

Нейросетевая сегментация объектов хозяйственной инфраструктуры на спутниковых изображениях высокого разрешения.

Козуб В.А. (1, 2), Гвоздев О.Г. (3, 1), Матвеев И.А. (4), Мурынин А.Б. (4, 1)

(1) НИИ "АЭРОКОСМОС", Москва, Российская Федерация

(2) Московский физико-технический институт (государственный университет), Москва, Российская Федерация

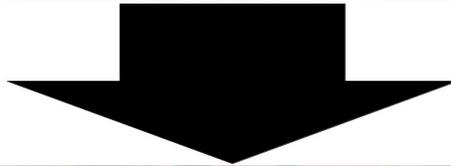
(3) МИИГАиК, Москва, Российская Федерация

(4) Федеральный исследовательский центр "Информатика и управление" РАН (ФИЦ ИУ РАН), Москва, Россия

Работа выполнена при финансовой поддержке Российской Федерации в лице Минобрнауки России в рамках соглашения №075–15–2020–776.

Постановка задачи

Создать нейросетевую модель, определяющую принадлежность каждого пикселя исходного изображения к одному или нескольким классам (для заданного набора классов).



Назначение

- Задача решается в рамках комплексной методики построения трехмерных цифровых моделей местности хозяйственной инфраструктуры.
- Методика ориентирована на жилищную, железнодорожную и автотранспортную составляющие хозяйственной структуры.
- Построение производится по одному изображению и метаданным съемки.

Применение:

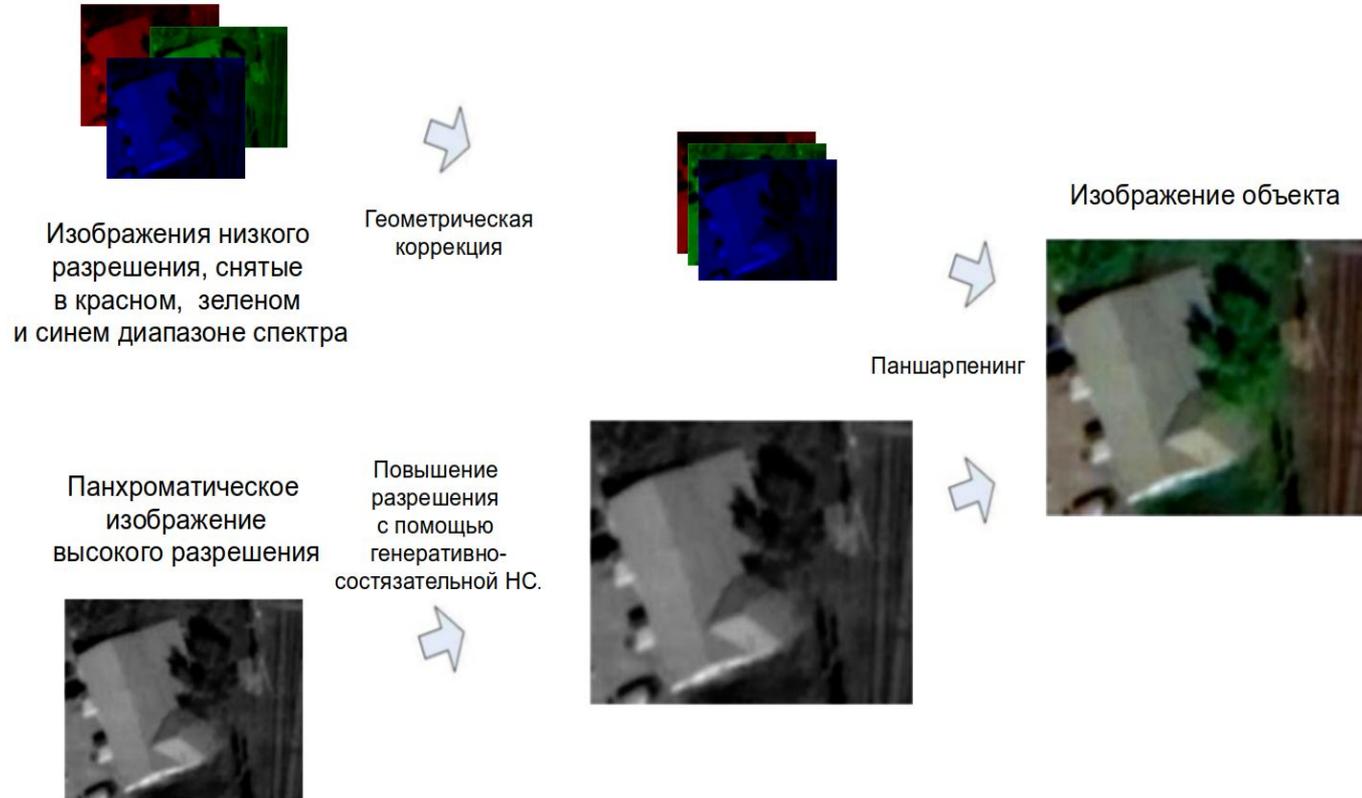
- оценка населенности района
- моделирование распространения акустических и радиоволн
- создание более удобных для пользователей ГИС

Методика восстановления трехмерной формы объектов.

