**Отчет о создании прототипа нейронной сети**

1. Тема, описание задачи

Тема проекта: **"Классификация дефектов на поверхности стали"**. Задача Object Detection: необходимо на снимках поверхности стали определить bounding box, в котором находится дефект. Всего 10 классов различных дефектов, соответственно задача также заключается в том, чтобы научиться максимально точно классифицировать найденный дефект.

1. База (если нет конфиденциальных данных)

База для данной задачи уже собрана:

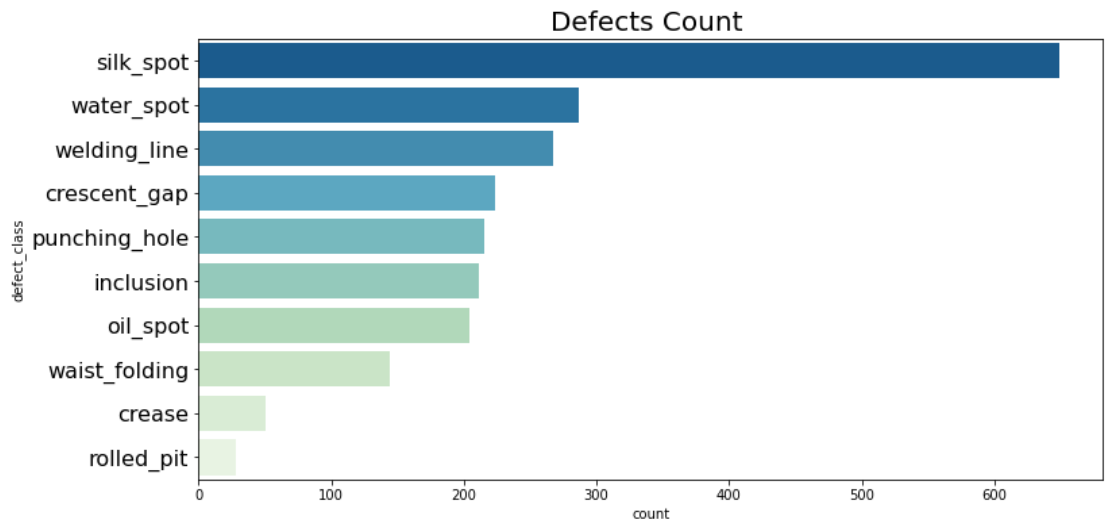
https://www.kaggle.com/datasets/zhangyunsheng/defects-class-and-location

Она содержит 2280 снимков поверхности стали и html-файлы с разметкой bounding boxes для дефектов на поверхности стали.

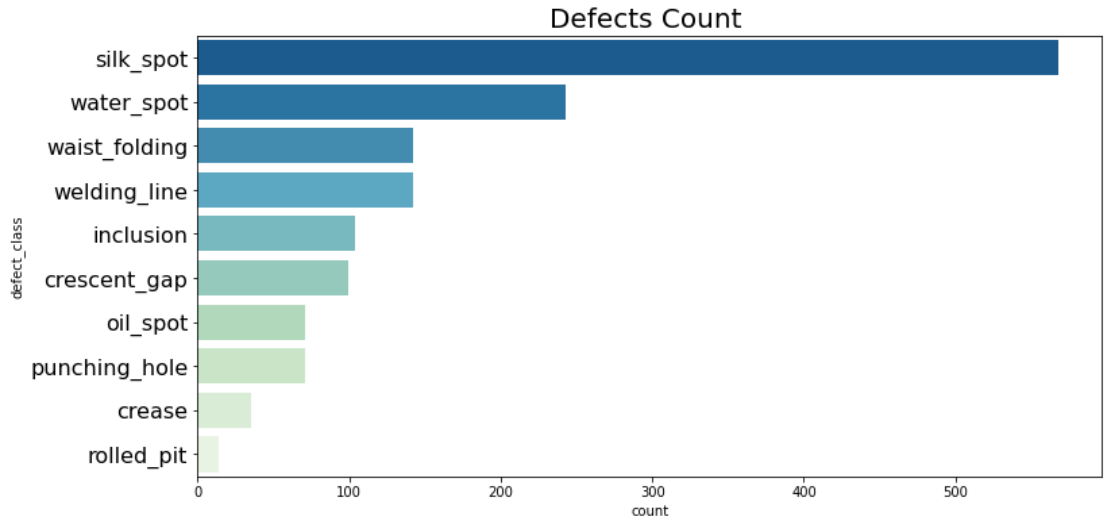
Всего в задачи представлено 10 разных классов дефектов:

* water\_spot
* inclusion
* waist\_folding
* punching\_hole
* silk\_spot
* welding\_line
* oil\_spot
* crease
* crescent\_gap
* rolled\_pit

При этом на части изображений есть сразу несколько разных дефектов и при этом для данных изображений нет возможности однозначно определить какой bounding box к какому типу дефектов относится. Если оставить только изображения с одним bounding box, то количество примеров сократится до 1489 изображений. Также стоит отметить, что классы крайне не сбалансированы: самый популярный класс silk\_spot содержит 649 примеров изображений, а самый редкий класс rolled\_pit - 28 примеров.



