

Институт математики, механики и
компьютерных наук им. И.И. Воровича

Прикладная математика и информатика

Кафедра теории упругости

Определение и сегментация трещин на поверхности при помощи свёрточной нейронной сети

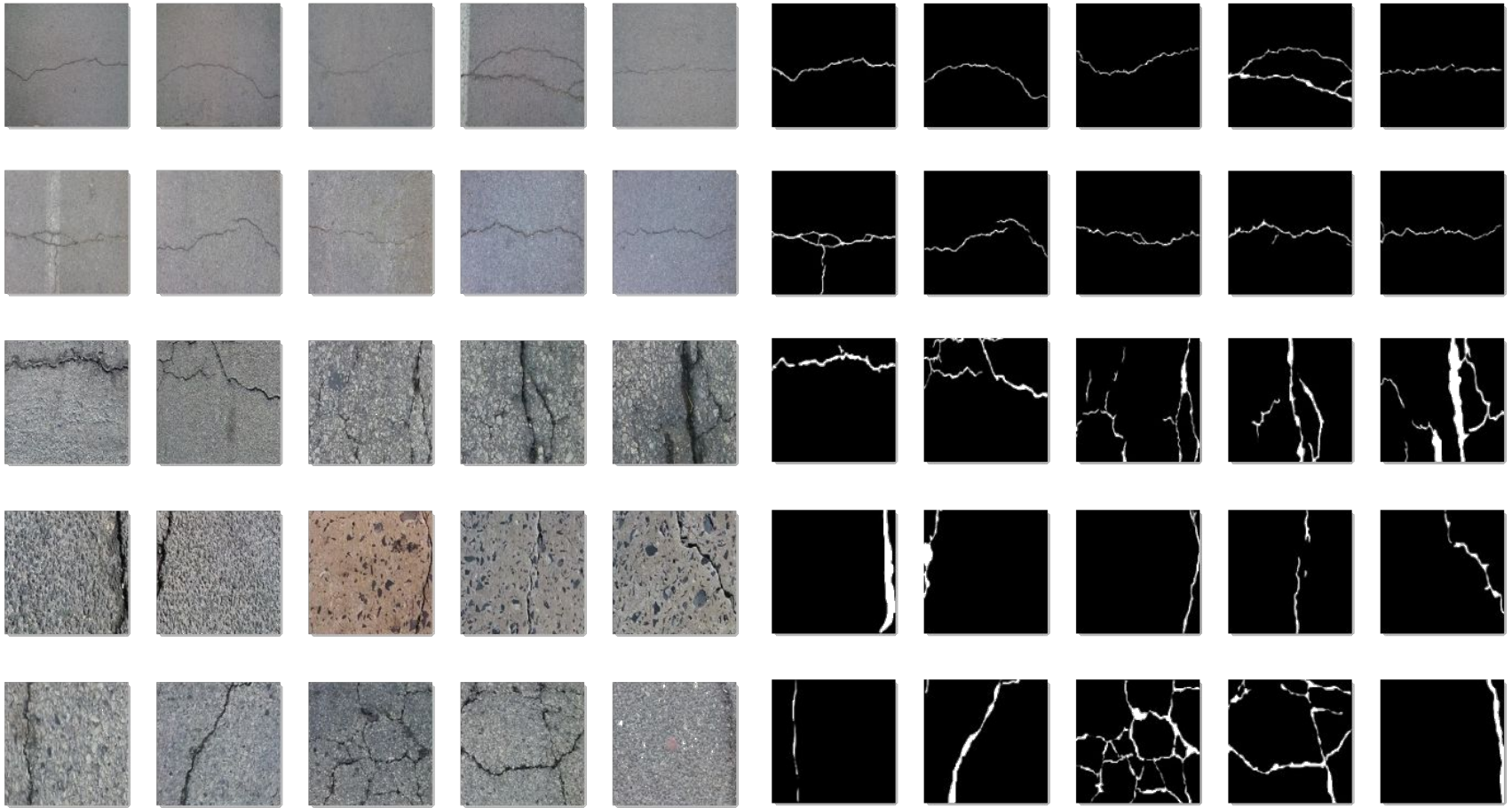
Студент: Пандов Вячеслав Дмитриевич
Научный руководитель: д.ф.-м.н., доц. Карякин Михаил Игорьевич

