

Использование методов глубокого обучения для анализа изображений аэрокосмического мониторинга в интересах обнаружения элементов новизны на местности

аспирант 1 курса Отырба Р. Р.
д. т. н., профессор Сирота А.А.
к. т. н., доцент Дрюченко М. А.

Воронеж
ВГУ - 10.02.2022

Описание проблемы

Обнаружение изменений – это аналитический процесс, целью которого является обнаружения новых или исчезнувших объектов на изображениях, зарегистрированных в разные моменты времени для одной и той же целевой области.

Описание проблемы

- Классический ручной поиск элементов новизны выполняется с использованием инструментальных методов при непосредственном участии оператора-интерпретатора изображений, что является трудоёмким и неточным процессом при частых изменениях.
- Известных алгоритмических и программных решений, касающихся обнаружения изменений на местности, существует не так много.

Описание проблемы

Большинство решений:

- Слишком сильно полагаются на геопространственную информацию, предоставляемую с изображениями;
- Ограничены конкретными топографическими объектами и геометрическими аспектами;

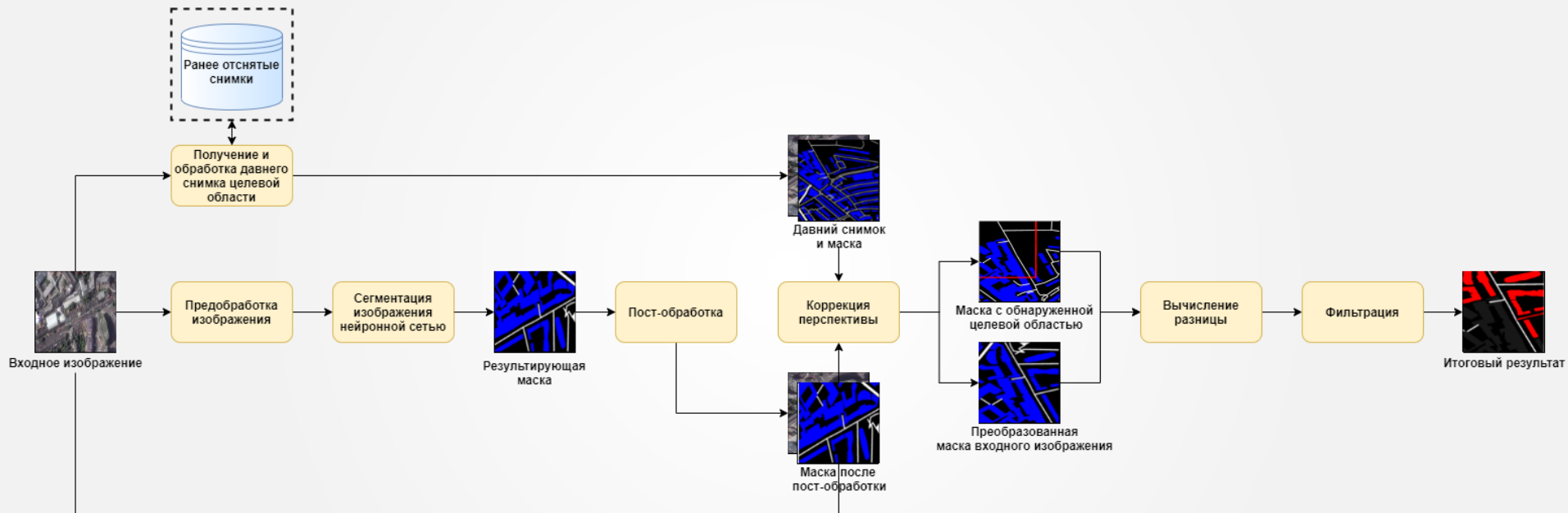
Актуальность

Актуальность обусловлена необходимостью эффективного решения множества различных прикладных задач для мониторинга и понимания изменений на местности, причём оперативно и с наименьшими затратами.

Цель работы

Целью данного исследования является разработка математического и программного обеспечения в интересах создания системы аэрокосмического мониторинга изменений на местности с использованием методов глубокого обучения.

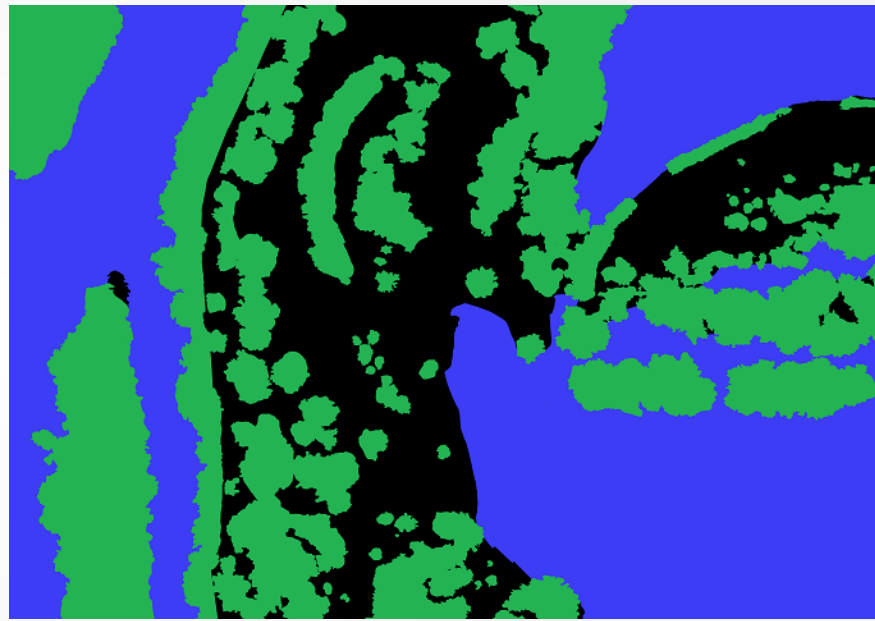
Предлагаемый алгоритм функционирования системы