*Распределение кол-ва примеров для полной базы*

**

*Распределение кол-ва примеров для базы, где на каждом изображении есть только один вид дефекта*

1. Параметризация данных

Для дальнейшей обработки и обучения сетей данные собраны в датафрейм. В нем сохранены:

- название файла со снимком стали

- полный путь до файла с изображением

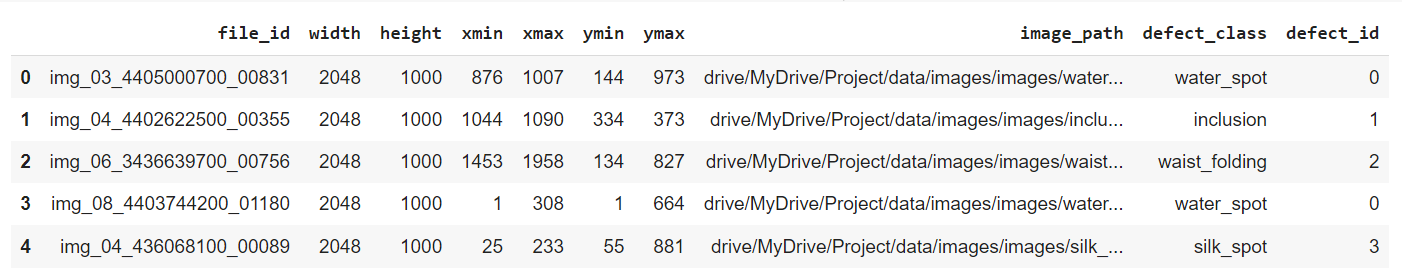
- ширина и высота изображения

- границы bounding box: xmin, xmax, ymin, ymax

- название дефекта

- сопоставленный названию дефекта номер класса дефекта

Пример первых пяти строк датафрейма:



1. Архитектура нейросети, графики обучения

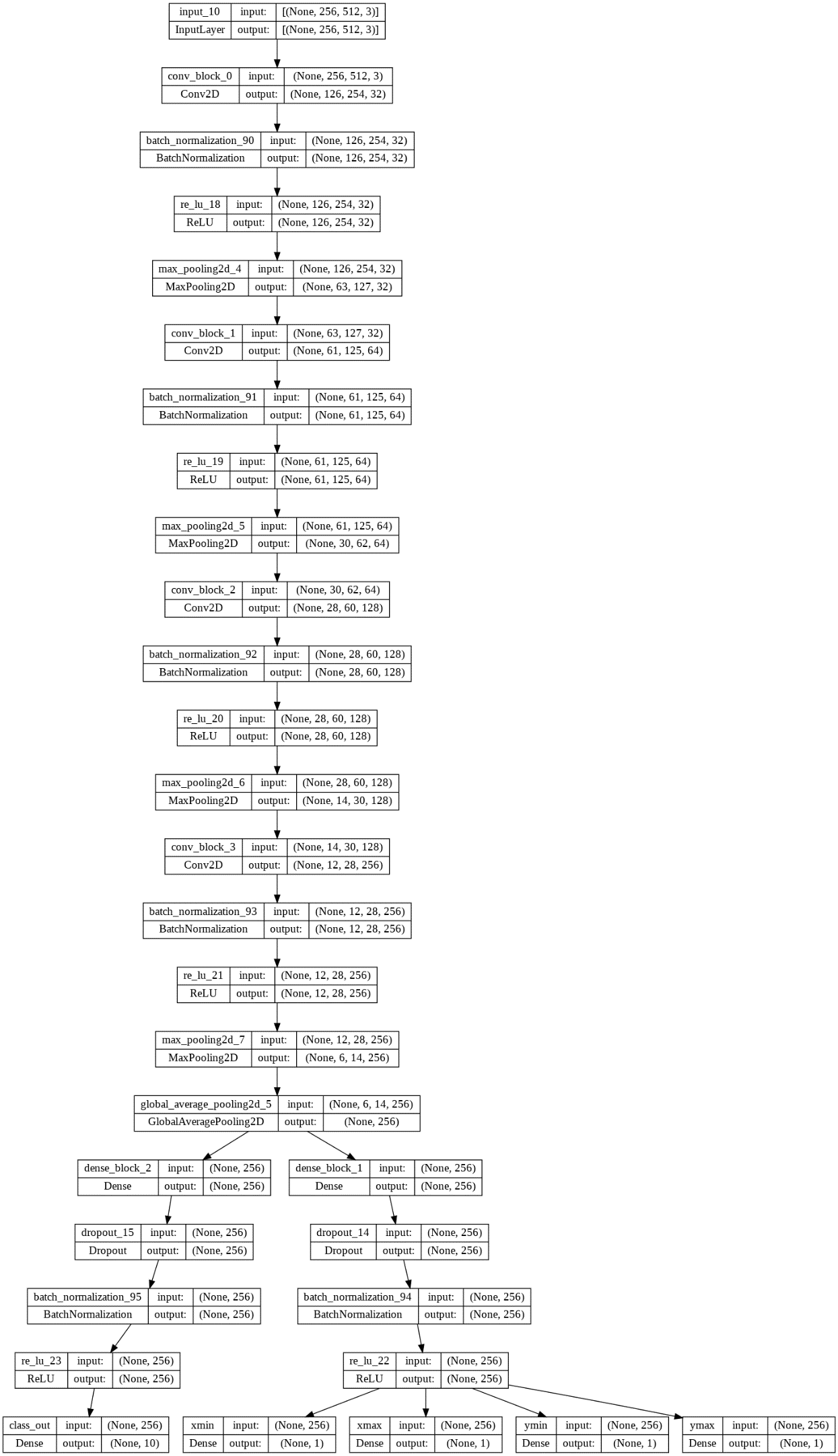
Было построено две нейросети с двумя разными архитектурами.