Данные с КА ДЗЗ Ресурс-П:

- разрешение R, G, B каналов:
2-3 м/пк

- разрешение панхроматической
съемки: 0.7 м/пк



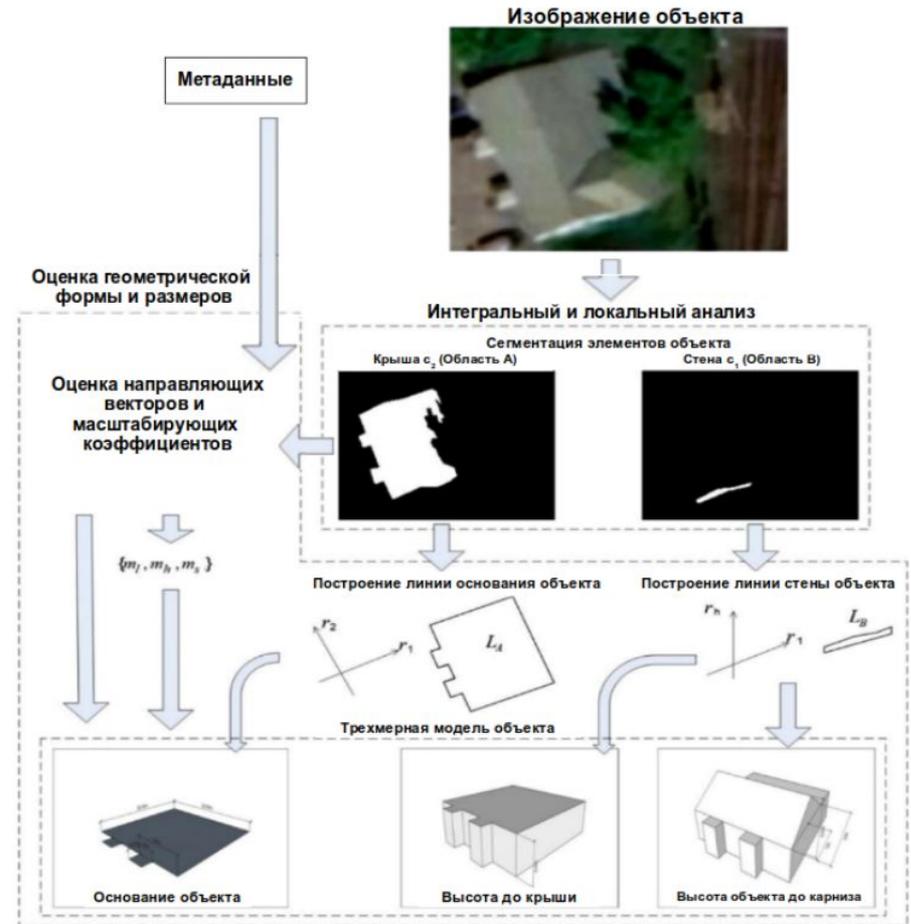
Первый этап: комплексирование данных.

Методика восстановления трехмерной формы объектов.

Второй этап: сегментация.

Третий этап: оценка геометрической формы и размеров:

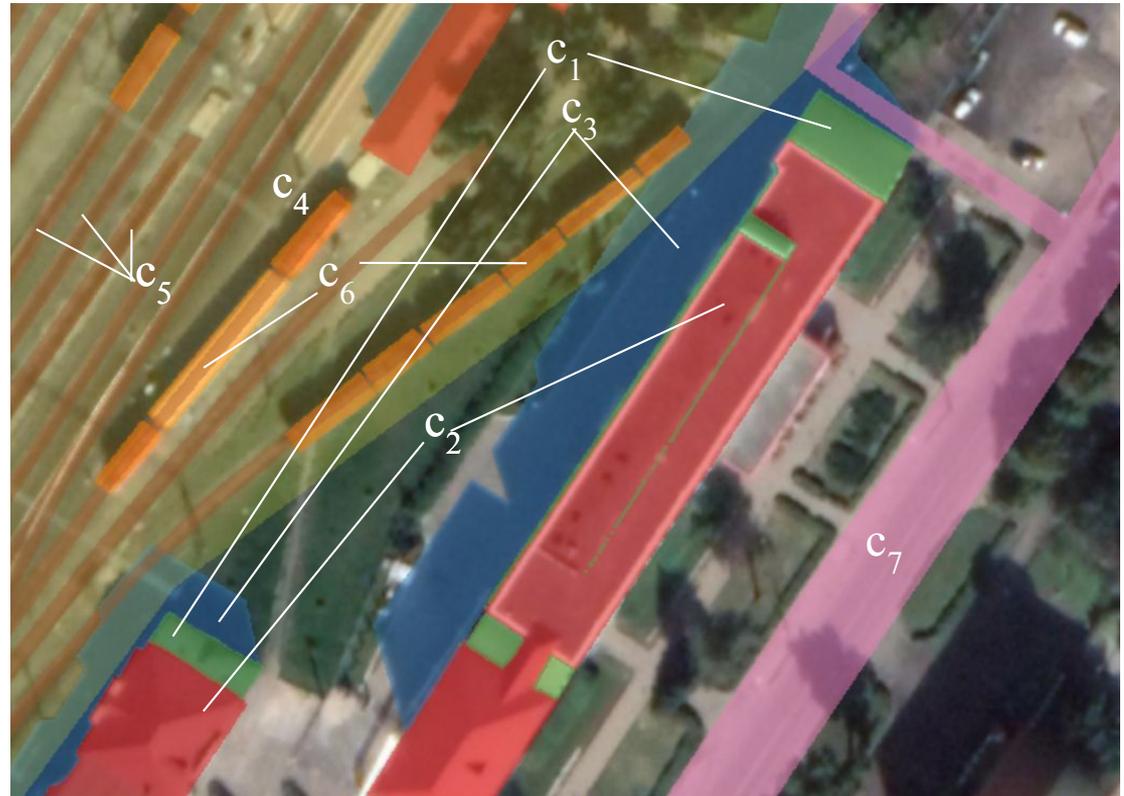
- Оценка масштабирующих коэффициентов (Горизонтальный, вертикальный и теневой)
- Оценка направляющих векторов (вертикальное направление и направление падения тени).
- Восстановление графа автомобильных и железных дорог
- Векторизация контуров здания, оценка высоты и моделирование формы крыши. В зависимости от имеющихся данных (качества сегментации и угле съемки) возможен уровень детализации от LoD0 до LoD2.