Подготовка набора данных для обучения

- Было осуществлено создание собственных наборов данных, посредством ручной разметки снимков высокого разрешения (6792x4800) Google Earth Pro в онлайн сервисе CVAT;
- Была произведена нарезка снимков на фрагменты размером 512x512 с перекрытием на 100 пикселей для учёта контекста;
- Итоговый состав набора данных:
 - *Здания и дороги*: 10 снимков (~2000 размером 512x512)
 - *Растительность и вода*: 5 снимков (~1000 размером 512x512)

Подготовка набора данных для обучения



Примеры растровых и размеченных снимков

Подготовка набора данных для обучения

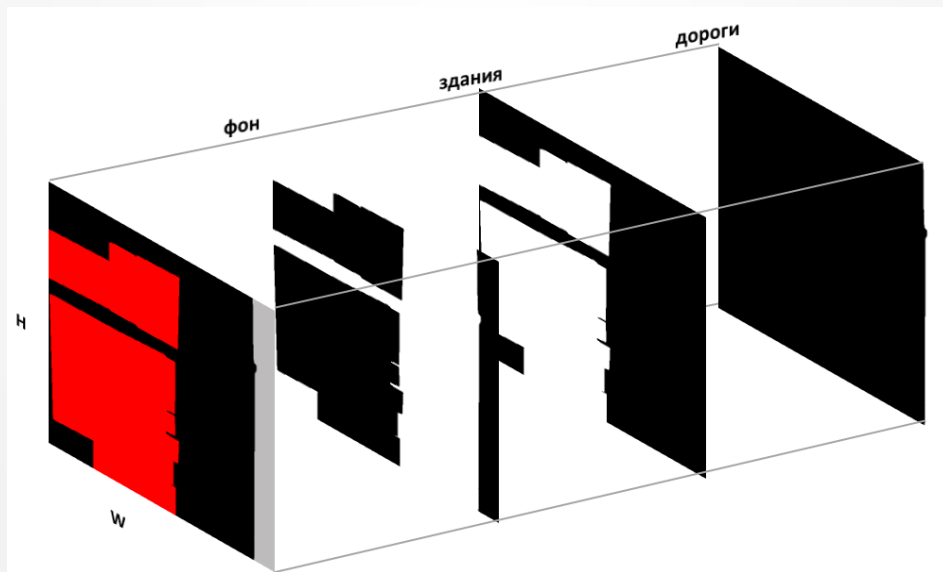
- *Аугментация данных:*
 - Горизонтальные и вертикальные отражения;
 - Случайные повороты на 90 градусов;
 - Масштабирование с случайными поворотами;
 - Уменьшение контрастности (от 5% до 20%);
 - Изменение цветового тона и насыщенности (до 20 %).

Подготовка набора данных для обучения

- *Нормализация растровых данных:*

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

- *One-Hot кодирование размеченных данных:*



Обучение нейронных сетей

Модели	Кодирующие сети
<ul style="list-style-type: none">• FCN• U-Net• SegNet• PSPNet• FPN• SpatialOCRNet• DeepLabV3+• DANet• ACFNet• Pix2Pix (cGAN)	<ul style="list-style-type: none">• DenseNet201,• ResNet101v2• ResNet152v2• EfficientNetb3• EfficientNetb5• EfficientNetb7