Постановка задачи



- Реконструкция
- Отбраковка
- Контроль качества
- Краш-тесты

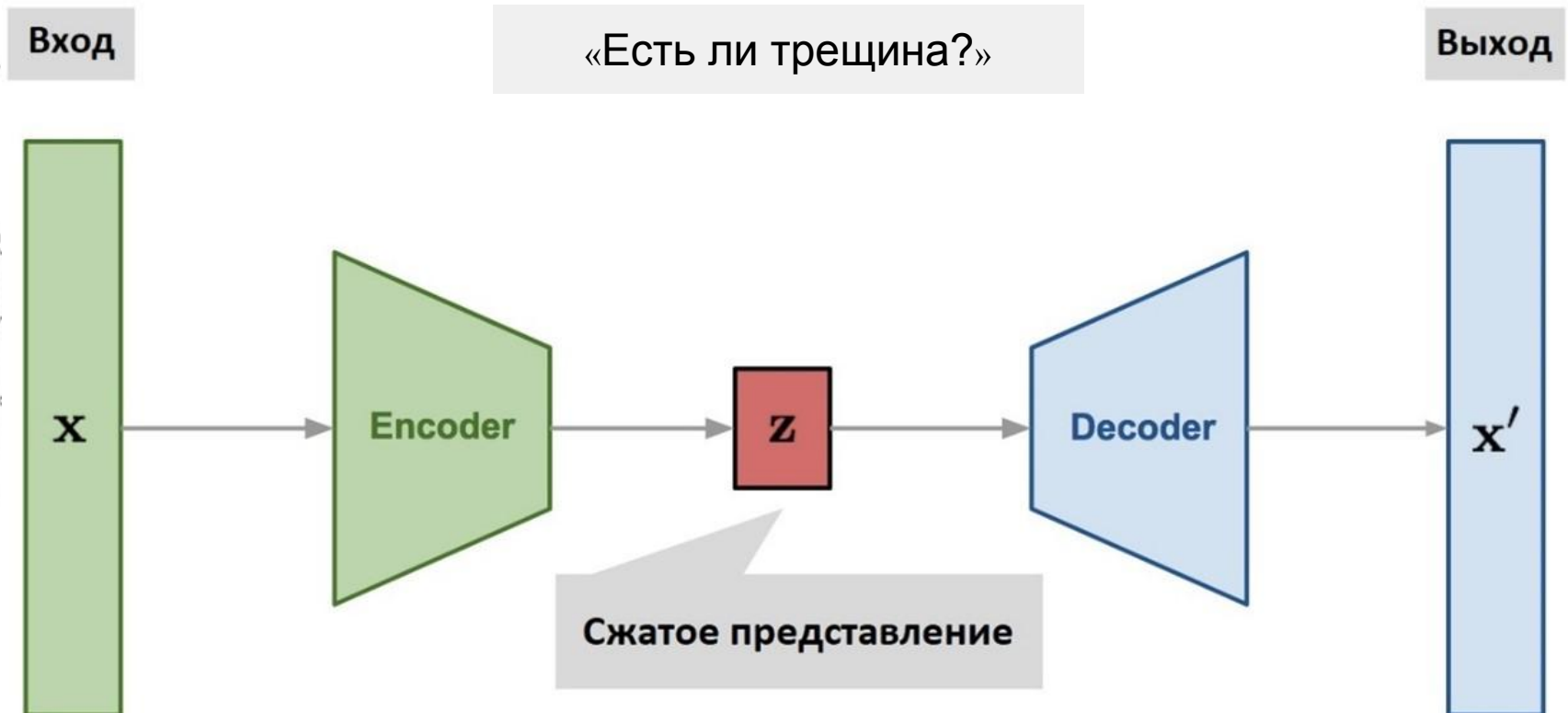
Объект исследования



Исходные изображения

Бинарные маски

Принцип архитектур «Variance Autoencoders»



