

## Проект обработки спутниковых снимков

**Основная цель проекта** – определение площади застройки на аэрофотоснимках.

Так, для решения поставленной задачи была выбрана архитектура модели сегментации U-Net из библиотеки `smp` (`segmentation_models_pytorch`) с предобученным энкодером ResNet50 с весами ImageNet. Реализован самостоятельный цикл обучения (Training Loop) выбранной архитектуры модели на данных **INRIA Aerial Image Labeling Dataset** (<https://project.inria.fr/aerialimagelabeling/contest/>). Важно отметить, что аэрофотоснимки (используемые в Проекте из INRIA Aerial Image Labeling Dataset) содержат встроенные геопространственные метаданные, которые включают информацию о проекции, разрешении пикселя и географической привязке.

Кроме того, следует отметить, что выбрана модель семантической сегментации, так как для решения поставленной задачи желательна пиксельная точность при определении контуров зданий. Площадь застройки рассчитывается как сумма всех пикселей, которые относятся к "зданиям", умноженная на площадь одного пикселя. В случае детекции для каждого здания определяются bounding boxes (включая фон), что приводит к существенному завышению общей площади застройки.

По результатам тестирования модели на тестовой выборке 1 (подготовлена на основе 10% данных INRIA Aerial Image Labeling Dataset из папки train) и тестовой выборки 2 (самостоятельно размеченные данные INRIA Aerial Image Labeling Dataset из папки test без масок) получены значения метрик, представленные в Таблице 1.

Таблица 1. Метрики на тестовых выборках

	IoU	F1-Score	BinaryAccuracy	BinaryRecall	BinaryPrecision
test	0.7471	0.8553	0.9535	0.8872	0.8255
test-2	0.7999	0.8888	0.9767	0.8745	0.9036

Модель демонстрирует относительно неплохое качество работы на двух независимых тестовых выборках, что подтверждает некоторую её надёжность и обобщающую способность. Основные метрики (IoU, F1-Score, BinaryAccuracy, BinaryRecall) имеют сопоставимые значения. Можно отметить небольшой рост BinaryPrecision на тестовой выборке 2, тем самым модель стала реже путать фон со зданиями (тени деревьев, дороги, облака и др.). Самостоятельно размеченные данные, вероятно немного лучше отображают наличие/отсутствие зданий в реальности, учитывая, что данные корректировались с использованием геопространственной привязки в ГИС QGIS. В свою очередь, BinaryRecall довольно высокий на двух выборках, то есть модель пропускает не очень большое количество реальных зданий.

Основные трудности автоматической сегментации зданий связаны с объектами, «текстурно» похожими на крыши:

- тени от объектов и рельефа;
- деревья и растительность с текстурой, похожей на крыши;
- асфальт, дороги, парковки (спектрально близки крышам);
- теплицы и ангары.

Наличие деревьев рядом с домами затрудняет идентификацию зданий: под густой кроной дома часто не идентифицируются моделью. Кроме того, довольно тяжело идентифицируются козырьки, навесы, пристройки и архитектурные элементы. Разнообразие форм крыш (многоскатные, плоские, купола и др.) также затрудняет идентификацию зданий. Также после модели отмечаются более сглаженные края зданий.

Поэтому, крайне важно стремиться к максимальной точности и согласованности при разметке изображений. Неточности могут привести к появлению шума в обучающих данных и ухудшению производительности модели. Повышению точности сегментации, возможно, может помочь изменение архитектуры модели (SegFormer / U-Net++ / DeepLabV3+), или использование при обучении Focal Loss.

Для реализации интерфейса для демонстрации работы модели используется библиотека Gradio. **Масштаб изображения** определяется с помощью библиотеки **rasterio**, которая извлекает встроенные геопространственные метаданные из изображений INRIA Aerial Image Labeling Dataset.

На примере изображения с маской (AerialImageDataset\train\images\vienna8.tif) показан результат работы интерфейса (рисунок 1) и расчет площади застройки из QGIS (рисунок 2). По результатам расчета площадь застройки составляет 1052036.28 м<sup>2</sup>, а в случае QGIS (по доступной маске, которая считается "истинной") – 1044297.27 м<sup>2</sup>. Таким образом, проведенная верификация показывает, что ошибка при расчете площади для выбранного изображения составляет менее 1%. В целом, с учетом того, что в «истинных масках» присутствуют неточности, при расчете общей площади застройки через реализованный интерфейс, ошибки, в основном, не превышают 5% от общей площади застройки.

## Building Segmentation with U-Net

Upload an aerial image (preferably GeoTIFF) to get a segmentation mask and building area calculation.