Кодирующие сети предобучены на крупном наборе данных ImageNet

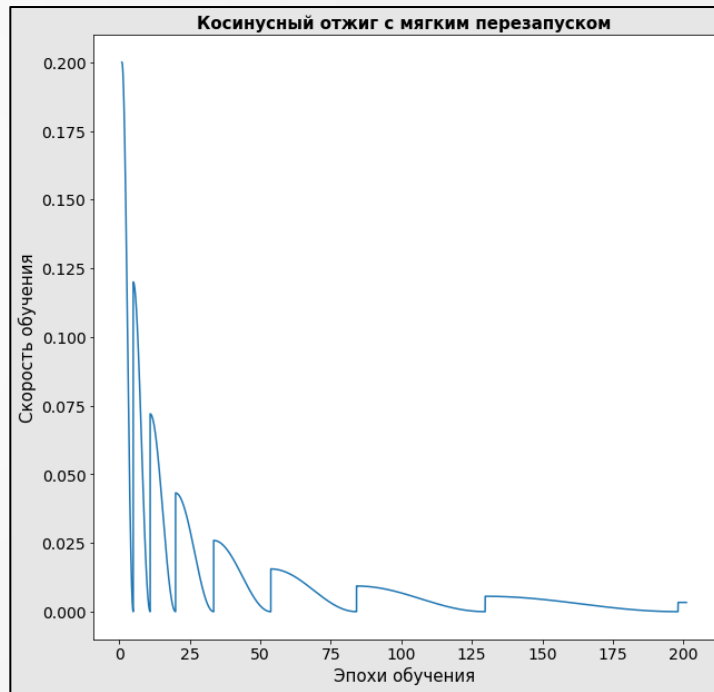
Обучение нейронной сети

- **Оптимизатор:** Момент Нестерова

$$\Delta_{p+1} = \gamma \cdot \Delta_p + \eta \cdot \nabla L(w_p - \gamma \Delta_p)$$
$$w_{p+1} = w_p - \Delta_{p+1}$$

- **Схема отжига скорости обучения:**

$$\eta_t = \eta_{\min}^i + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i) \cdot \left(1 + \cos\left(\frac{T_{\text{cur}}}{T_i} \pi\right)\right);$$



Обучение нейронной сети

- **Метрика:** IoU (Intersection-over-Union)

$$IoU = \frac{1}{N} \cdot \sum_{c=0}^C \omega(c) \cdot \frac{I(c)}{U(c)}$$

где:

$$w(c) = \frac{1}{C} \cdot \frac{\sum_{c=1}^C r(c)}{r(c)} - \text{весовой коэффициент класса};$$

$$r(c) = \frac{\sum_{f=n}^F \sum_{n=1}^N g_n(c)(f)}{\sum_{f=n}^F \sum_{n=1}^N g_n(f)} - \text{коэффициент соотношения пикселей класса к общему количеству пикселей}$$

$$I(c) = \sum_{n=0}^N p_n(c) \cdot g_n(c) - \text{область пересечения};$$

$$U(c) = \left(\sum_{n=0}^N p_n(c) + g_n(c) \right) - \text{область объединения};$$

Обучение нейронной сети

- **Функция ошибки:** Функция Дайса и Фокальная ошибка

$$L = L_{Dice} + \lambda L_{Focal}$$

$$L_{Dice} = 1 - \frac{1}{N} \cdot \sum_{c=0}^C \omega(c) \frac{(1 + \beta)^+ TP_p(c)}{(1 + \beta) \cdot TP_p(c) + \beta^2 \cdot FN_p(c) + FP_p(c)}$$

$$L_{Focal} = - \frac{1}{N} \sum_{c=0}^C \sum_{n=1}^N g_n(c) \cdot (1 - p_n(c))^2 \cdot \log(p_n(c))$$

где:

$$TP_p(c) = \sum_{n=0}^N p_n(c) \cdot g_n(c) - \text{положительный класс распознан как положительный}$$

$$FN_p(c) = \sum_{n=0}^N (1 - p_n(c)) \cdot g_n(c) - \text{отрицательный класс распознан как отрицательный}$$

$$FP_p(c) = \sum_{n=0}^N p_n(c) \cdot (1 - g_n(c)) - \text{отрицательный класс был распознан как положительный}$$

Результаты обучения нейронных сетей

Здания и дороги

Модель	Loss	mIoU
FCN (DenseNet201)	0.29930	0.8208
U-Net (Classic)	0.37015	0.6571
SegNet (Classic)	0.35474	0.6680
PSPNet (DenseNet201)	0.35753	0.6565
FPN (DenseNet201)	0.31262	0.7527
DANet (DenseNet201)	0.29077	0.7636
OCRNet (DenseNet201)	0.21951	0.8806
ACFNet (DenseNet201)	0.27441	0.8010
DeepLabV3+ (DenseNet201)	0.22694	0.8582
Pix2Pix	0.27461	0.7621

Густая растительность и вода

DeepLabV3+ (DenseNet201)	0.30472	0.7576
SpatialOCRNet (DenseNet201)	0.29729	0.7789
Pix2Pix	0.31861	0.7458

Здания и дороги

Кодирующая сеть	Loss	mIoU
VGG19	0.28788	0.7714
ResNet101v2	0.27074	0.7758
ResNet152v2	0.25539	0.8012
DenseNet201	0.22694	0.8582
EfficientNetb3	0.29736	0.7784
EfficientNetb5	0.26763	0.8084
EfficientNetb7	0.24701	0.8224

Пост-обработка

Алгоритмы пост-обработки:

- Полносвязные условные случайные поля (CF-CRF);
- Рекуррентные условные случайные поля (RNN-CRF);
- Сегментационная модель рафинирования CascadePSP.

Ручная обработка:

1. Пороговая обработка маски;
2. Медианная фильтрация;
3. Поиск контуров и очистка от лишних объектов;
4. Морфологические операции;
5. Модифицированный алгоритм упрощения геометрии Дугласа-Пекера, сохраняющий топологию объекта;
6. Медианная фильтрация в цикле.

