



ФЕДЕРАЛЬНОЕ АГЕНТСТВО ВОЗДУШНОГО ТРАНСПОРТА

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ» (МГТУ ГА)**

Выпускная квалификационная работа

**Распознавание морских судов на аэрофотоснимках
методами компьютерного зрения**

Выполнил:

студент группы ПМб 4-1
Фейзуллин К.М.

Научный руководитель:

к.ф.-м.н., доцент
Филонов П. В.

Москва, 2021

Актуальность

На 1 апреля 2015 года насчитывалось 87 тыс. только торговых судов



Сервис, изучающий использование семантической сегментации для определения местоположения корабля по аэрофотоснимкам – **«MarineTraffic»**

Решение данной задачи, позволяющее вести независимое наблюдение за потоком кораблей, позволит избежать застоя морских судов, как это было недавно в Суэцком канале. Данный инцидент стоил компаниям в совокупности 400 млн \$ за каждый час бездвижного ожидания.

Цель работы

- Построение алгоритма, позволяющий зафиксировать морское судно на изображении, а также определить его точное местоположение на фотографии.

Задачи:

- Описание набора данных;
- Выбор метрик качества;
- Выбор используемых моделей и исследование их обучения;
- Сравнение полученных результатов.

Описание данных

Image: 350e77b90.jpg



Ground Truth

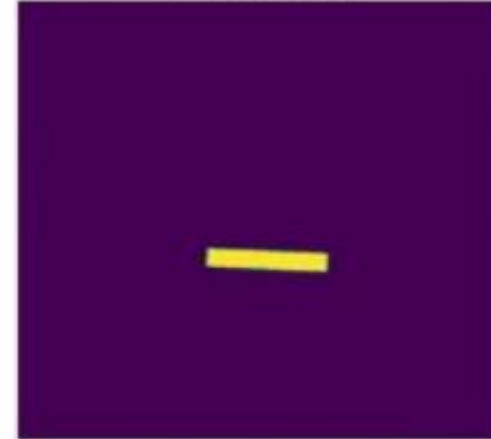


Image: 69ea9b5fa.jpg



Ground Truth



Образец изображений

Образец с декодированными пикселями
истинных ответов из табличных данных

Описание наборов данных после предобработки

	ImageId	exist_ship
0	00031f145.jpg	[0.0, 1.0]
1	0005d6d95.jpg	[0.0, 1.0]
2	000baef0c.jpg	[1.0, 0.0]
3	000fd9827.jpg	[0.0, 1.0]
4	0010e88ce.jpg	[1.0, 0.0]

Образец после применения быстрого кодирования (One-Hot Encoding) данных для задачи классификации

```
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][0][0][0][0]...
[0][0][0][0][0][0][1][1][1][0]...
[0][0][0][0][0][0][1][1][1][1]...
[0][0][0][0][0][0][1][1][1][1]...
[0][0][0][0][0][0][1][1][1][1]...
[0][0][0][0][0][0][0][1][1][1]...
```

Образец истинных ответов после декодирования данных для одного изображения задачи семантической сегментации

Выбор метрик качества

Матрица ошибок

Результат классификации	Фактический класс	
	Положительный	Отрицательный
Положительный	True Positive	False Positive
Отрицательный	False Negative	True Negative

Точность (accuracy)

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP}$$

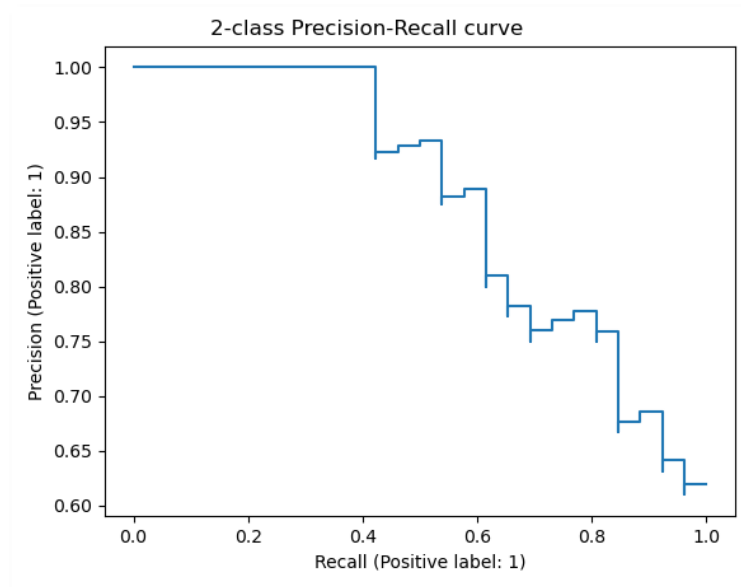
Для задачи семантической сегментации:

$$DICE = \frac{2|A \cap B|}{|A \cup B|} \rightarrow DICE = \frac{2TP}{2TP+FN+FP}.$$

Точность (precision) и полнота (recall)

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

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



Математическая постановка задачи

1. Классификация:

Есть выборка изображений X и выборка правильных ответов Y . Пусть $\xi: \Omega \rightarrow X$ – случайное изображение из X . И пусть $\eta: \Omega \rightarrow Y$ – случайный правильный ответ из Y . Тогда определим случайную величину $(\xi, \eta): \Omega \rightarrow (X, Y)$ с распределением $p(x|y)$. Задачей классификации будет являться нахождение $p(x|y)$ при заданном наборе элементов $D = \{(x_i, y_i) \sim p(x|y), i = 1, N\}$.

