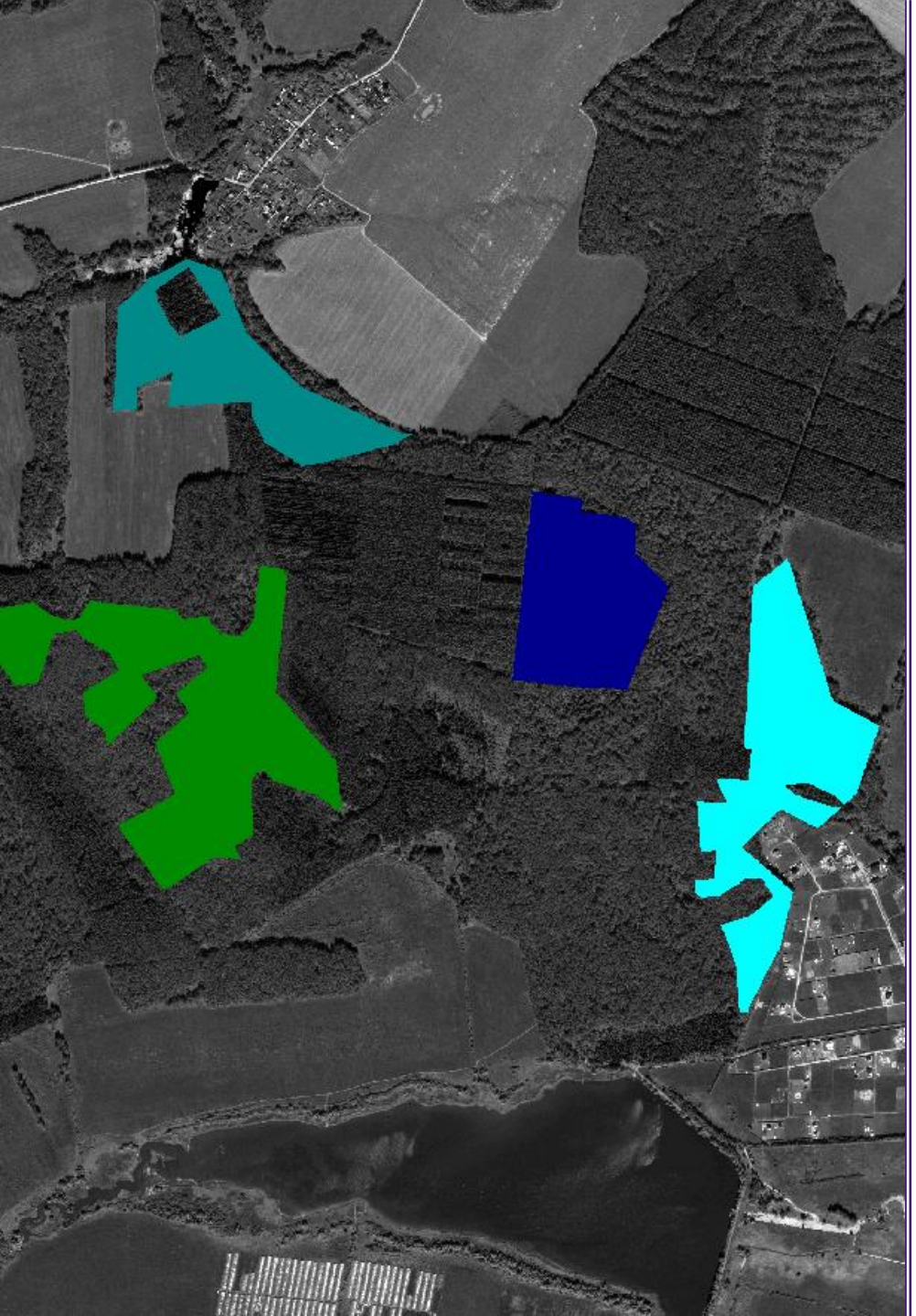


Междоменная адаптация моделей глубокого обучения при сегментации спутниковых данных с ограниченной разметкой