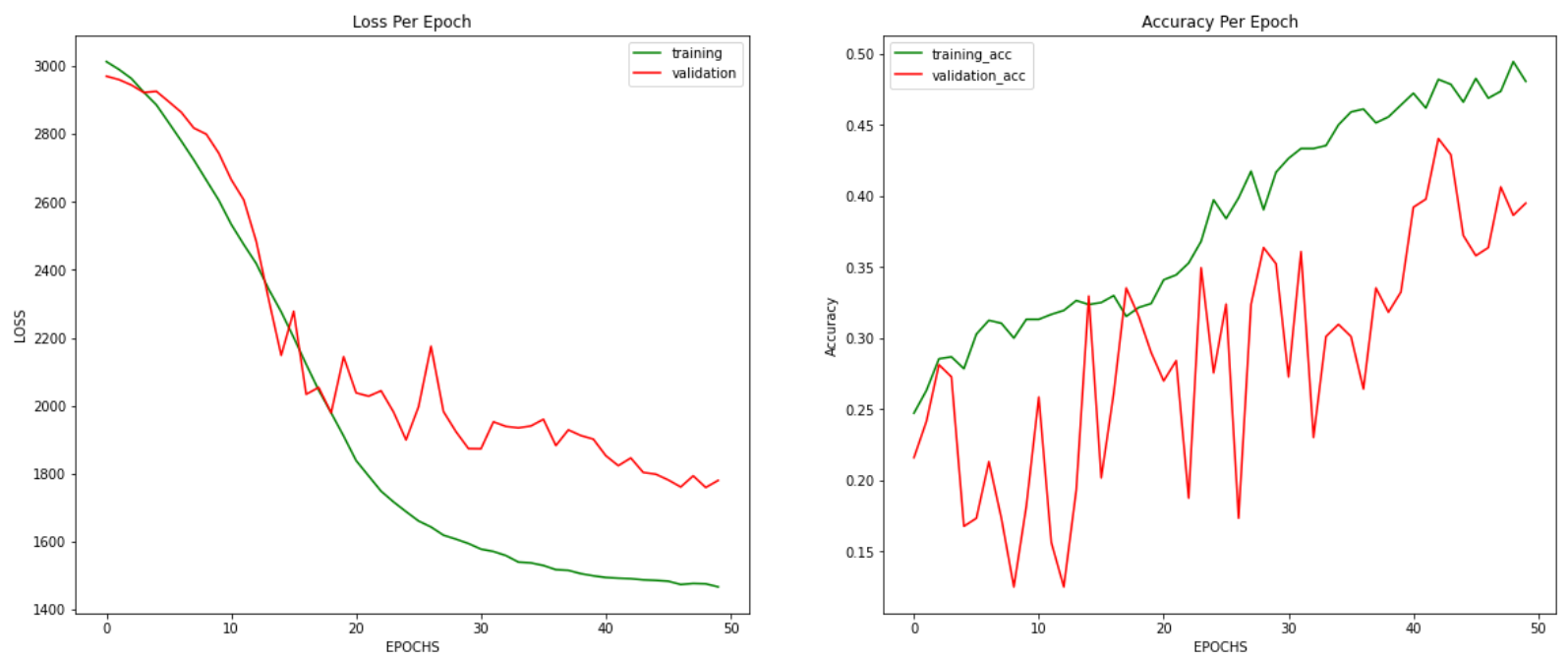
Также в ходе исследования выяснилось, что часть изображений содержит сразу несколько дефектов, но в этом случае нет возможности определить класс каждого из дефектов. И поэтому было решено оставить только те изображения, которые содержат один дефект.