Пост-обработка

Результаты пост-обработки			
Метод	Loss	mIoU	Время работы
DeepLabV3+	0.21669	0.8929	17 сек.
OCRNet	0.21397	0.9102	20 сек.
DeepLabV3+ & RNN-CRF	0.21689	0.8992	312 сек
Ensemble{DeepLabV3+ & OCRNet}	0.21255	0.9163	37 сек.
Ensemble + FC-CRF	0.22154	0.8791	125 сек.
Ensemble + FC-CRF + Ручная обработка	0.21746	0.8956	140 сек.
Ensemble + CascadePSP	0.21059	0.9248	221 сек.
Ensemble + CascadePSP + Ручная обработка	0.20939	0.9301	258 сек.

Пост-обработка



Ensemble

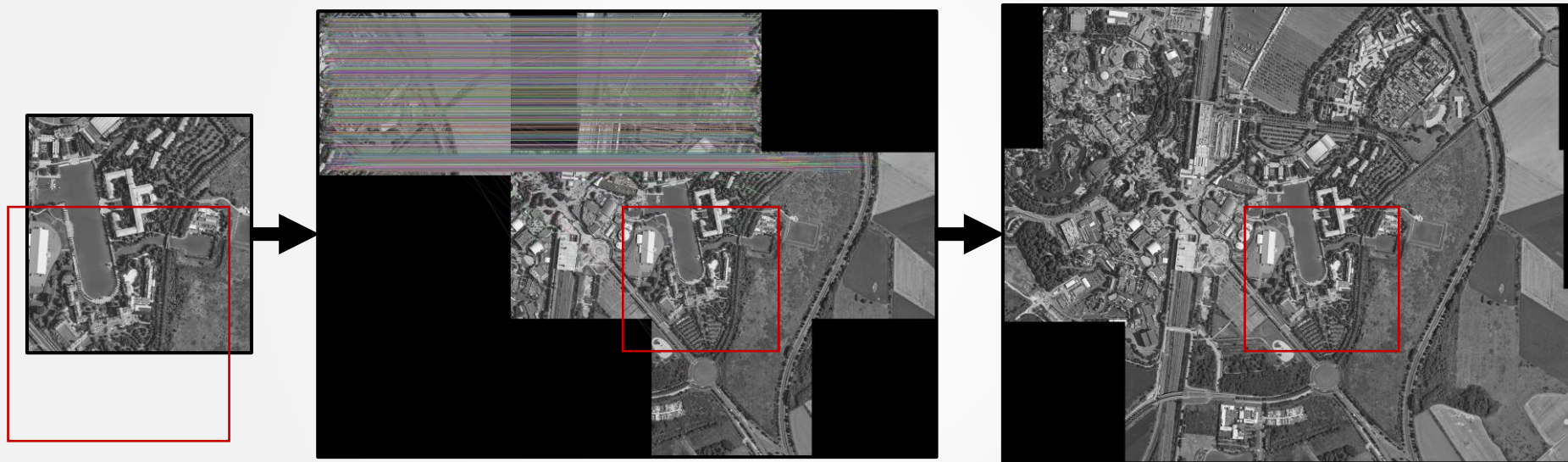
CF-CRF

CascadePSP

Получение и обработка ранее отснятого снимка целевой области

- Для получения ранее отснятого снимка целевой области используются GPS метаданные снимка, с помощью которых находится наиболее кратчайшее расстояние до координаты целевой области;
- Расчёт производится с помощью вычисления геодезического расстояния, в которой используется эллипсоидальная модель Земли WGS-84.

Получение и обработка ранее отснятого снимка целевой области



Процесс сшивки снимков

Коррекция перспективы

Для коррекции перспективы производится оценка матрицы гомографии. Для этого используются:

- **Алгоритмы выявления локальных признаков изображения:**
 - *SIFT, SURF, KAZE, AKAZE, ORB.*
- **Методы сопоставления признаков:**
 - *Brute-Force Matcher* или *FLANN Based Matcher*
- **Методы вычисления расстояния между дескрипторами:**
 - *Расстояние Хэмминга* или *Евклидова расстояние.*
- **Методы оценки матрицы гомографии:**
 - *RANSAC, LMEDS, RHO*