2. Сегментация:

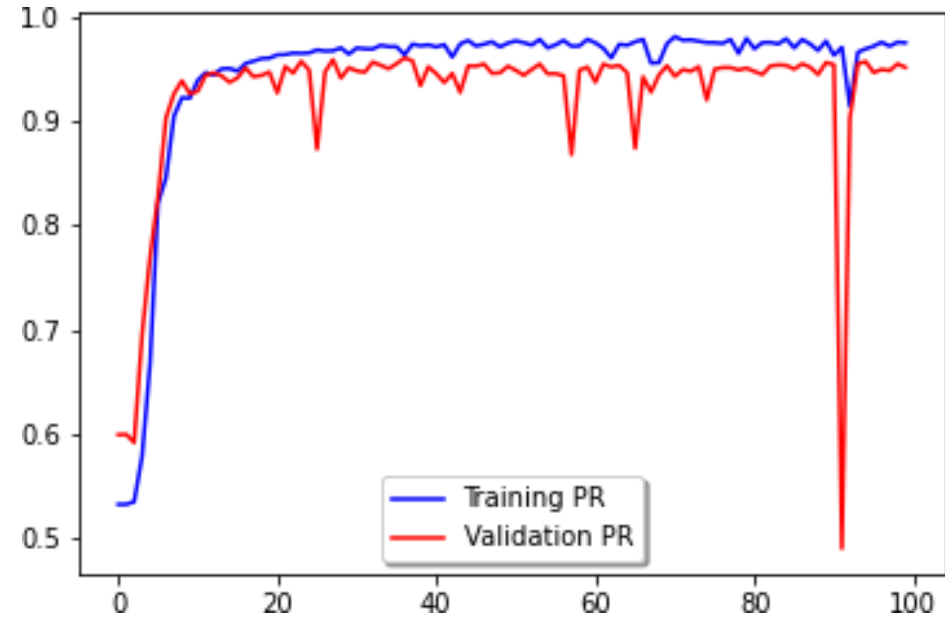
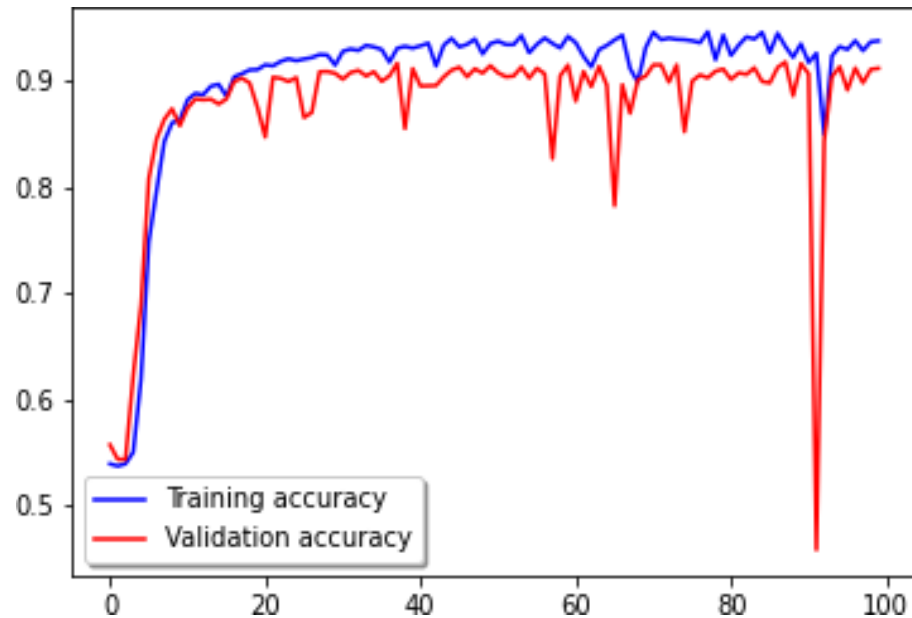
Постановка задачи классификации, где y_i – матрица правильных ответов для одного изображения.