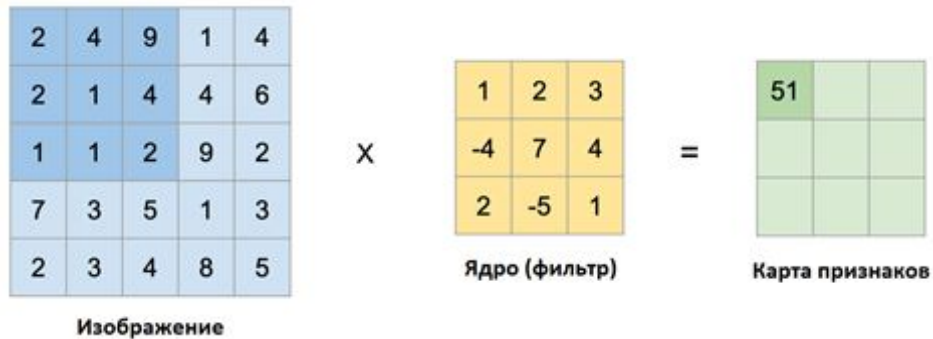
- **Кодировщик.** Сжимает входную информацию в так называемое «сжатое представление».
- **Декодер.** Реконструирует сжатое представление к необходимому конечному виду.

Цели и задачи

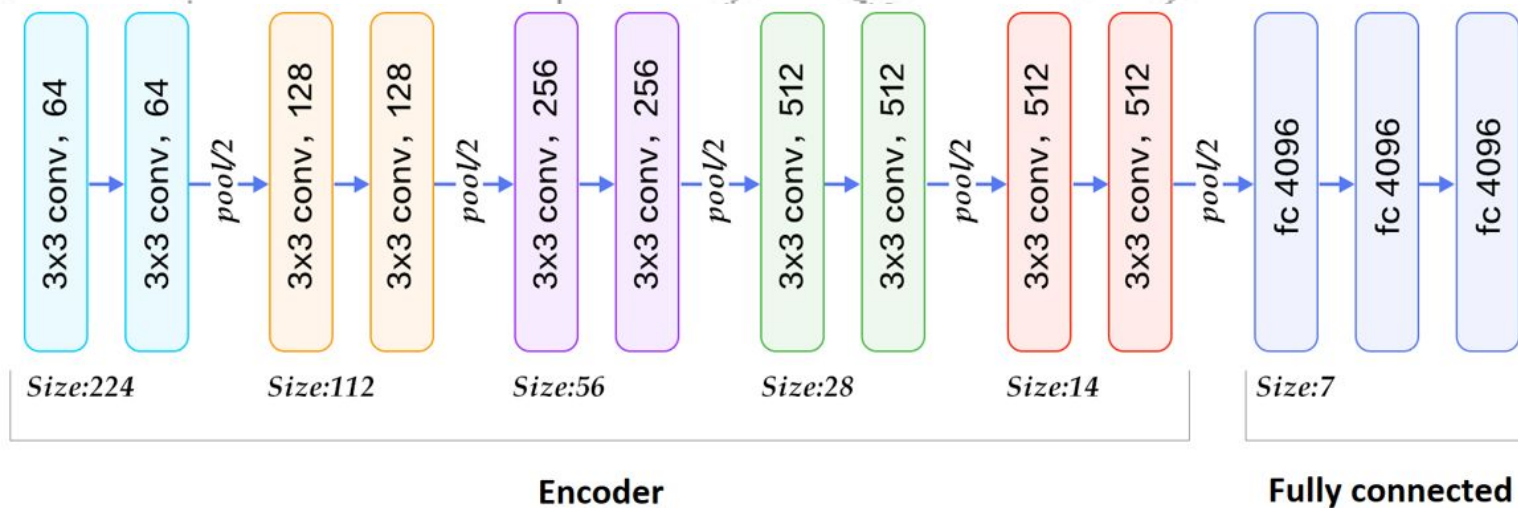
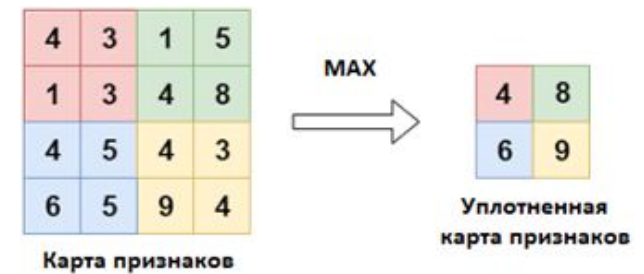
1. Реализовать и обучить модель для сегментации трещин на поверхности по изображению.
2. Модифицировать модель из п.1 для решения задачи определения трещин на тех же изображениях поверхности. Обучить такую модель.