Поскольку вторая архитектура нейросети показала более высокое качество, то только для данной архитектуры было проведено второе обучение на урезанном датасете (где на каждом изображении только один дефект).

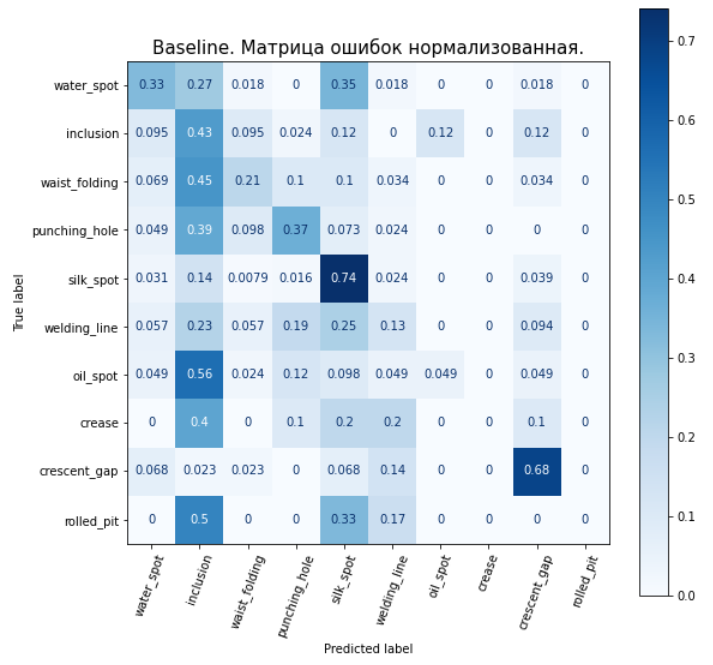
1. Первая архитектура построена на использовании сверточных слоев (Conv2D – BatchNormalization – activation ReLU - MaxPooling2D) и полносвязных слоев (Dense – Dropout – BatchNormalization - activation ReLU)

Архитектура сети:





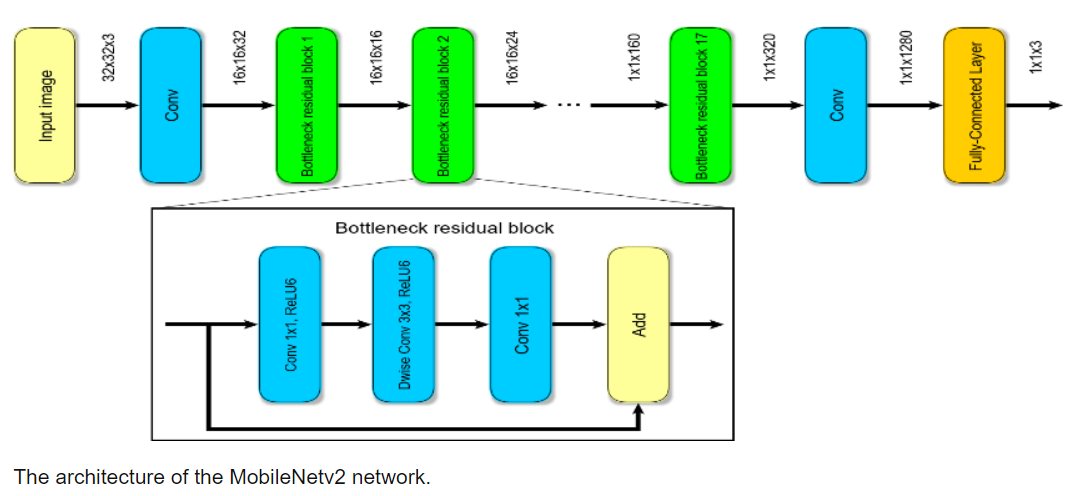
*Графики обучения первой нейросети на полном датасете на 50 эпохах*