Коррекция перспективы

Результаты тестирования на 5 парах изображений

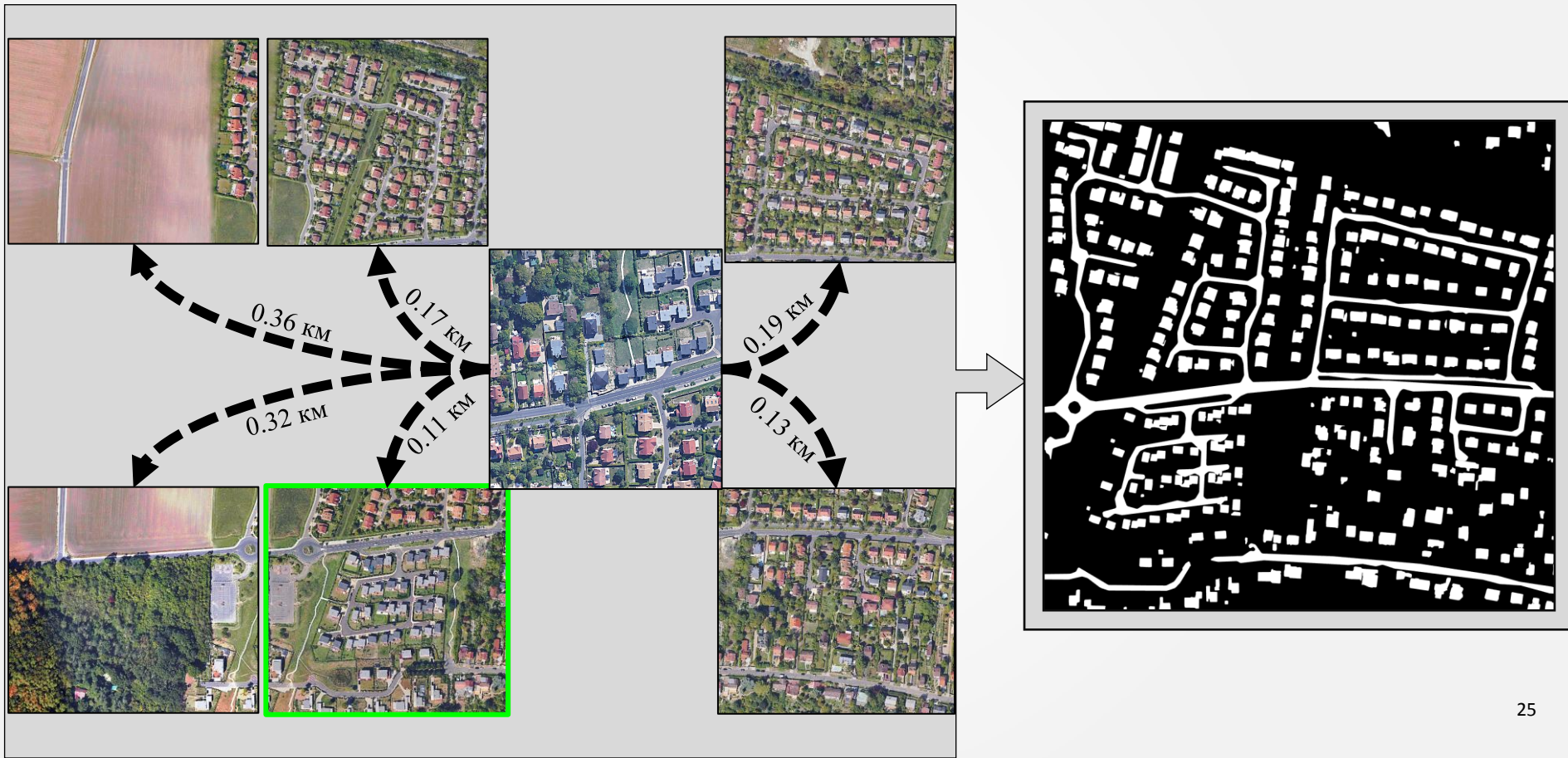
Метод	Метод оценки гомографии	Среднее время работы	Количество удачных совпадений при разных факторах					
			Исходный	Шум и Яркость	Масштаб (x1.4)	Угол (150°)	Масштаб и угол	Все факторы
AKAZE	LMEDS	49.7 с.	•51/71	•18/29	•9/13	•37/45	•10/15	•7/8
			•47/48	•11/14	•8/11	•25/26	•6/11	•0/0
			•5/9	•0/0	•0/0	•0/0	•0/0	•0/0
			•46/58	•11/15	•26/30	•42/45	•20/24	•5/8
			•42/52	•16/23	•7/9	•26/34	•9/13	•9/11
SIFT	RHO	55.1 с.	•163/196	•101/111	•85/106	•169/185	•86/101	•43/49
			•117/131	•82/89	•34/54	•102/120	•47/55	•44/54
			•24/29	•12/13	•9/11	•21/33	•10/11	•7/8
			•190/204	•20/22	•83/89	•152/156	•91/100	•26/32
			•82/117	•60/73	•50/62	•77/89	•36/45	•26/36
ORB	RANSAC	3.3 с.	<5	<5	<5	<5	<5	<5