Описание экспериментальной установки

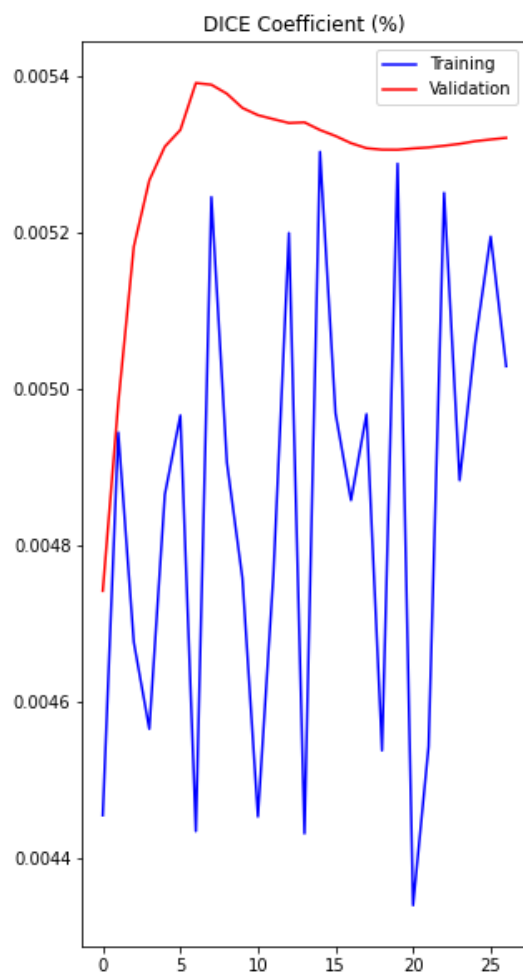
- Создание моделей и обучение реализовано на языке Python 3.6 с помощью библиотек Tensorflow, Keras, в среде программирования Jupyter Notebook, в облачной вычислительной среде Google Colab с следующими вычислительными характеристиками:
 - GPU NVIDIA T4 – 16 GB vRAM
 - 25 GB RAM
 - Ограничение на время работы – 24 часа
- Обучение каждой модели идет 50 эпох, за редким исключением для дополнительного исследования берется 100 эпох. На данных в 20 тыс. изображений для сравнения моделей одна эпоха длится 5 – 15 минут. На наборе в 70 тыс. изображений одна эпоха длится 30 – 60 минут.
- Данные для сравнения моделей загружались из облачного диска Google Drive. Данные для финального обучения модели загружались с открытого набора данных предоставленных Airbus на веб-ресурсе Kaggle.

Определение необходимого и достаточного уровня метрик качества для задачи классификации и результат исследования модели VGG16

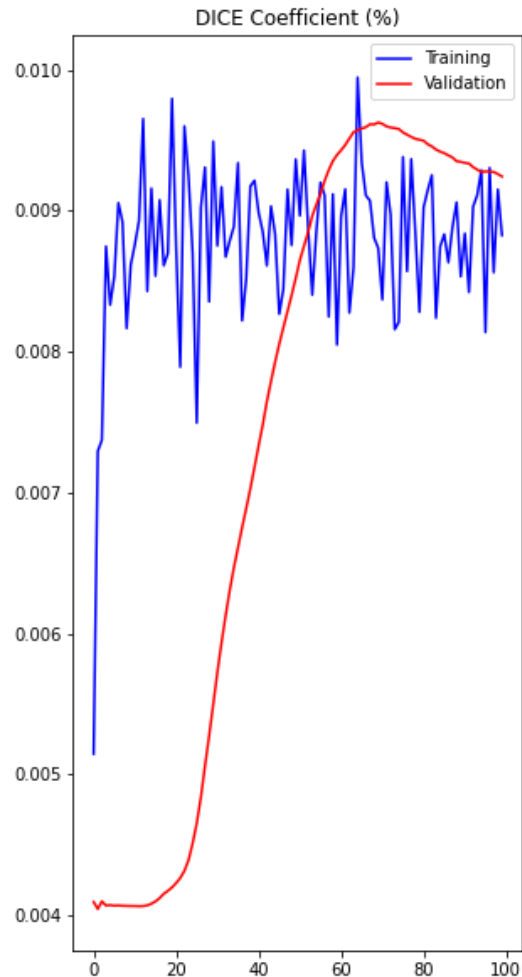
- Решение с помощью подбрасывания монеты дает значение метрики ассурасу равное 0.49
- Решение с использованием человеческих ресурсов, где человек отвечает, есть ли морское судно на изображении или нет, дает значение метрики ассурасу равное 0.78
- Во время исследования сети VGG16 удалось достичь значений метрики ассурасу равное 0.92



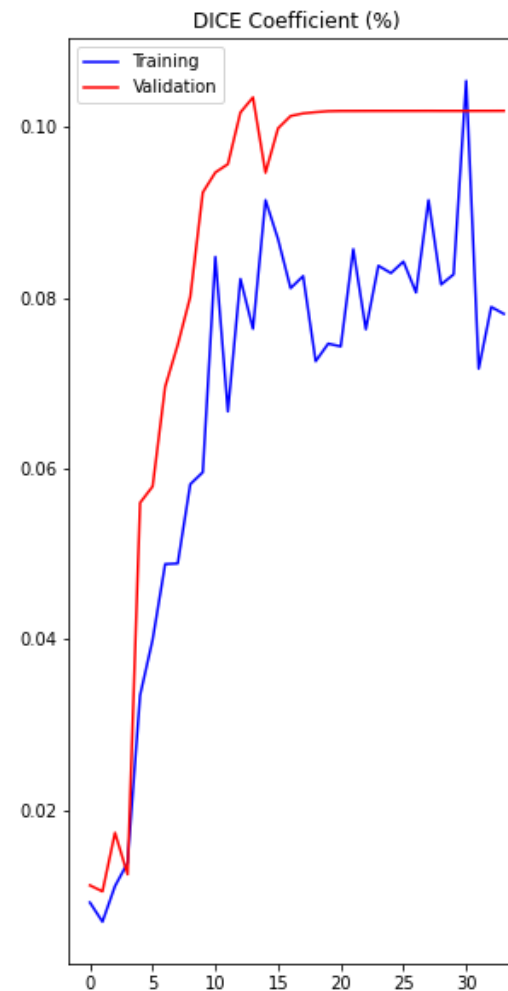
Сравнение результатов метрики сетей для сегментации изображений