**

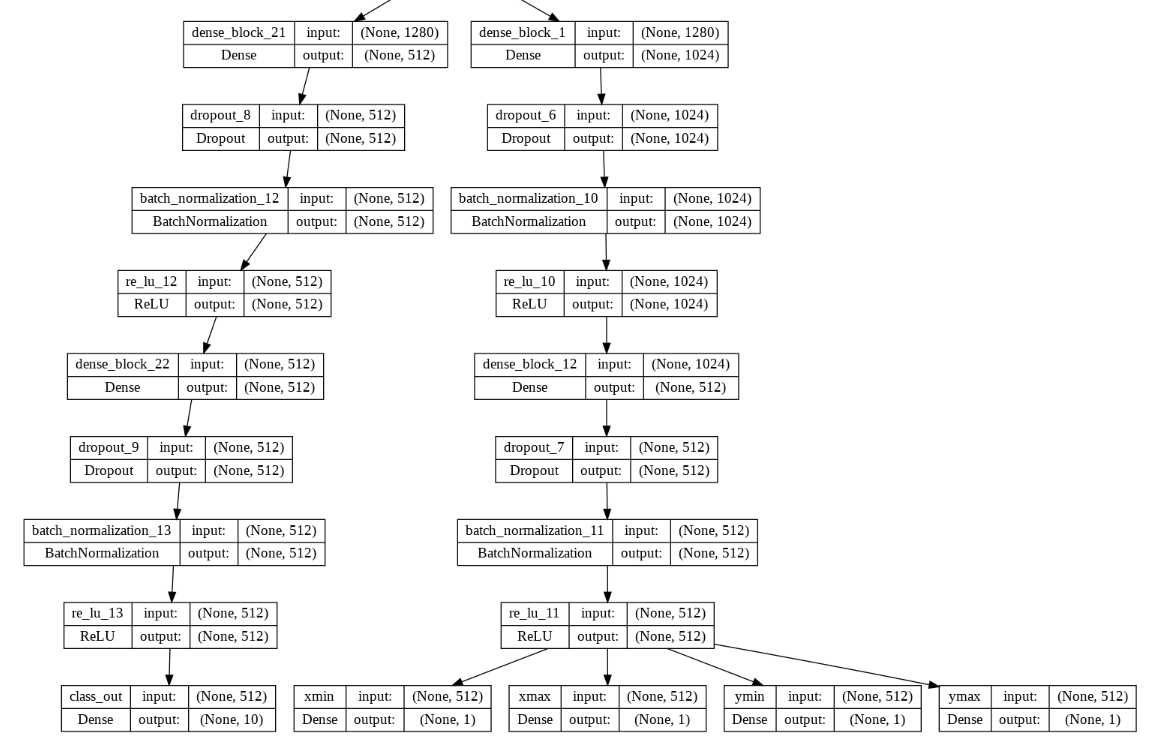
*Матрица ошибок предсказания класса дефекта (первая нейросеть, 50 эпох)*

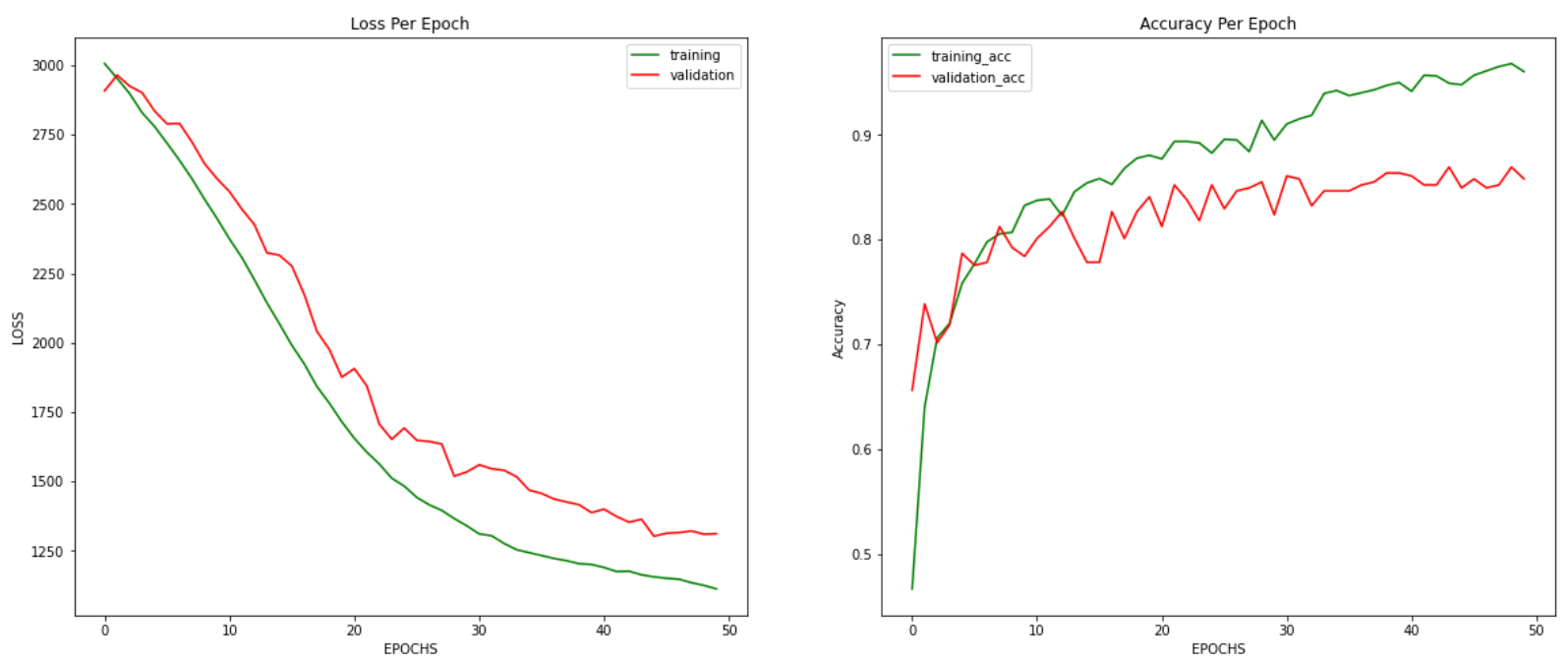
1. Вторая архитектура берет в качестве основы предобученную MobileNetV2 и добавляет полносвязные слои (Dense – Dropout – BatchNormalization - activation ReLU) для предсказания класса дефекта и границ bounding box