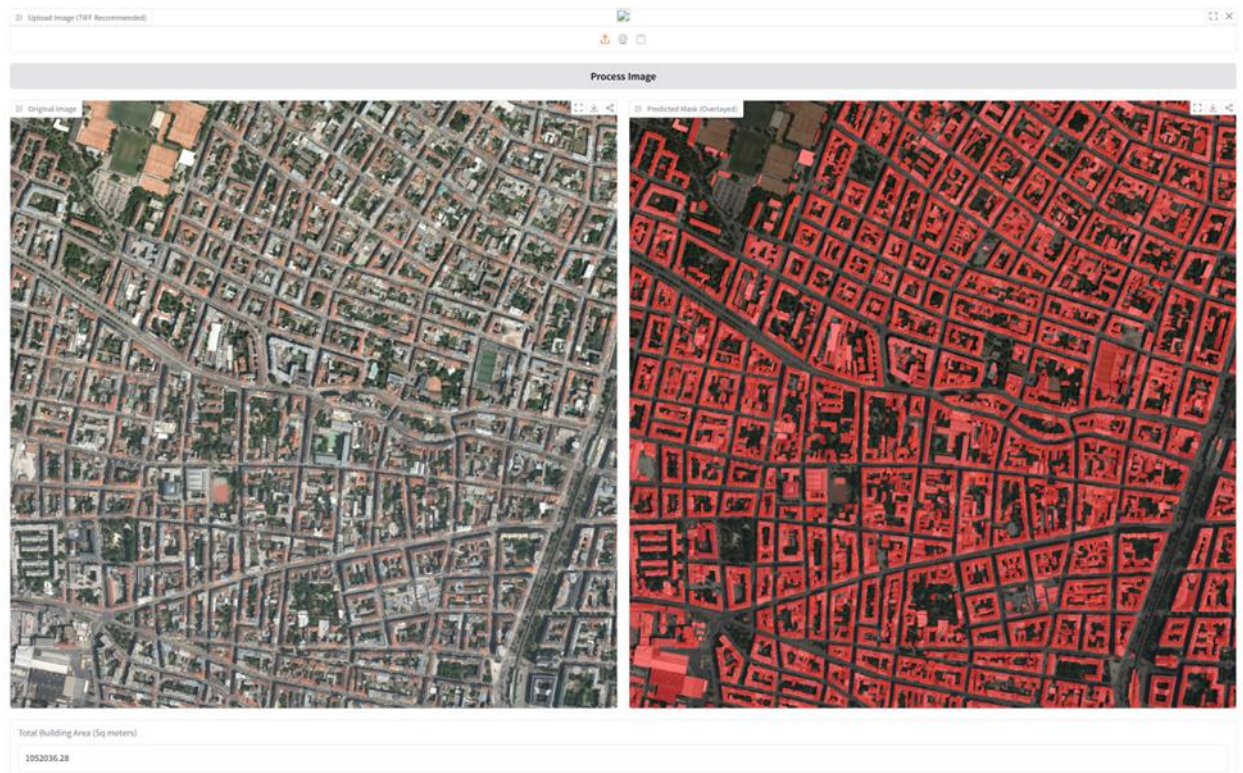


Рисунок 1 – Результат расчета площади застройки через реализованный интерфейс

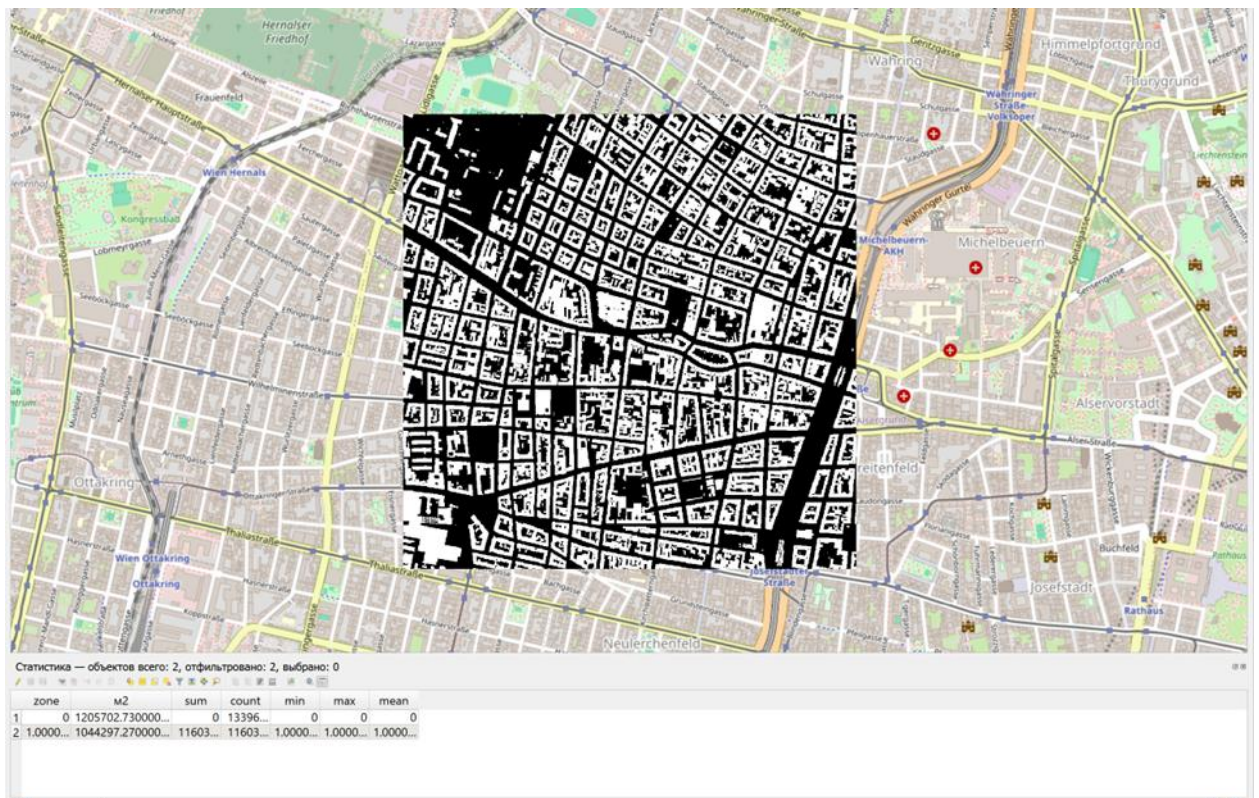


Рисунок 2 – Результаты определения площади по "истинной" маске в QGIS: в таблице представлена площадь для "1" (зданий).