

Отчет по задаче сегментации домашних животных

1. Анализ данных и возможные проблемы

Основные проблемы:

- **Дисбаланс классов:**
- **Разный размер изображений:** Нужно масштабировать все изображения и маски, чтобы они имели одинаковый размер.
- **Качество изображений и масок:** Некоторые изображения могут быть размытыми, а маски — неполными, что усложняет обучение модели.

2. Предобработка данных

Предобработка данных включает несколько этапов:

- **Изменение размера:** Все изображения и маски были уменьшены до размера 128x128 пикселей, чтобы снизить нагрузку на модель и ускорить обучение.
- **Аугментация:** Применяется случайное горизонтальное отражение изображений и масок, чтобы улучшить разнообразие данных и повысить устойчивость модели.
- **Нормализация:** Изображения нормализуются в диапазон $[0, 1]$, а маски корректируются под нулевую индексацию.
- **Буферизация и кэширование:** Для ускорения загрузки данных и помогает быстрее обучать модель.

3. Используемая модель и гиперпараметры

Для решения задачи выбрана архитектура **U-Net**.

Я выбрал модель U-Net, потому что она идеально подходит для задач сегментации изображений. Она была разработана специально для точного выделения объектов на изображениях, даже если данных не так много. В модели есть два основных блока — энкодер (который извлекает признаки) и декодер (который восстанавливает изображение), что позволяет хорошо работать с деталями на разных уровнях. U-Net также хорошо справляется с изображениями, где объекты могут быть разных размеров и форм, что важно для сегментации домашних питомцев.

Гиперпараметры:

- Количество фильтров в слоях меняется от 64 до 512.
- Для предотвращения переобучения используется **Dropout** (0.3).
- В скрытых слоях используется активация **ReLU**, в выходном — **Softmax**.

4. Оценка модели

Для оценки качества работы модели использовалась метрика **accuracy**. Однако для задач сегментации более подходящими являются метрики **IoU** или **Dice coefficient**, которые лучше показывают, насколько точно модель выделяет объекты. Из-за ограничений видеопамати этих метрик в расчетах не использовал.

5. Визуализация предсказаний

- Оригинальное изображение.
- Истинную маску.
- Предсказанную моделью маску.

Это позволяет визуально сравнивать, насколько хорошо модель сегментирует объекты.

6. Проблемы с аппаратными ресурсами

Из-за ограничений видеопамати возникли следующие проблемы:

- **Размер изображений:** Не удалось протестировать модель на изображениях размером 256x256, поскольку видеокарта не поддерживает такие большие данные. В итоге использовался размер 128x128.
- **Метрики сегментации:** Из-за тех же ограничений не было возможности использовать более сложные метрики (IoU, Dice coefficient), что могло бы улучшить точность сегментации.

Заключение

Модель **U-Net** показала хорошие результаты с изображениями размером 128x128. Для улучшения качества сегментации бы попробовал:

- Использовать изображения размером 256x256, 512x512 при наличии более мощной видеокарты.
- Применить метрики **IoU** и **Dice coefficient** для более точной оценки.
- Использовать дополнительные методы аугментации и регуляризации для улучшения обобщающих способностей модели.