Рассматриваемые классы

Целевые классы:

- Стены здания (c_1)
- Крыши зданий (c_2)
- Тени здания (c_3)
- Область ЖД инфраструктуры (c_4)
 - Рельсы (c_5)
 - Вагоны (c_6)
- Автомобильная дорога (c_7)



Крыши

Тени

Стены

ЖД пути

ЖД
инфрастр.

Дороги

Вагоны

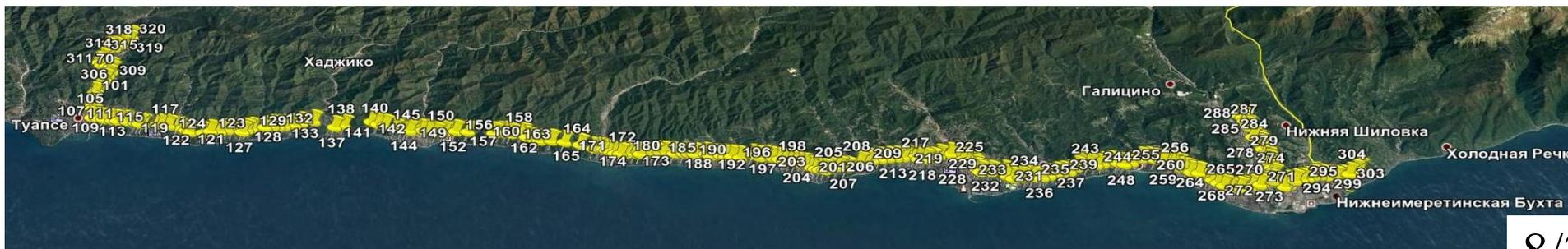
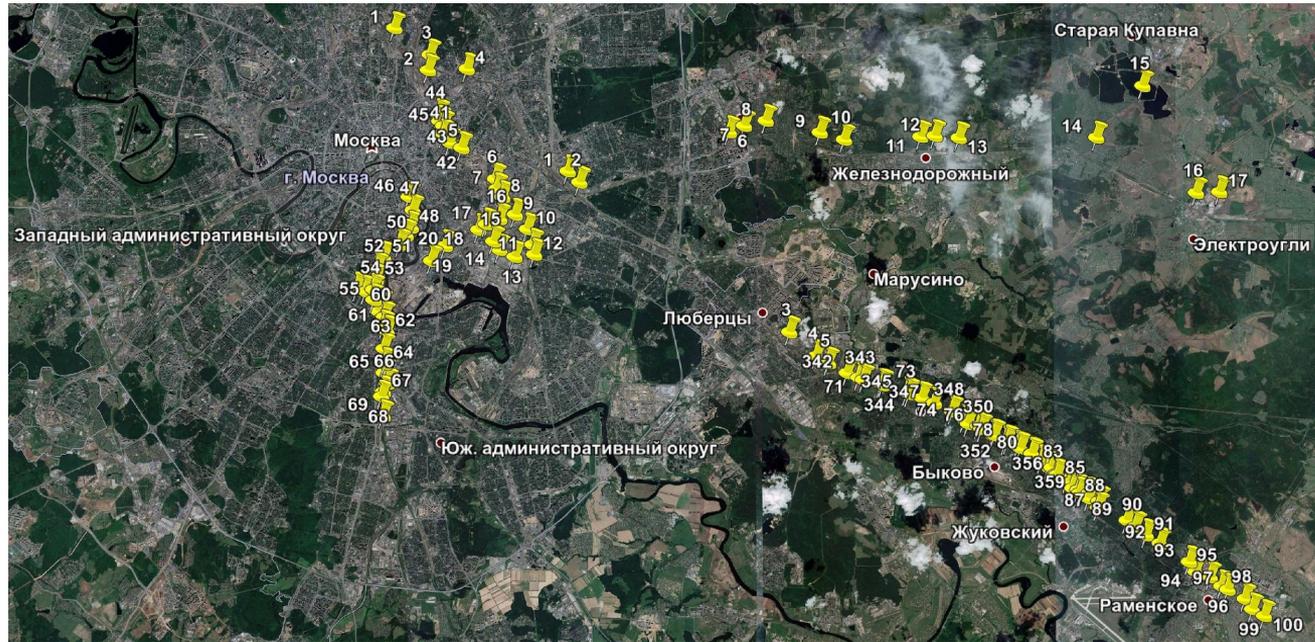
Особенности задачи