Архитектура VGG-13

- Свертка



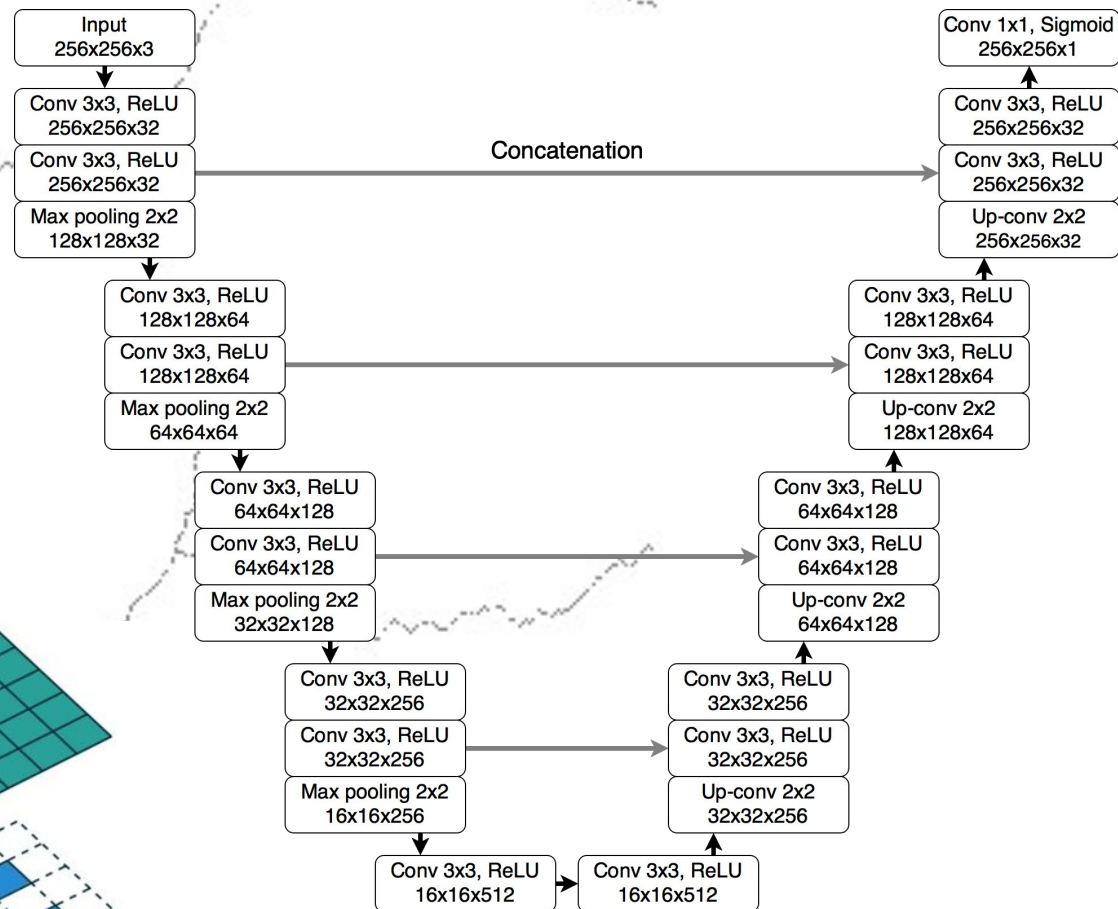
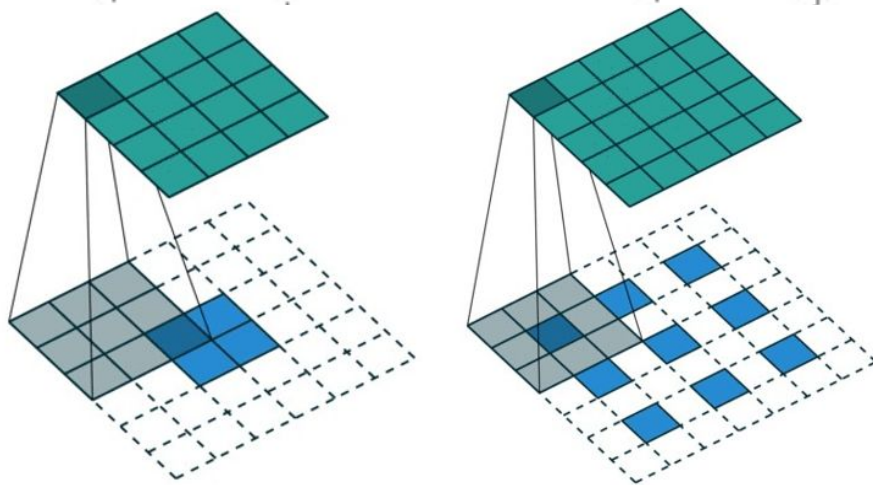
- Субдискретизация