U - Net

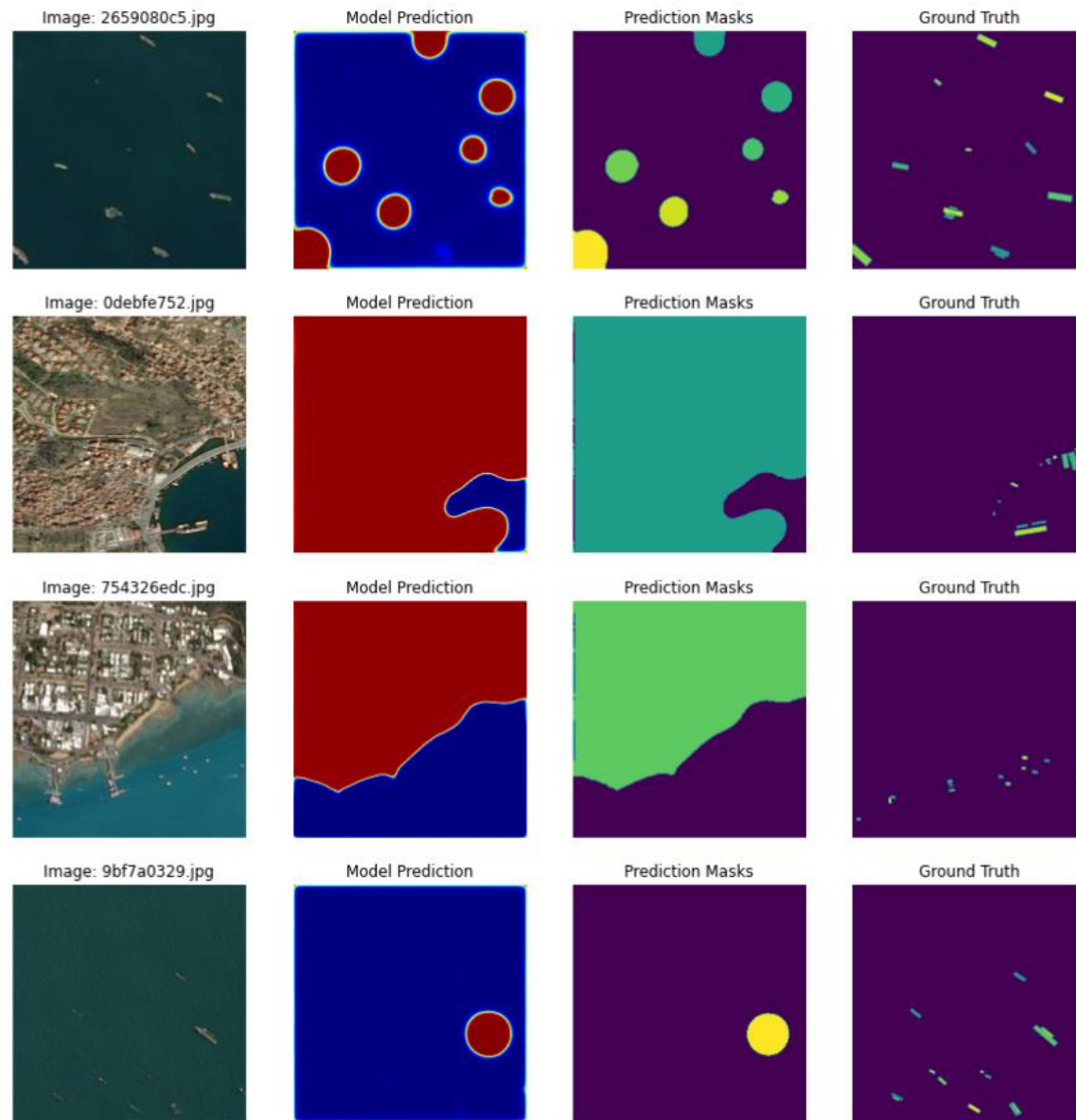


SegNet



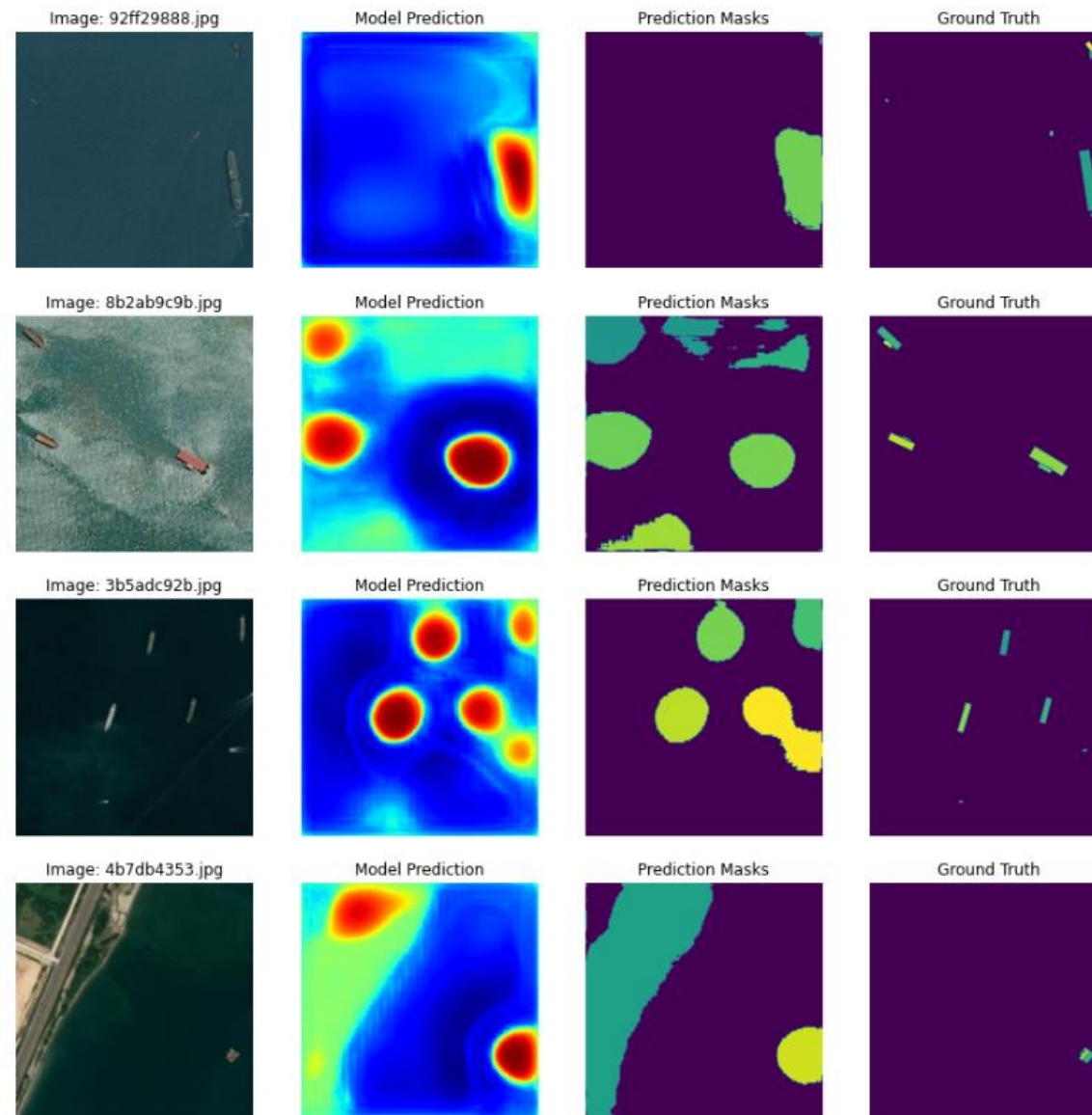
FCN VGG16

Сравнение результатов обработки сетей для сегментации изображений - 1