Среди особенностей сегментации искусственных объектов на спутниковых снимках можно выделить:

- Произвольная ориентация объектов.
- Нужно учитывать контекст.
- Разнообразие форм.
- Наличие теней.
- Работа в большом диапазоне географических и погодных условий.

Обучающая выборка

- 502 размеченных вручную снимка
- Открытые данные из ГИС Google Earth.
- Снимки были загублены до разрешения 35 см/пк.
- География:
 - Московская область
 - Тульская область
 - ЖД-пути Туапсе-Адлер.



Дисбаланс классов

- Обучающие классы не сбалансированы
- Частично сбалансировать классы получилось за счет специального метода аугментации.

Класс	Крыши зданий	Тени зданий	Стены зданий	Автомобильные дороги	ЖД пути	вагоны	ЖД инфраструктура	Фон
Мощность класса в % (до аугментации)	5.86	2.07	0.63	5.53	1.99	0.47	11.17	72.28
Мощность класса в % (после аугментации)	11.27	4.07	1.72	6.42	3.95	1.30	20.13	51.17

Аугментация

N — кол-во эпох аугментации. Для i -ой эпохи выполняем:

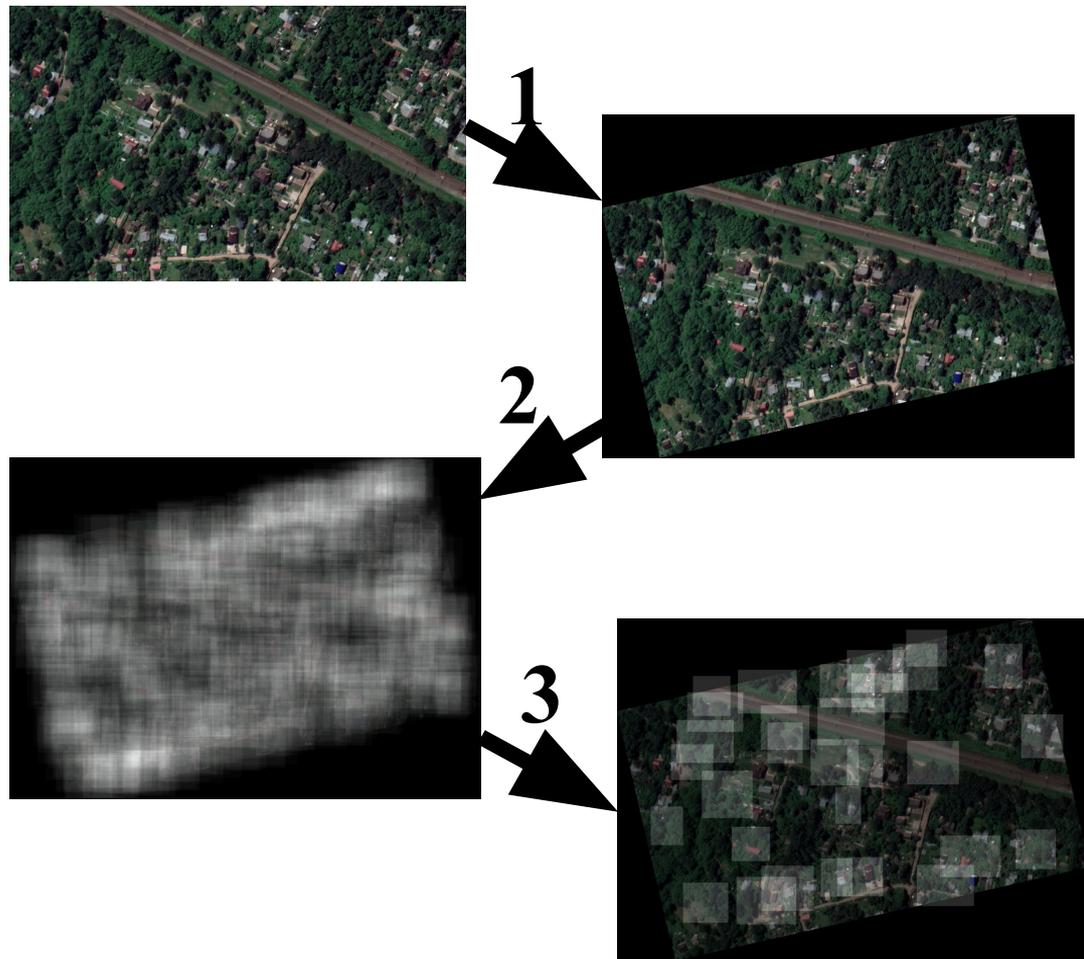
1. Поворот на случайный угол в диапазоне от $360i/N$ до $360(i+1)/N$ градусов.
2. Сэмплирование S случайных квадратных областей в заданном диапазоне размеров. Сэмплирование происходит по одному изображению за раз. Для каждого сэмпла проверяется доля информативных пикселей. Для каждой сэмплированной s области подсчитывается ее значимость:

$$I_s = \sum_{k=1}^C \frac{p_k}{Pr_k^\gamma}$$

где p_k - процент пикселей из класса k в сэмпле s , а Pr_k - процент пикселей класса k во всей обучающей выборке, γ — положительно число.