Коррекция перспективы

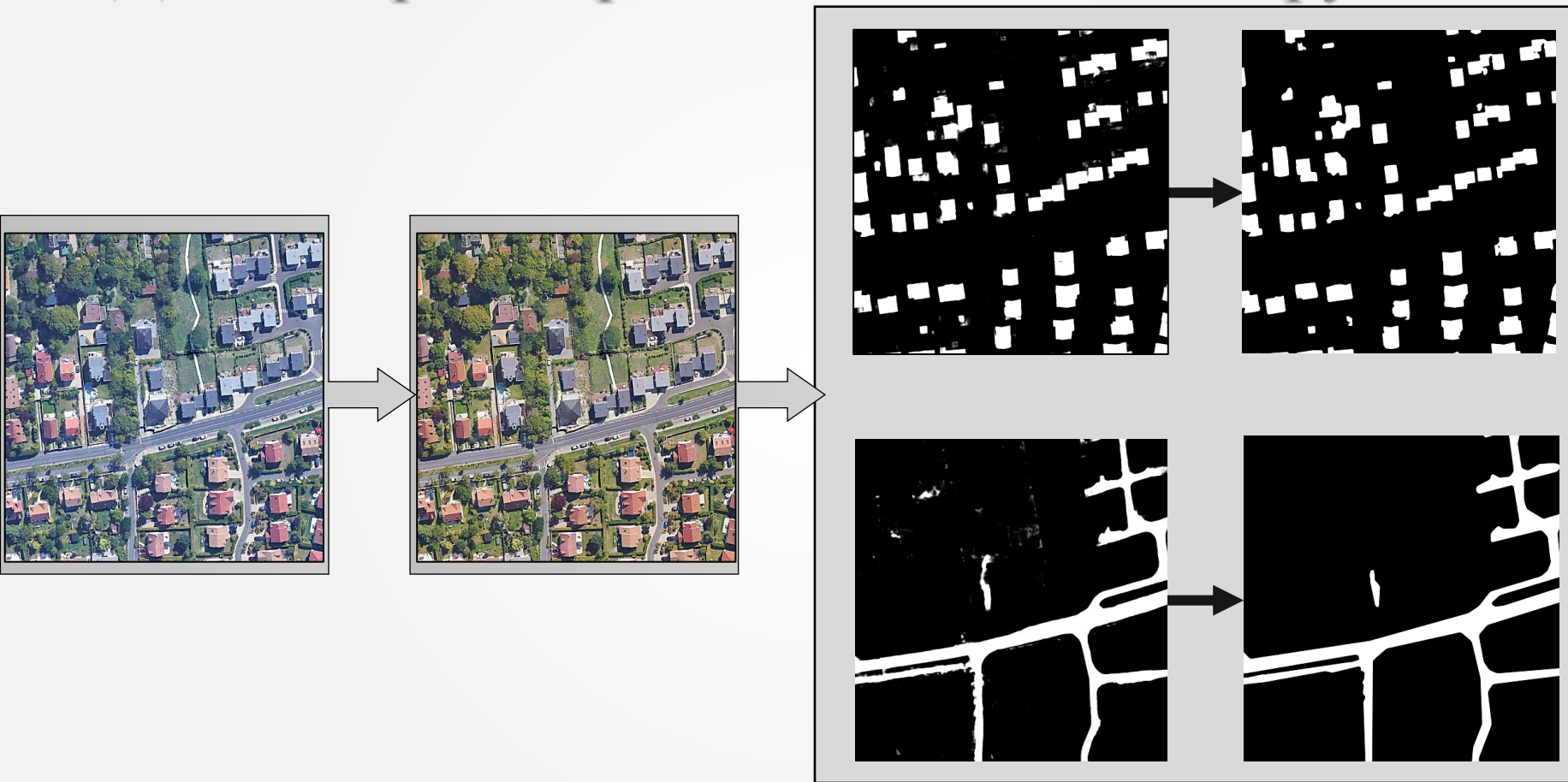


Процесс коррекции перспективы

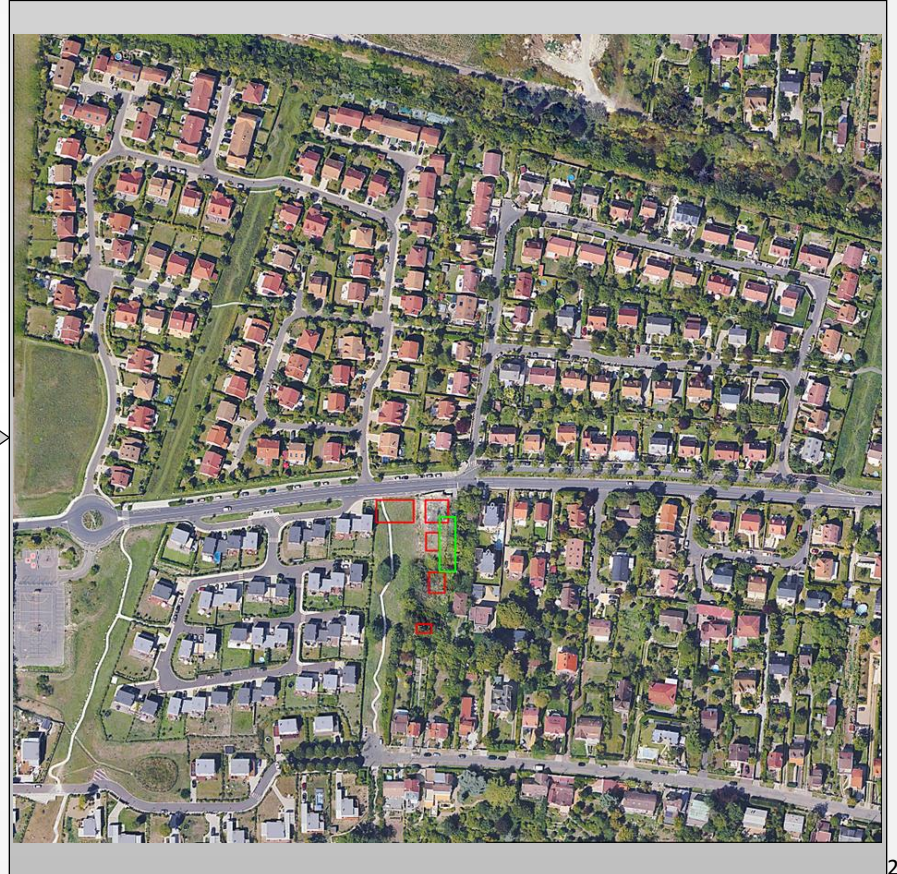