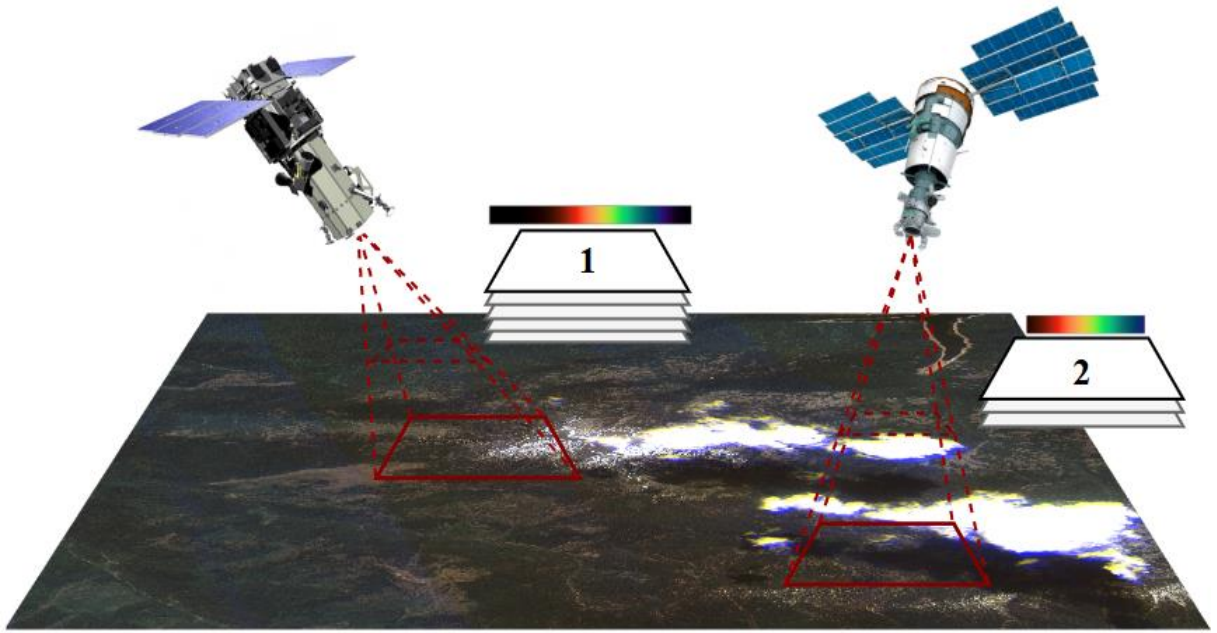
Данил Андреевич Шатилов[*]

Научный руководитель – к.т.н. С.А. Рылов[*]

[*]Федеральный исследовательский центр информационных и вычислительных технологий, г. Новосибирск

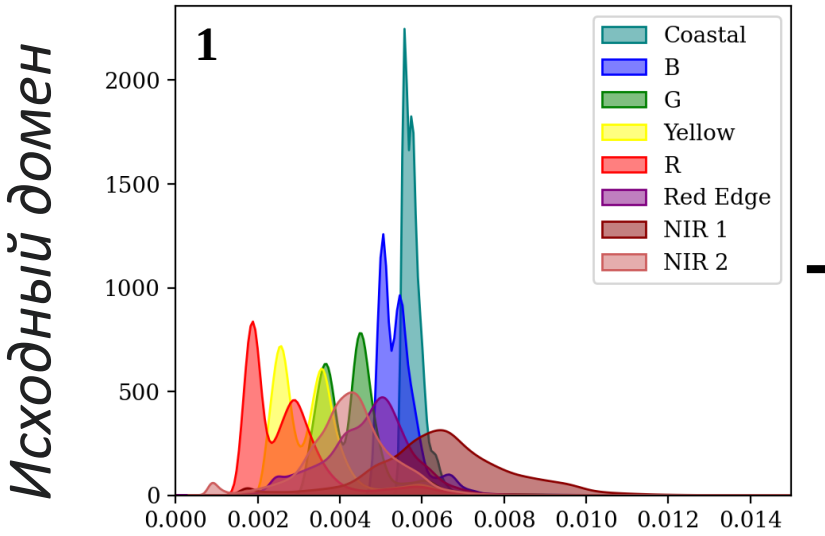
ноябрь 2026 г.