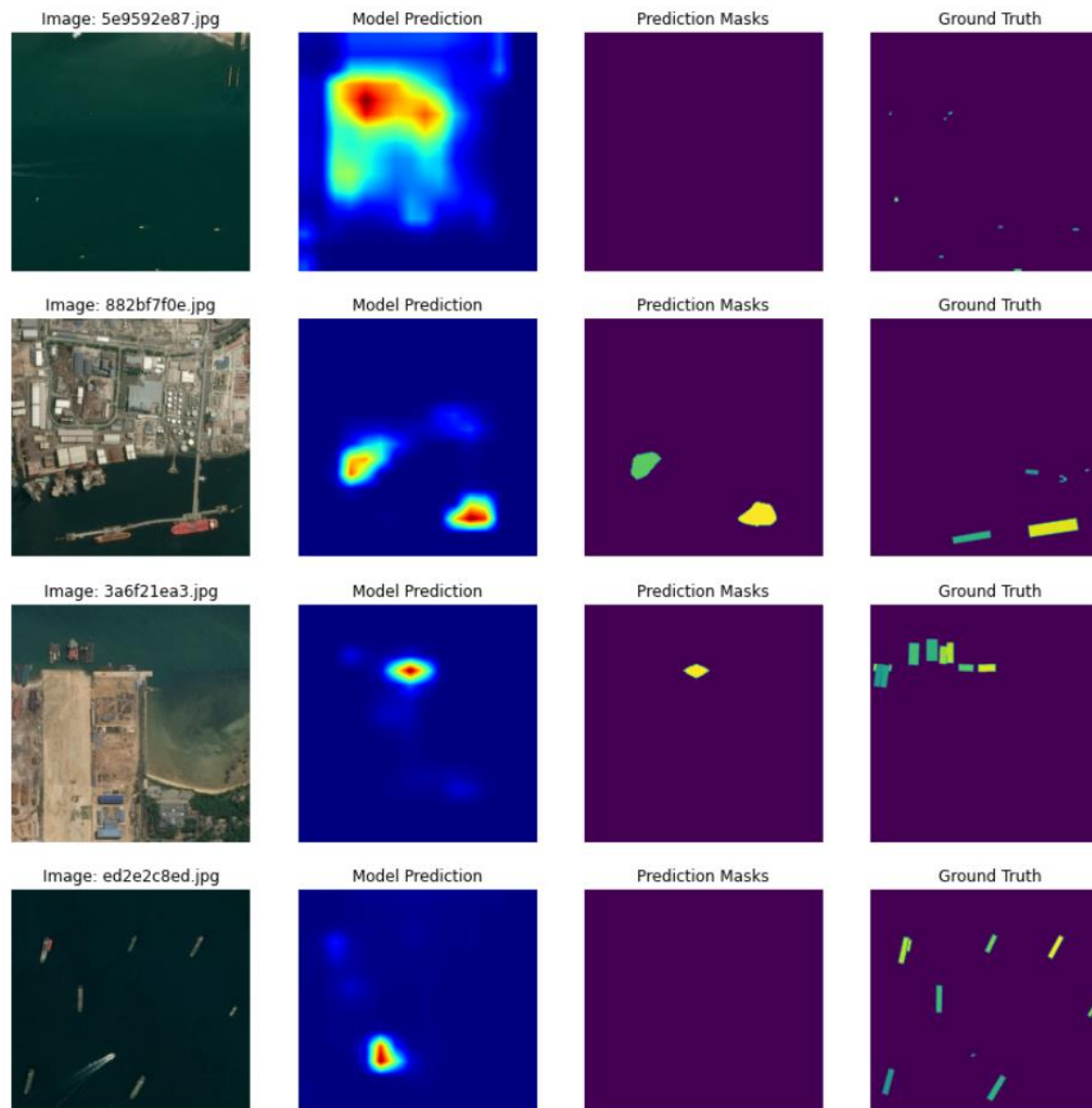
U – Net

Сравнение результатов обработки сетей для сегментации изображений - 2



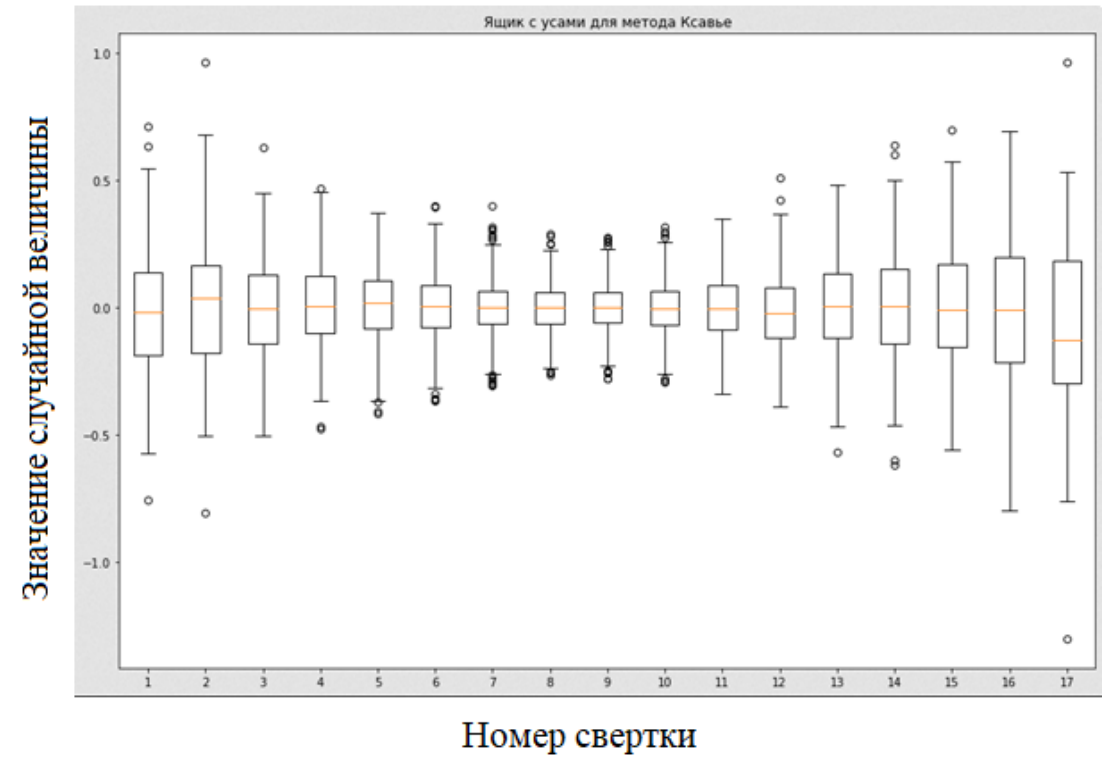