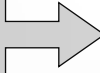
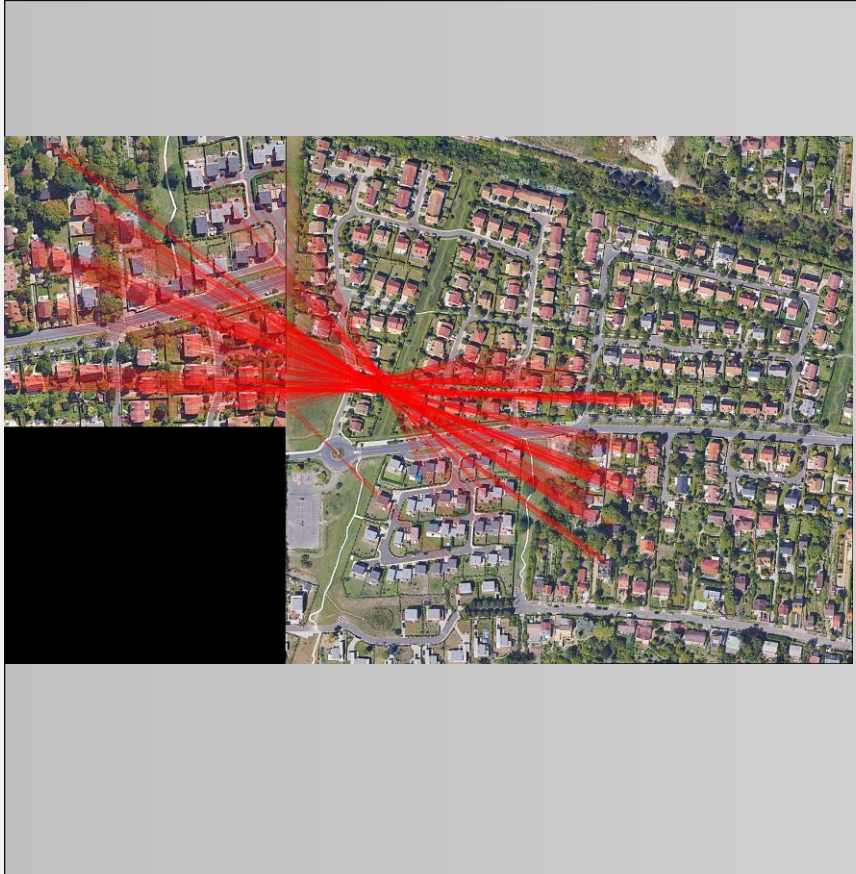
Демонстрация работы системы обнаружения