Архитектура U-Net

- VGG Encoder

- Обратная свертка



Критерий классификации: бинарная кросс-энтропия

- Бинарная кросс-энтропия

$$BCE = -t \log(p) + (1 - t) \log(1 - p),$$

$$BCE \in (0, +\infty),$$

$$BCE \rightarrow \min,$$

$$p = \sigma(y),$$

- Производная бинарной кросс-энтропии

$$\frac{\partial BCE}{\partial y} = \frac{\partial BCE}{\partial p} \cdot \frac{\partial p}{\partial y} =$$

$$= -\frac{t}{p} \cdot p \cdot (1 - p) + \frac{1 - t}{1 - p} \cdot p \cdot (1 - p) =$$

$$= -t \cdot (1 - p) + (1 - t) \cdot p =$$

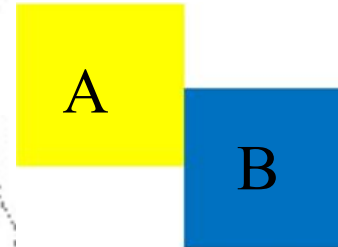
$$= -t + t \cdot p + p - t \cdot p =$$

$$= p - t$$

Критерий сегментации: коэффициент Сёренсена

- С точки зрения множеств

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$



Dice = 0



Dice = 0.25



Dice = 1

- С точки зрения пикселей

$$Dice = \frac{2pt + 1}{p + t + 1},$$

$$t \in \{0,1\}, \quad p = \sigma(y), \quad y \in \mathbb{R}$$

Dice Index = 0.83828



Метрики качества

Сегментация

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Классификация

	$t = 1$	$t = 0$
$y = 1$	True Positive (TP)	False Positive (FP)
$y = 0$	False Negative (FN)	True Negative (TN)

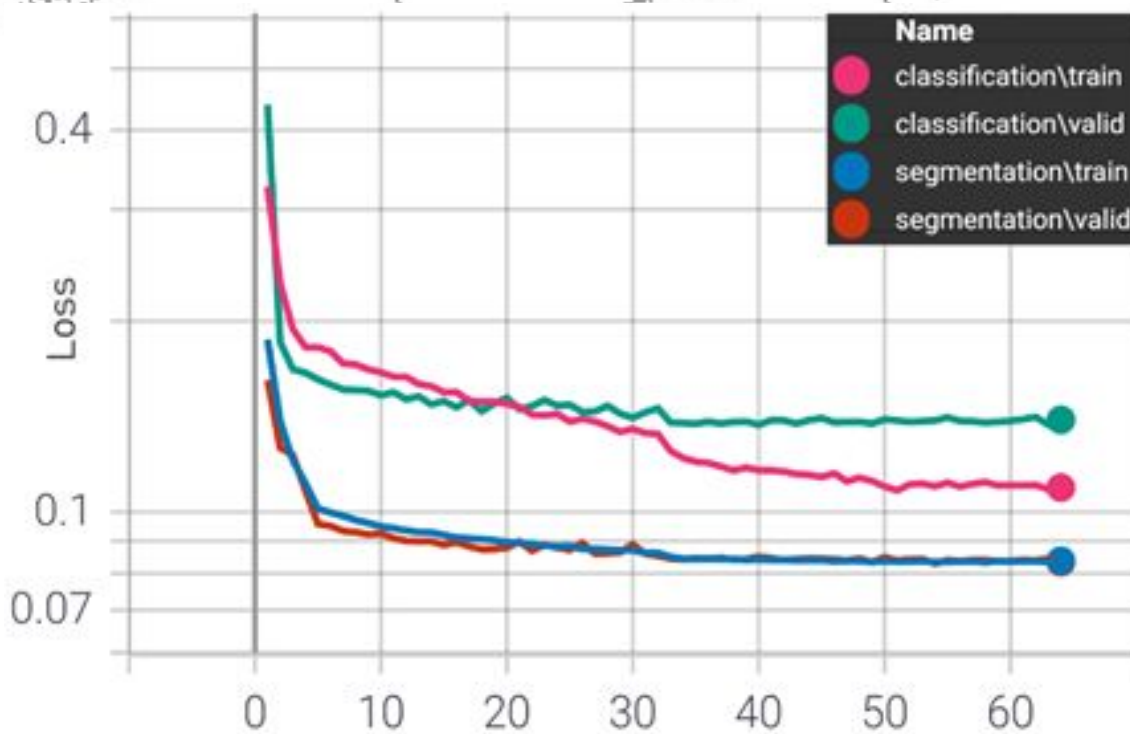
$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F_{\beta} = (1 + \beta^2) \cdot \frac{precision \cdot recall}{(\beta^2 \cdot precision) + recall}$$