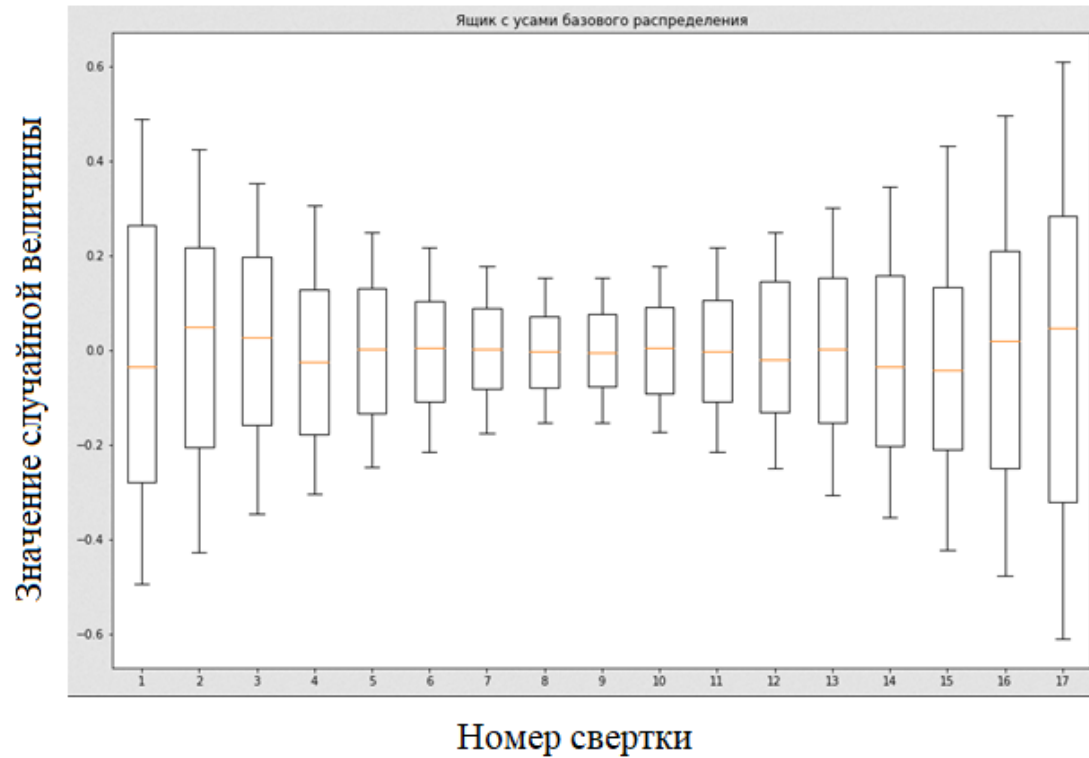
SegNet

Сравнение результатов обработки сетей для сегментации изображений - 3



FCN VGG16

Сравнение инициализации базовым методом (равномерным) и методом Ксавье