Архитектура сети:

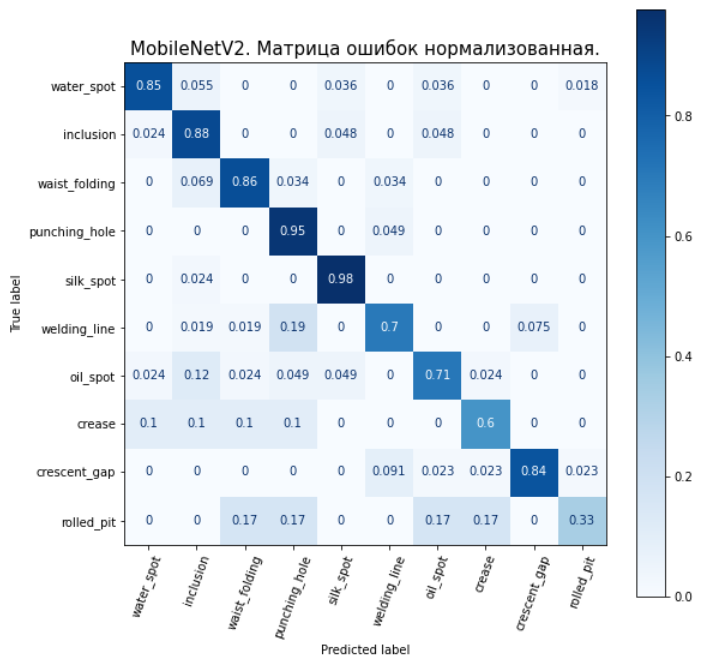


Далее к MobileNetv2 сети добавляется следующая архитектура:

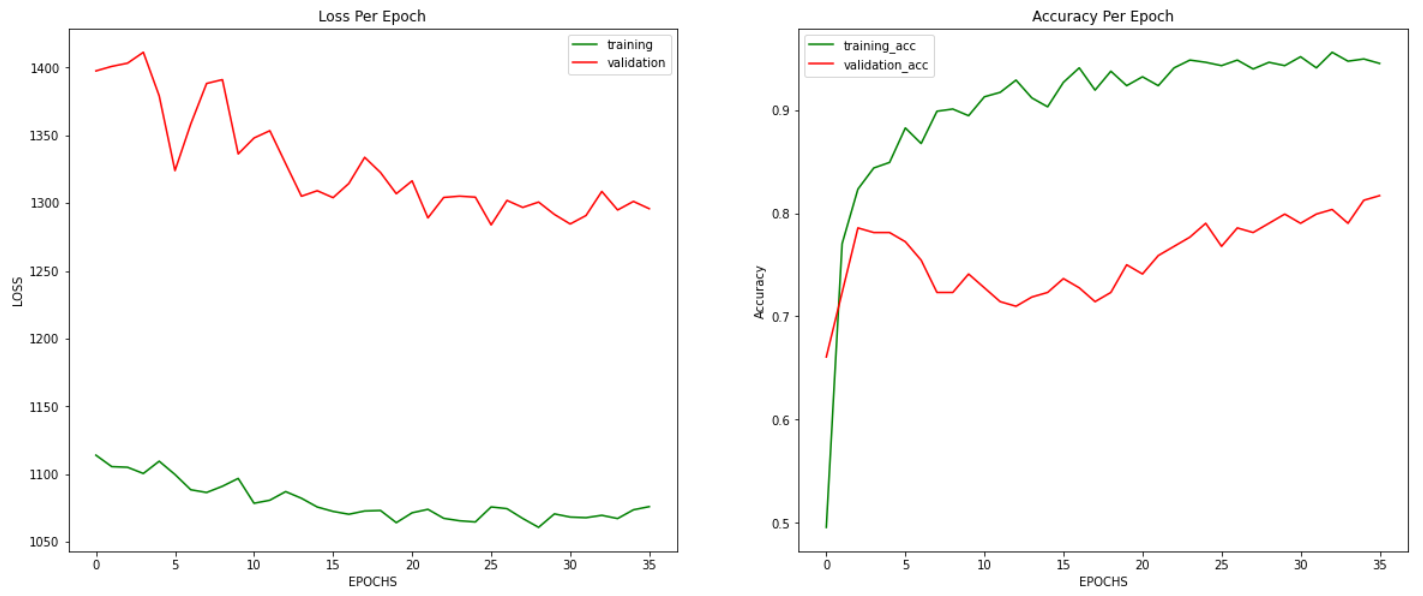




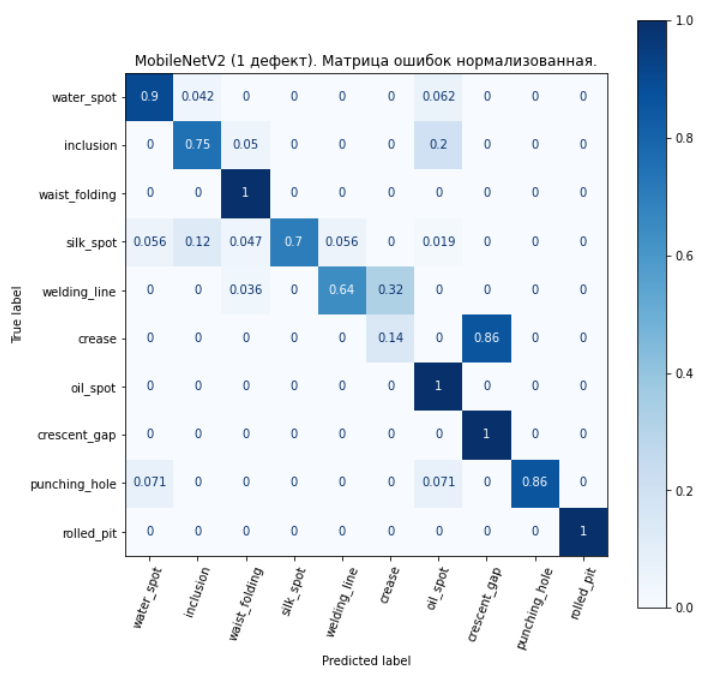
*Графики обучения второй нейросети на полном датасете на 50 эпохах*



*Матрица ошибок предсказания класса дефекта (вторая нейросеть, полный датасет, 50 эпох)*

**

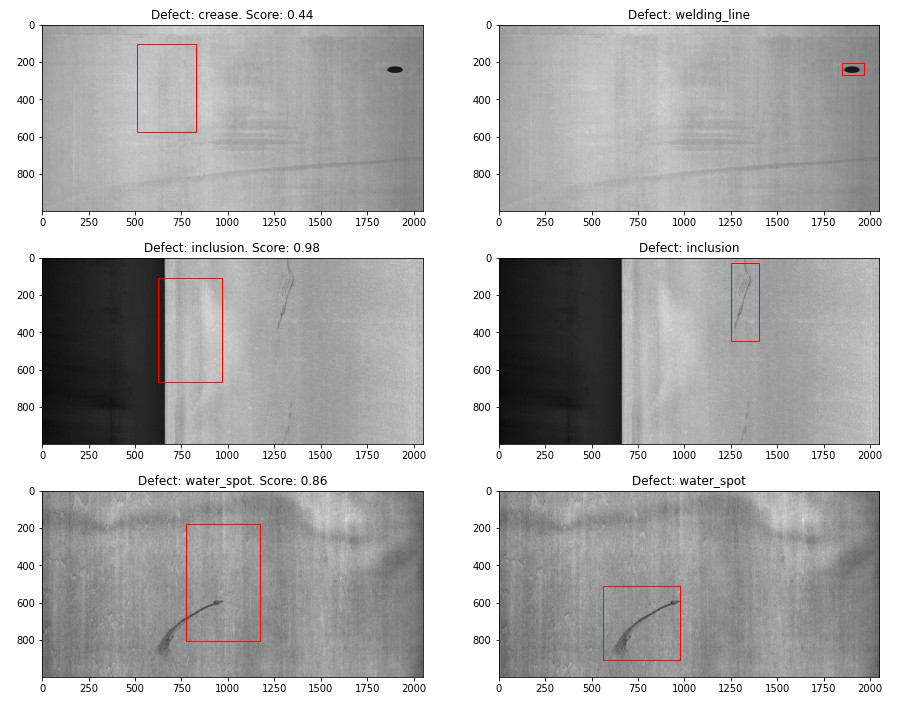
*Графики обучения второй нейросети на сокращенном датасете на 50 эпохах*