Результаты



SEGMENTATION

LOSS 0.083

BCE 0.053

DICE 0.647

IOU 0.481

CLASSIFICATION

	Class 0	Class 1	Mean
PRECISION	0.920	0.966	0.943
RECALL	0.947	0.948	0.947
F1	0.933	0.957	0.945
LOSS	0.137		

Результаты

1

1

1

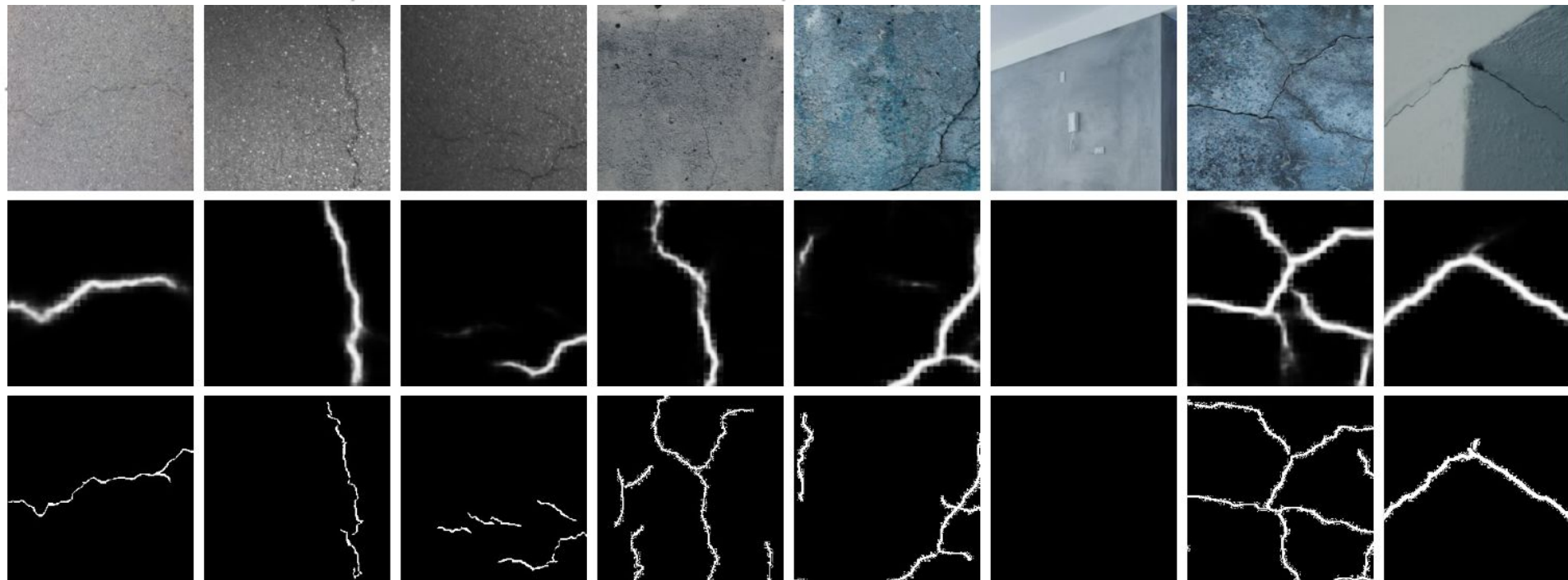
1

1

0

1

1



1) Входные изображения

2) Предсказанные маски

3) Истинные маски

Заключение

По результатам исследования можно сделать следующие выводы:

1. Была реализована и успешно обучена модель архитектуры U-Net для решения задачи сегментации .
2. Кодировщик, обученный на задаче сегментации, может успешно применяться в задаче классификации, если к нему добавить полносвязные нейронные слои.
3. Полносвязная нейронная сеть, добавленная после уже обученного кодировщика, успешно анализирует информацию, исходящую из кодировщика и способна интерпретировать её для решения задачи классификации.