$$a = \sqrt{\frac{6}{\text{size}[i-1] + \text{size}[i]}}$$

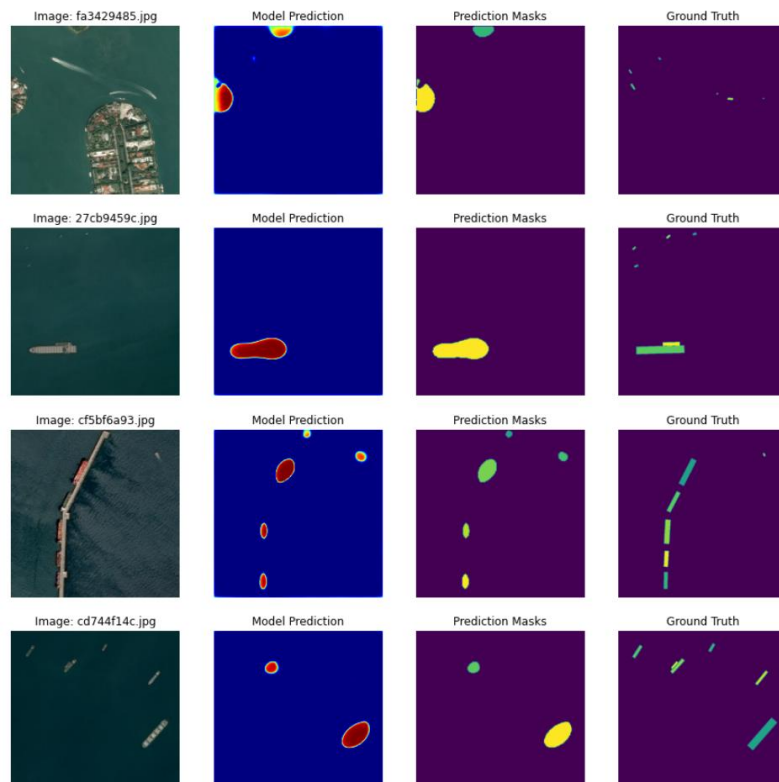
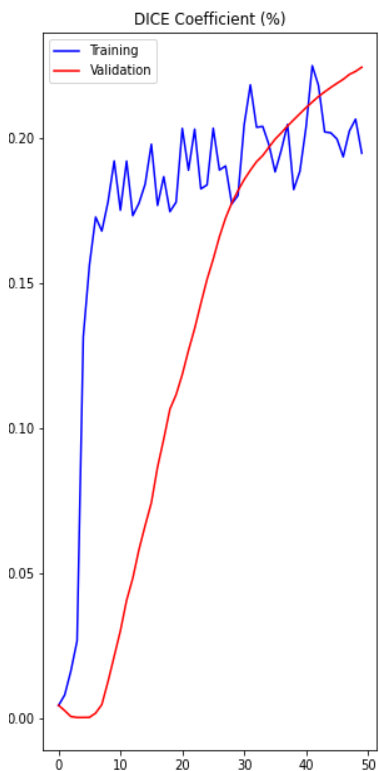
$$\xi = U(-a, a)$$