**

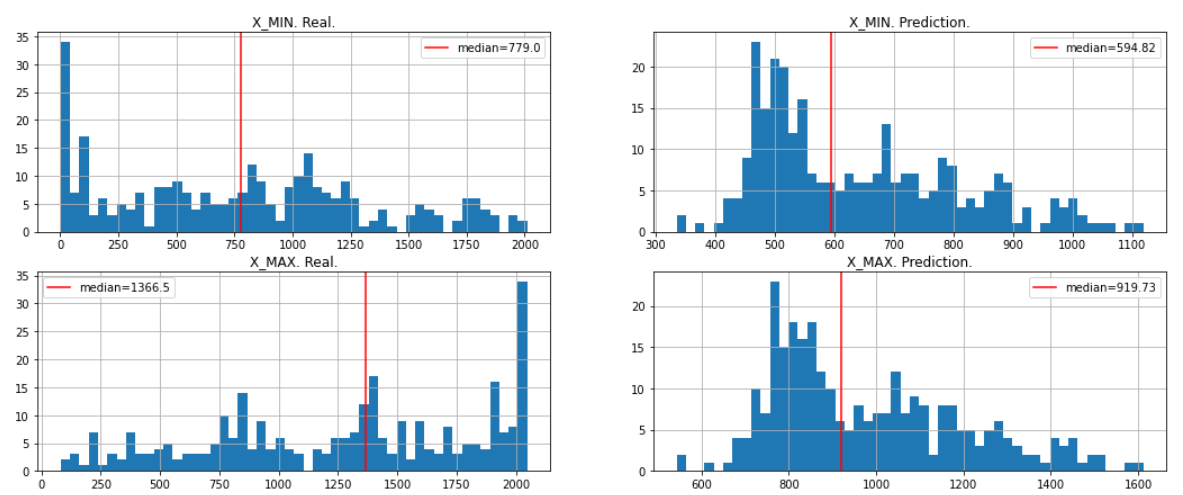
*Матрица ошибок предсказания класса дефекта (вторая нейросеть, сокращенный датасет, 50 эпох)*

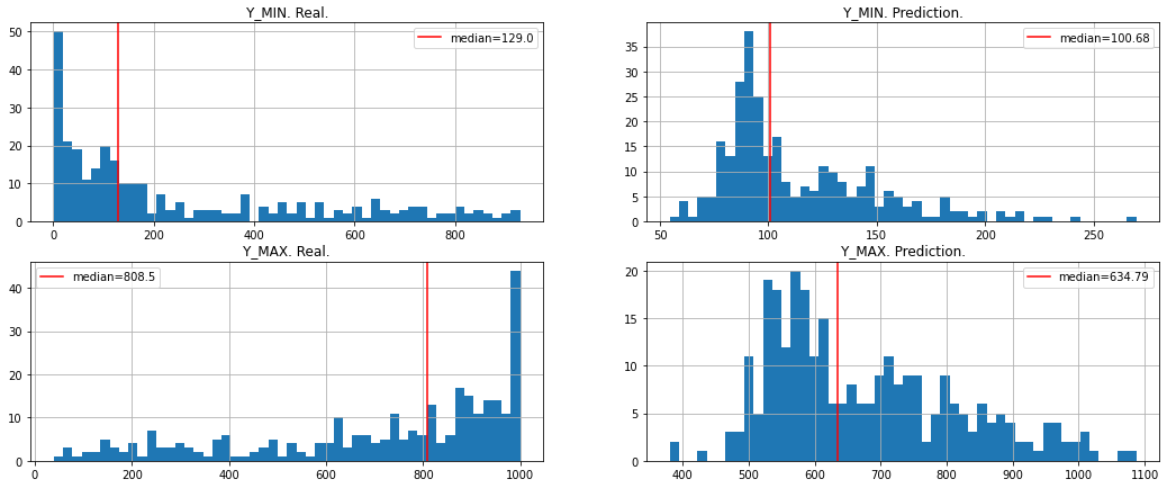
1. Выводы

С одной стороны, в соответствии с последней приведенной матрицей ошибок, мы видим высокое качество классификации классов дефектов (за исключением класса ‘crease’). С другой, наблюдается низкое качество по прогнозированию местонахождения bounding boxes. Примеры определения bounding boxes:



Распределения значений xmin, xmax, ymin, ymax для оценки разброса реальных и предсказанных значений:





Видим, что реальные распределения значений xmin, xmax, ymin, ymax имеют в основном более "широкий" разброс (крайние значения) по сравнению с предсказанными и сильно отличные медианные значения. Это говорит об очень низком качестве предсказания bounding boxes.

1. План дальнейшей работы

Далее планируется обучить последнюю нейросеть на большем количестве эпох с целью повышения качества предсказания рамок для дефектов, а также попробовать применить архитектуру YOLOv3.