3. Случайными образом, с вероятностями пропорциональными значимостям выбираются M сэмплов.

4. Выбранные сэмплы масштабируются к одному размеру и добавляются к обучающему набору.



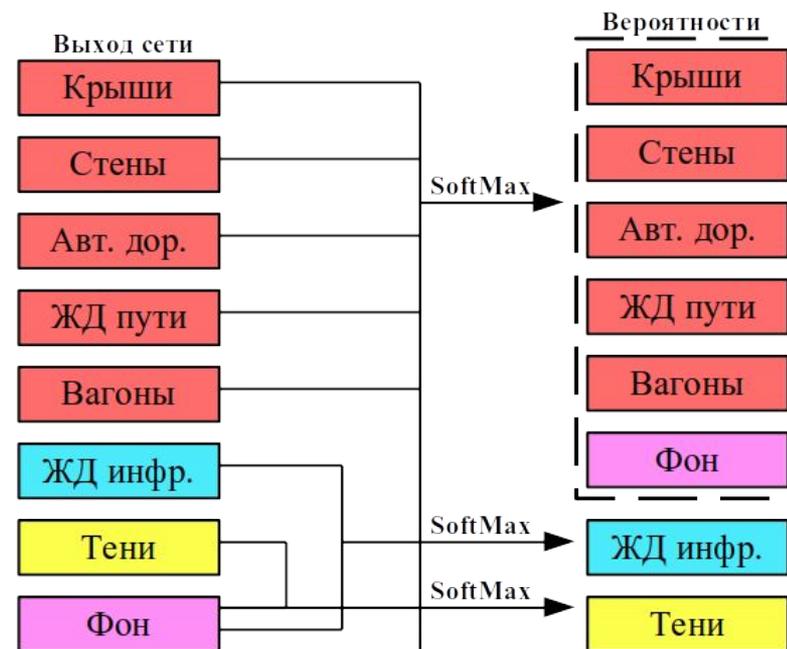
Нейросетевая архитектура

1	Вход	3	25	SoftMax	c
2	3x3 свертка	48	24	1x1 свертка	c
3	Dense блок (4 слоя)	112	23	Dense блок (4 слоя)	256
4	Понижающий блок	112	22	Повышающий блок	384
5	Dense блок (5 слоев)	192	21	Dense блок (5 слоев)	384
6	Понижающий блок	192	20	Повышающий блок	578
7	Dense блок (7 слоев)	304	19	Dense блок (7 слоев)	578
8	Понижающий блок	304	18	Повышающий блок	816
9	Dense блок (10 слоев)	464	17	Dense блок (10 слоев)	816
10	Понижающий блок	464	16	Повышающий блок	1088
11	Dense блок (12 слоев)	656	15	Dense блок (12 слоев)	1088
12	Понижающий блок	656	14	Повышающий блок	896
13	Dense блок (15 слоев)	896			

- тип слоя

 - кол-во каналов

- порядковый номер слоя



Обучение

Функция потерь – Focal Loss:

$$FL(P) = \frac{1}{CHW} \sum_{i,j,k=1,1,1}^{C,H,W} w_i T_{i,j,k} (1 - P_{i,j,k})^\gamma \ln(P_{i,j,k})$$

где C – кол-во классов, H , W – размеры изображения, T – ответы, P – предсказанные вероятности, w – веса классов, γ – неотрицательный параметр.

Подбор порогов:

$$TPR = \frac{TP}{P} \quad FPR = \frac{FP}{N} \quad TP + FP = P$$

$$F_1 - score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} = \frac{2 \cdot TPR \cdot P}{TPR \cdot P + (FPR \cdot N + P)} \quad Thresholds = \operatorname{argmax}(F_1 - score)$$

где TP – истинно-положительные ответы, FP – ложно-положительные ответы, FN – ложно-отрицательные ответы, P – кол-во пикселей объекта, N – кол-во пикселей фона.

