

# Краткий итоговый отчет по экспериментам в задаче сегментации построек Майя по спутниковым снимкам

Виктор Шляхин

## Задача

Предлагается сегментировать постройки Майя на разных участках местности размером 240 на 240 метров. Для каждого такого участка имеется снимок с лидара размера 480x480 + снимок со спутника sentinel1 размера 24x24 + снимок со спутника sentinel2 размера 24x24. В снимке с лидара находится всего три канала, как в обычном трехцветном снимке.

## Метрика

За основную метрику в исследовании было взято среднее значение Intersection over Union (IoU) на валидационной выборке.

## Эксперименты

В процессе исследований были проведены следующие эксперименты:

1. Построен бейзлайн на основе предобученной модели DeepLabV3 (последний слой изменен для выдачи масок трех классов) с оптимизатором Adam, learning rate = 0.001 и шедюлером learning rate OneCycleLR. Размер батча - 2. В качестве функции потерь применялось среднеквадратичное отклонение. В дальнейших экспериментах использовались такие же параметры. Модель обучалась 5 эпох.

Таблица 1. Показатели качества бейзлайна при разных функциях потерь на 3 эпохах обучения:

Функция потерь	Mean IoU
BCE	0.6880
MSE	0.7012
Dice	0.4334
Focal	0.6041

2. В бейзлайн добавлена аугментация тренировочных и валидационных картинок путем произвольного ресайза изображений в пределах [125, 500] и дальнейшего рандомного обрезания в размер 250\*250. Это увеличивает вариативность выборки и позволяет модели обучаться лучше. Обучение происходило на 5 эпохах.

Далее эксперименты были направлены на использование при обучении данных со спутников sentinel1 и sentinel2.

3. Стандартное трехканальное изображение с лидара было поточечно геометрически усреднено в один канал, в качестве второго и третьего канала были взяты 97 и 103 каналы с sentinel1 (медианы по VV и VH диапазонам за 2017-2020 годы). Полученные трехканальные изображения использовались для 5 эпох тренировки DeepLabV3.

4. Первый слой backbone DeepLabV3 был расширен на 3 входящих канала. Веса добавленных каналов были инициализированы из равномерного распределения в соответствии со стандартными рекомендациями по такой инициализации. На вход модели подавались шестиканальные картинки, содержащие 3 rgb канала, а также 97 и

103 канала с sentinel1 (медианы по VV и VH диапазонам за 2017-2020 годы) и канал 0 со спутника sentinel2. Обучение длилось 5 эпох.

5. Были дообучены две предобученные модели DeepLabV3 (последний слой изменен для выдачи масок трех классов) – первая на снимках лидар, вторая на 97 и 103 каналах с sentinel1 (медианы по VV и VH диапазонам за 2017-2020 годы) и канале 0 со спутника sentinel2. Выходы моделей после применения сигмоиды были геометрически усреднены. Обучение длилось 5 эпох.

Таблица 2. Сводные результаты экспериментов.

	Эксперимент	Mean IoU
1	Бейзлайн	0.68
2	Бейзлайн с аугментацией	0.82
3	Усреднение rgb в один канал и добавление 2 каналов sentinel1	0.81
4	Расширение входящего слоя модели до 6 каналов	0.83
5	Обучение двух моделей	0.81

#### Выводы

Наилучший результат по метрике IoU показало расширение входящего слоя модели до 6 каналов. Однако, количественно улучшение невелико по сравнению с другими подходами, включая аугментированный бейзлайн без использования дополнительных каналов.