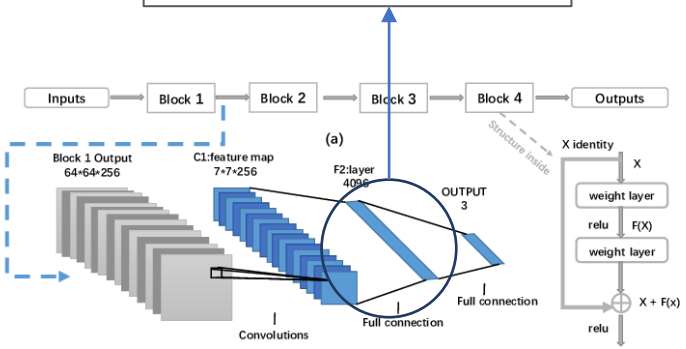
Binary Classification Output

Negative Class

SVM Classifier

Ternary Classification Output

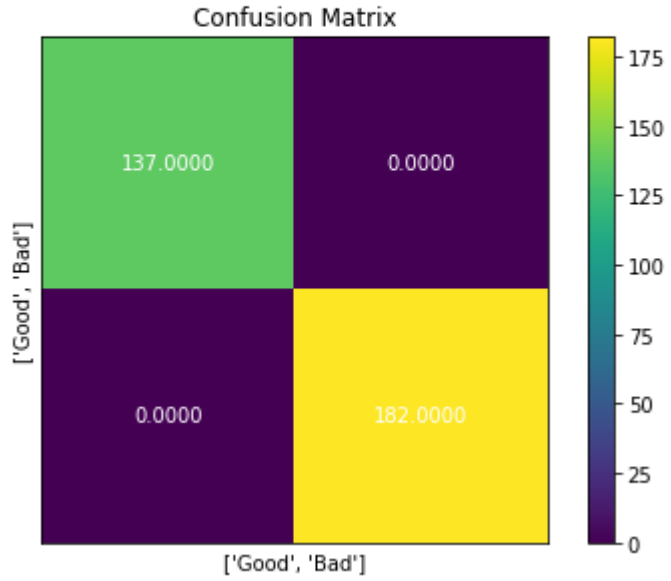




Performance Metrics

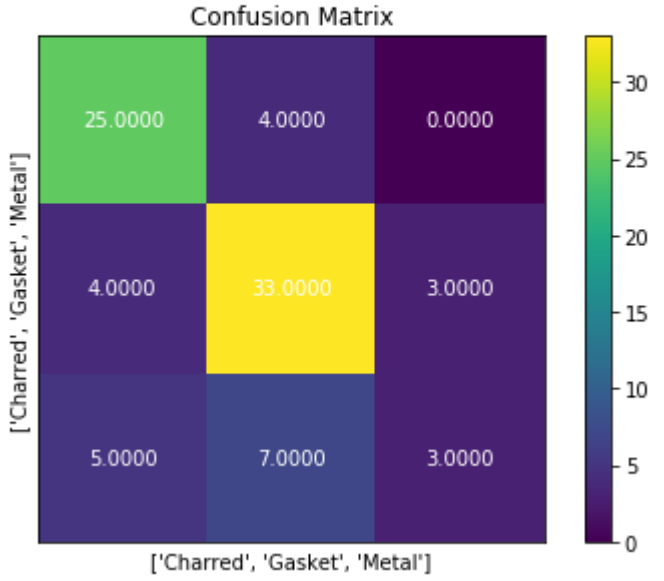
Binary Classification Metrics

```
Acuuracy: 1.0
Precision: 1.0
Recall: 1.0
F1_score: 1.0
Specificity: 1.0
```



Ternary Classification Metrics

```
Acuuracy: 0.7261904761904762
Precision: [0.73529412 0.75       0.5       ]
Recall: [0.86206897 0.825      0.2       ]
F1_score: [0.79365079 0.78571429 0.28571429]
```



Design choices

- Tried PCA first for feature extraction and Initial SVM for Classification with PCA reduced feature set. Poor results.
- Hints from the organizers and the paper suggested Neural Networks for Feature Extraction, which seems to be a common method.
- 3 channel grayscale to RGB conversion reduces dataset size and is compatible with the Resnet101, packing **thrice** the amount of information into one tensor.
- Tried a CNN for Ternary Classification, poor results due to limited samples. Switched to Linear Kernel based SVM.
- SVMs perform well for smaller datasets.
- Applied Further Data Augmentation such as mask-based cropping. Accuracy decreased due to remaining artefacts from the mask-cropping.
- Realized that static parts of the image do not play a role in the learning outcome of the models. So mask-cropping not entirely necessary for the sole purpose of size reduction.
- Optimized training for **Validation Loss/Validation accuracy**.

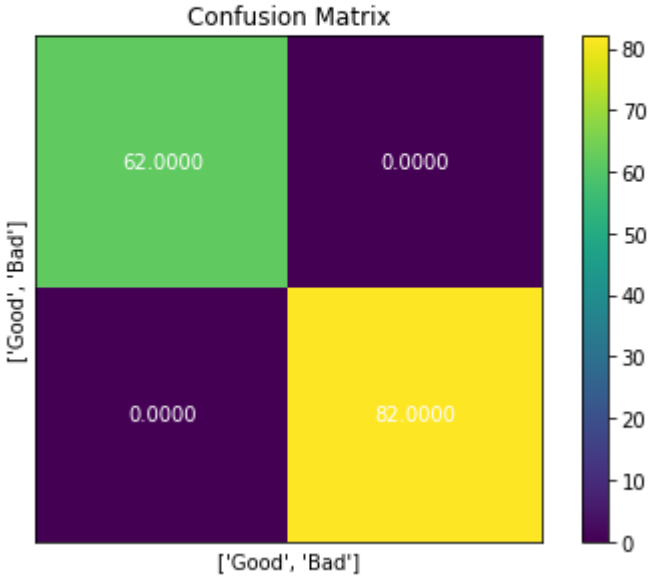
Additional Information

- Libraries Used
  - Pytorch, Keras and Scikit-Learn for Neural Network building, Classical Machine Learning models and for Classification Metrics
- What Did I Learn
  - Industrial application example of Computer Vision.
  - How to use state of the art Neural Networks for Feature Extraction.
  - Solidified knowledge in Transfer Learning.
  - Combining Classical Machine Learning and Deep Learning.
  - Image Processing aspects for Masking Region of Interests.
  - Differences between normal object detection and miniscule object detection methods.
  - Understanding where not to use CNNs.
  - Learn about Binary Entropy Losses and Categorical Entropy Losses

Test Dataset Performance Metrics

Binary Classification Metrics

```
Acuuracy: 1.0
Precision: 1.0
Recall: 1.0
F1_score: 1.0
Specificity: 1.0
```



Ternary Classification Metrics

```
Acuuracy: 0.7419354838709677
Precision: [0.88888889 0.67647059 0.7       ]
Recall: [0.66666667 0.88461538 0.58333333]
F1_score: [0.76190476 0.76666667 0.63636364]
```