Подбор параметров: Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$, lr_decay = 0.9995, lr = $4 \cdot 10^{-4}$ (первые 30 эпох, затем еще 10 эпох с начальным lr = $1 \cdot 10^{-5}$))

Вычисления выполнялись с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» ФИЦ ИУ РАН (ЦКП «Информатика»)

Оценка качества полученных результатов

F1-score

Класс \ Нейросеть	Крыши зданий	Тени зданий	Стены зданий	Автомобильные дороги	ж/д пути	Вагоны	ж/д инфра- структура
MultiResUnet [4]	0.750	0.703	0.575	0.623	0.582	0.554	0.750
FCDenseNet-103 [5]	0.778	0.704	0.578	0.680	0.606	0.588	0.794
предл. арх.	0.793	0.713	0.570	0.702	0.607	0.584	0.823

IoU

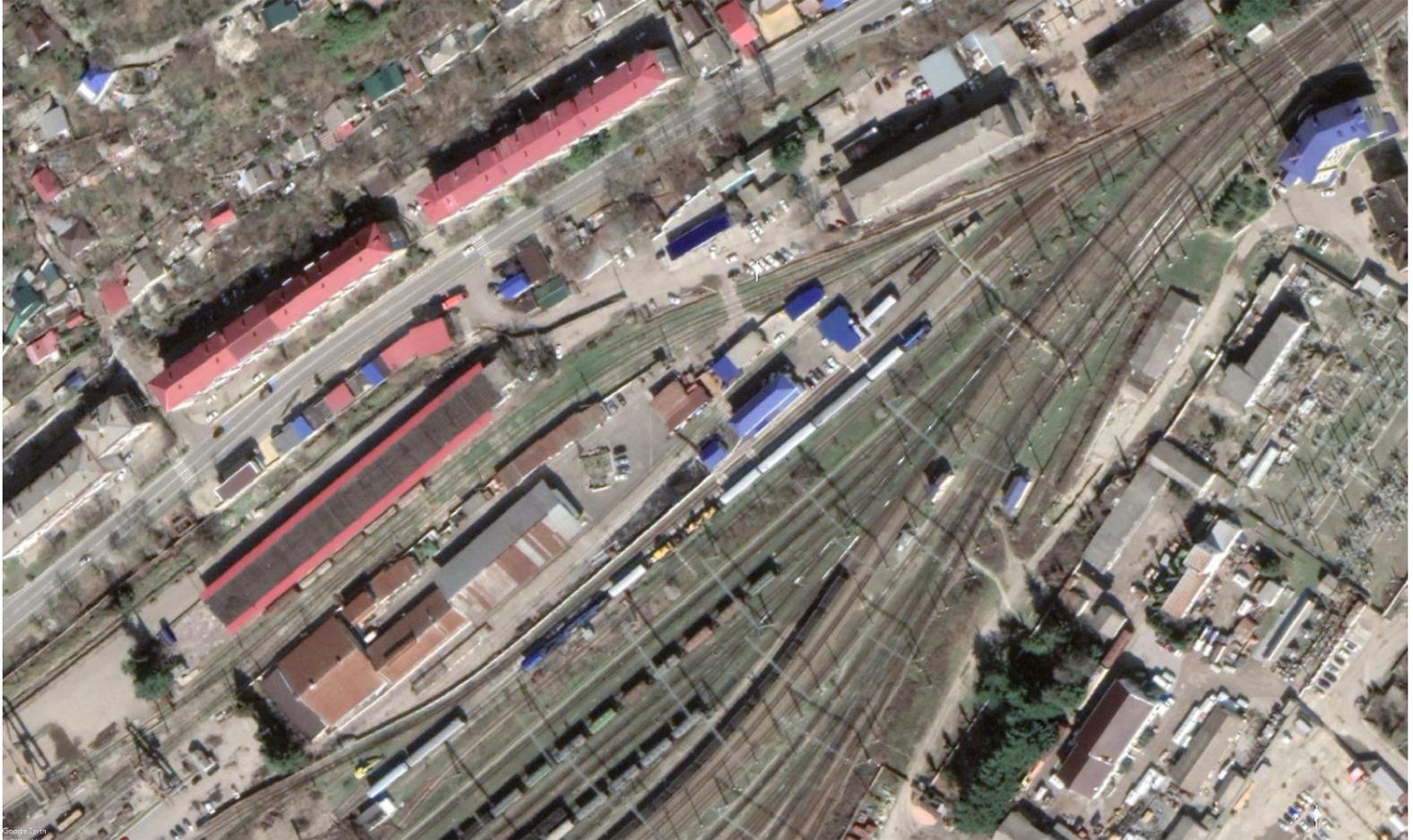
Класс \ Нейросеть	Крыши зданий	Тени зданий	Стены зданий	Автомобильные дороги	ж/д пути	Вагоны	ж/д инфра- структура
MultiResUnet [4]	0.636	0.559	0.414	0.514	0.442	0.420	0.712
FCDenseNet-103 [5]	0.663	0.558	0.418	0.527	0.461	0.433	0.741
предл. арх.	0.674	0.564	0.411	0.555	0.455	0.430	0.762