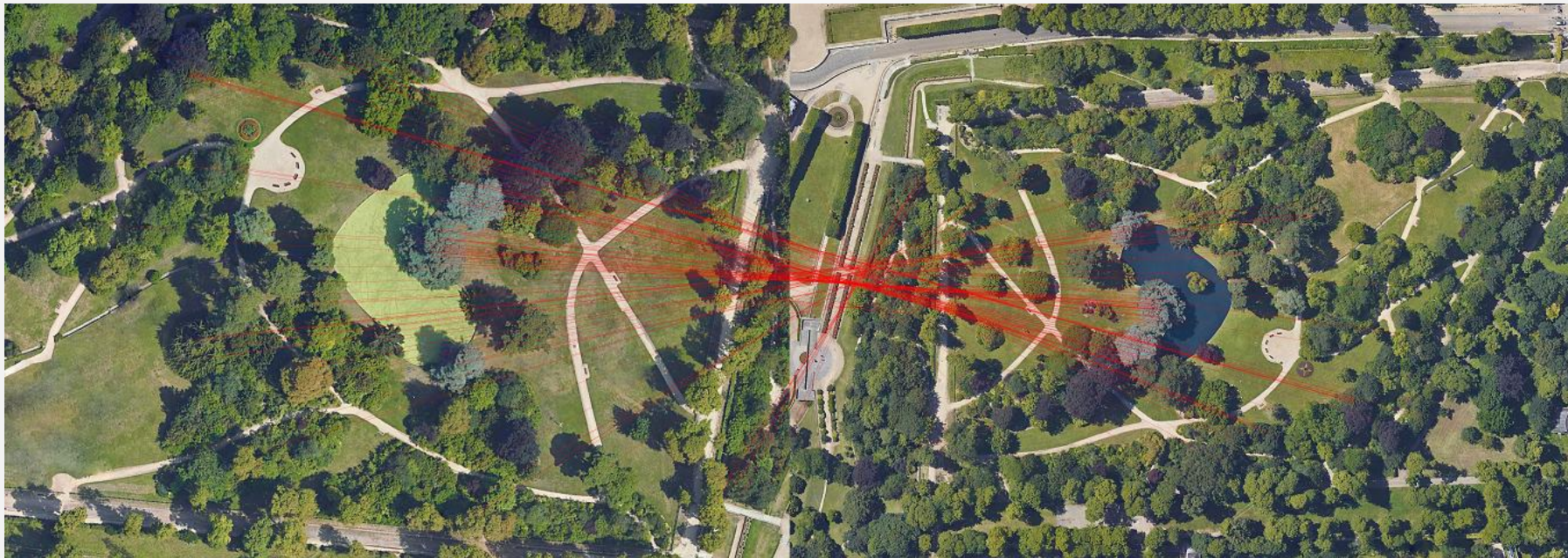
Демонстрация работы системы обнаружения



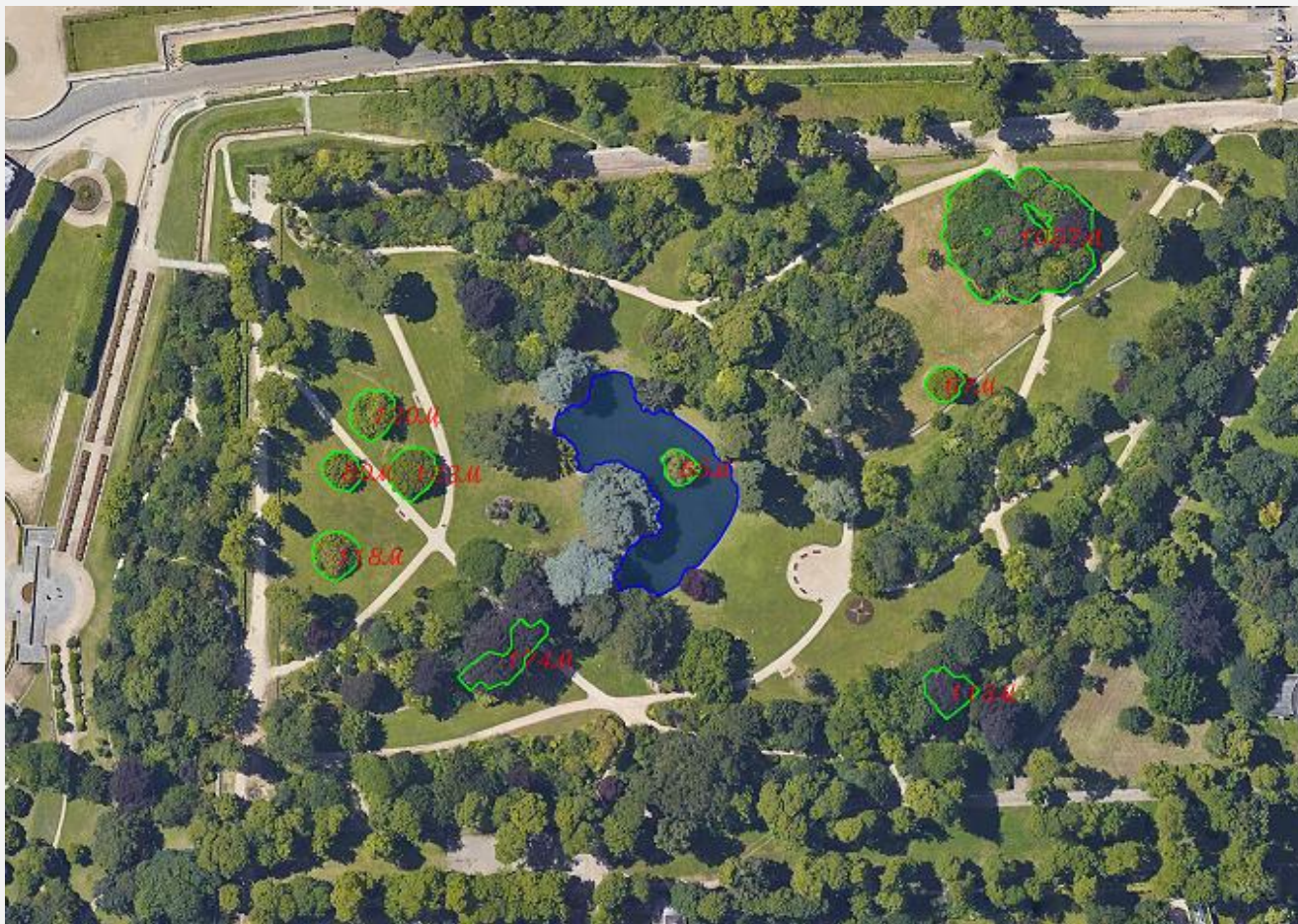
Демонстрация работы системы обнаружения



Демонстрация работы системы обнаружения



Демонстрация работы системы обнаружения



Заключение

Данная работа посвящена разработке математического и программного обеспечения системы для мониторинга элементов новизны на местности.

В результате было выполнено:

- Подготовлен собственный набор данных;
- Подготовлены модели глубоких нейронных сетей;
- Проведена пост-обработка;
- Организовано получение и обработка ранее отснятого снимка и процесс коррекции перспективы с итоговым обнаружением изменений.

Спасибо за внимание!