$$\sigma = \sqrt{\frac{2}{\text{size}[i-1] + \text{size}[i]}}$$

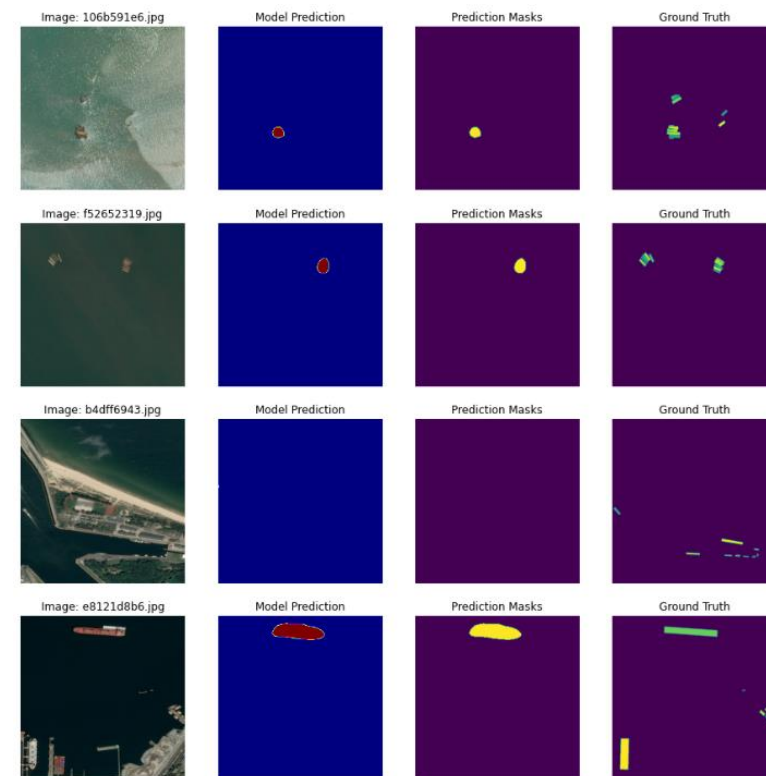
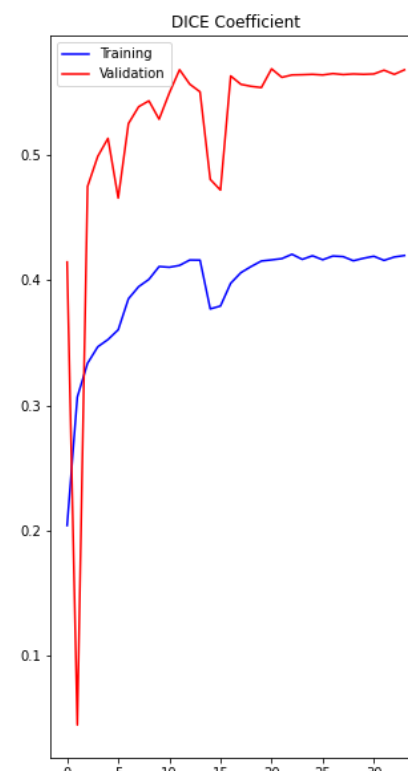
$$\mu = 0$$

$$\xi = N(\mu, \sigma)$$

Результат улучшения и – Net и обучения на большей выборке



После инициализации весов методом Ксавье



После увеличения выборки

Заключение

В данной работе была решена задача распознавания морских судов методами компьютерного зрения. Были приведены результаты работы метода бинарной классификации и метода семантической сегментации изображения.

- Для метода бинарной классификации были подобраны оптимальные гиперпараметры для данной задачи у модели VGG16, что позволило достичь точности классификации в 92% случаев.
- Для метода семантической сегментации была определена лучшая модель – это U – Net и результаты ее обучения для данной задачи были улучшены путем использования метода инициализации весов для слоев сверточной нейронной сети методом Ксавье, что увеличило значение метрики качества с 0,0054 до 0,25. А после увеличения обучающей выборки метрика качества достигла значения 0.56

Спасибо за внимание!