$$F_1 - score = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

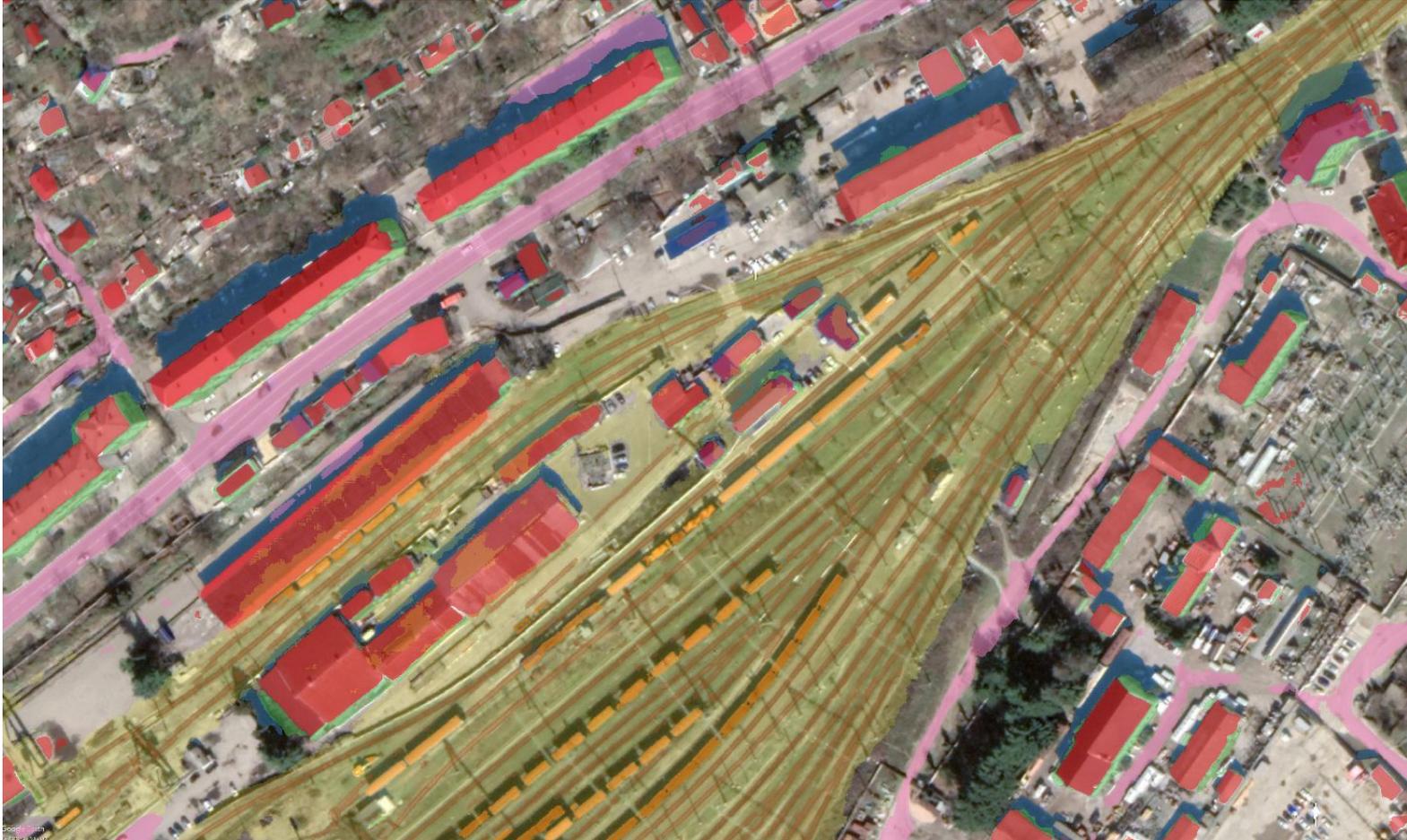
$$IoU = \frac{TP}{TP + FN + FP}$$

TP – истинно-положительные ответы, *FP* – ложно-положительные ответы,
FN - ложно-отрицательные ответы,

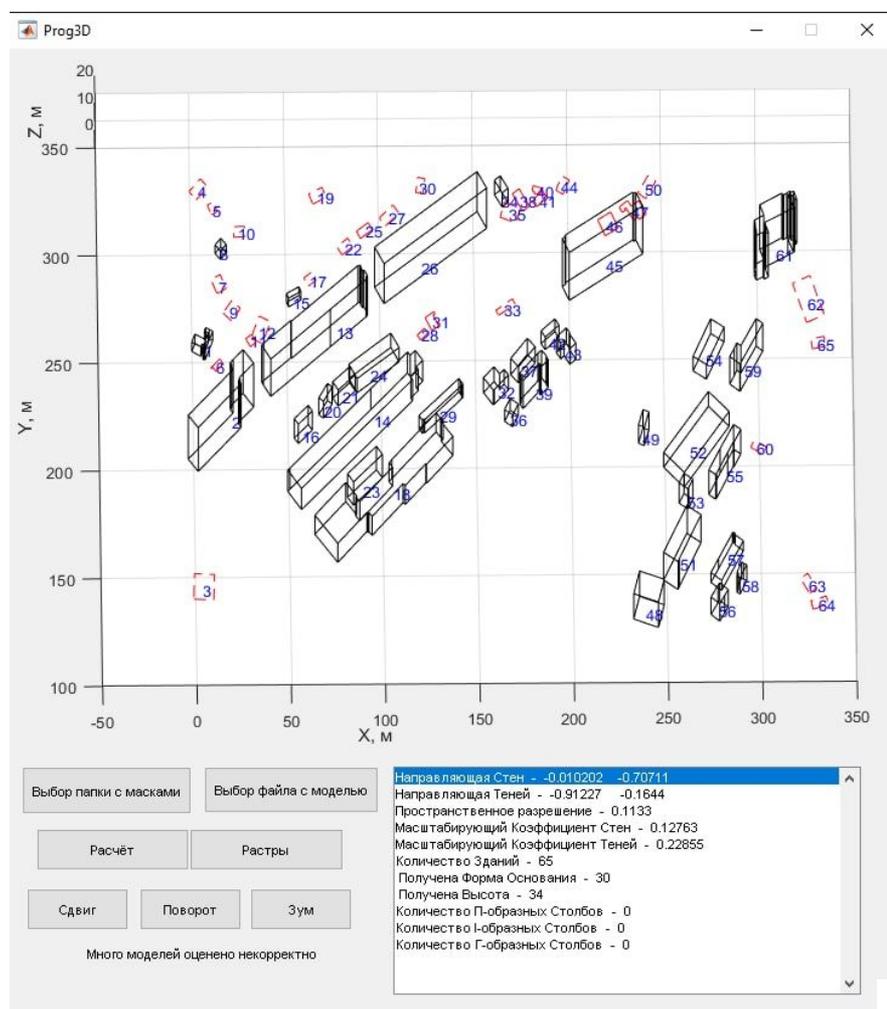
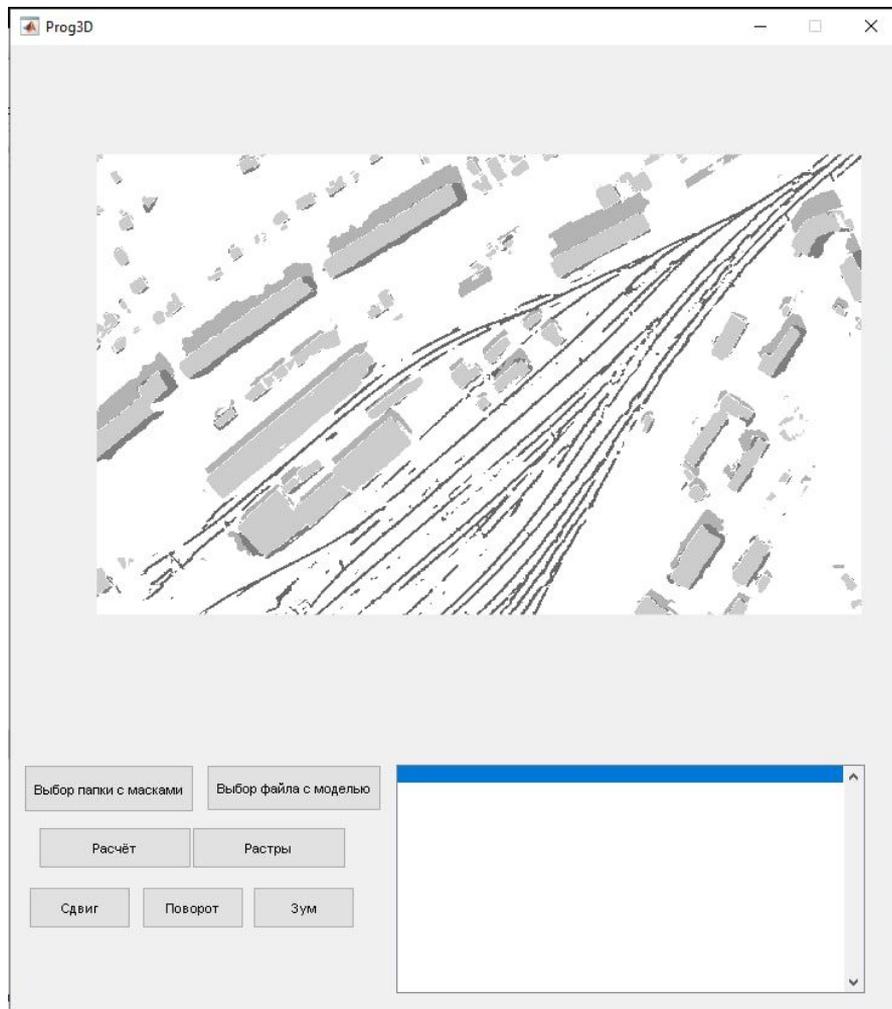
Пример работы макета программного комплекса



Пример работы макета программного комплекса



Пример работы макета программного комплекса



Выводы

Рассмотрена задача семантической сегментации, в рамках методики восстановления трехмерных цифровых моделей ригидных объектов хозяйственной инфраструктуры по одному спутниковому изображению.

Подготовлена обучающая выборка основанная на открытых данных геоинформационной системы Google Earth.

Проведена аугментация данных с применением разработанного метода вероятностной аугментации, позволившего частично сбалансировать выборку.

Обучены для задачи сегментации различные архитектуры НС. Произведена оценка качества результатов.

Дальнейшая развитие

- Работа с данными Ресурс-П.
- Использование более современных моделей с предобученными весами.
- Использование Deep WaterShed подхода для пообъектной сегментации зданий.
- Исследования возможностей применения для мониторинга импактных районов за полярным кругом.

Спасибо за внимание!