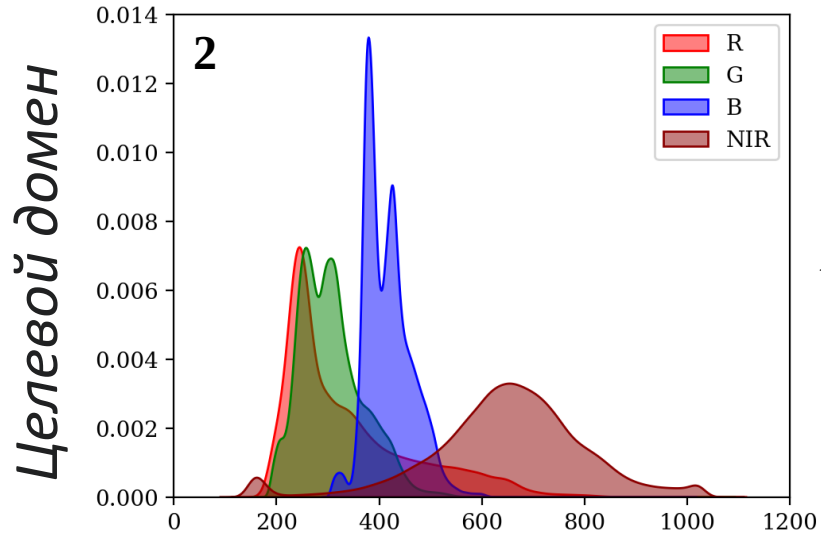
Проблема сдвига домена в задачах ДЗЗ

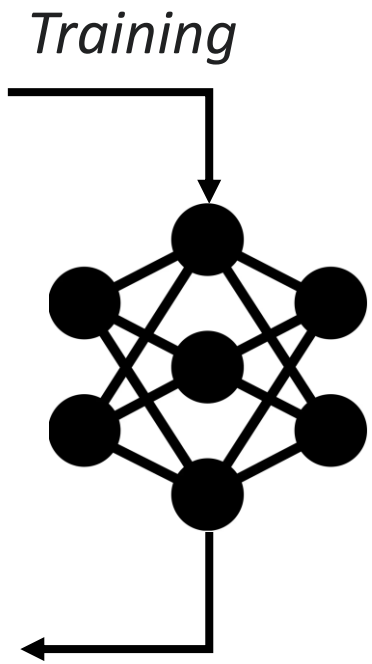


1. Географическая привязка;
2. Временной и сезонный сдвиг;
3. Атмосферные условия;
4. Гетерогенность сенсоров;
5. Различие уровней предобработки.

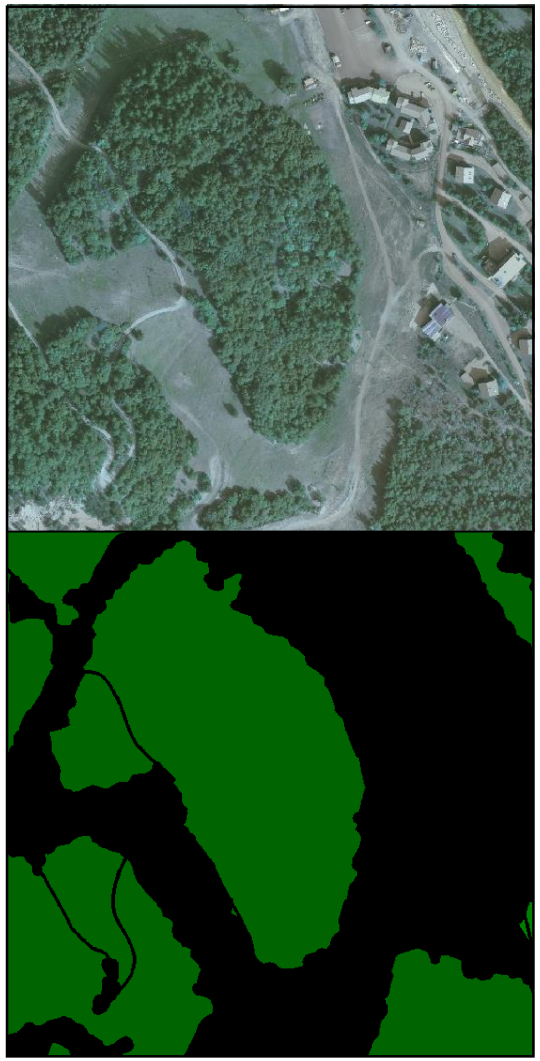
**Дополнительно*: дефицит разметки!



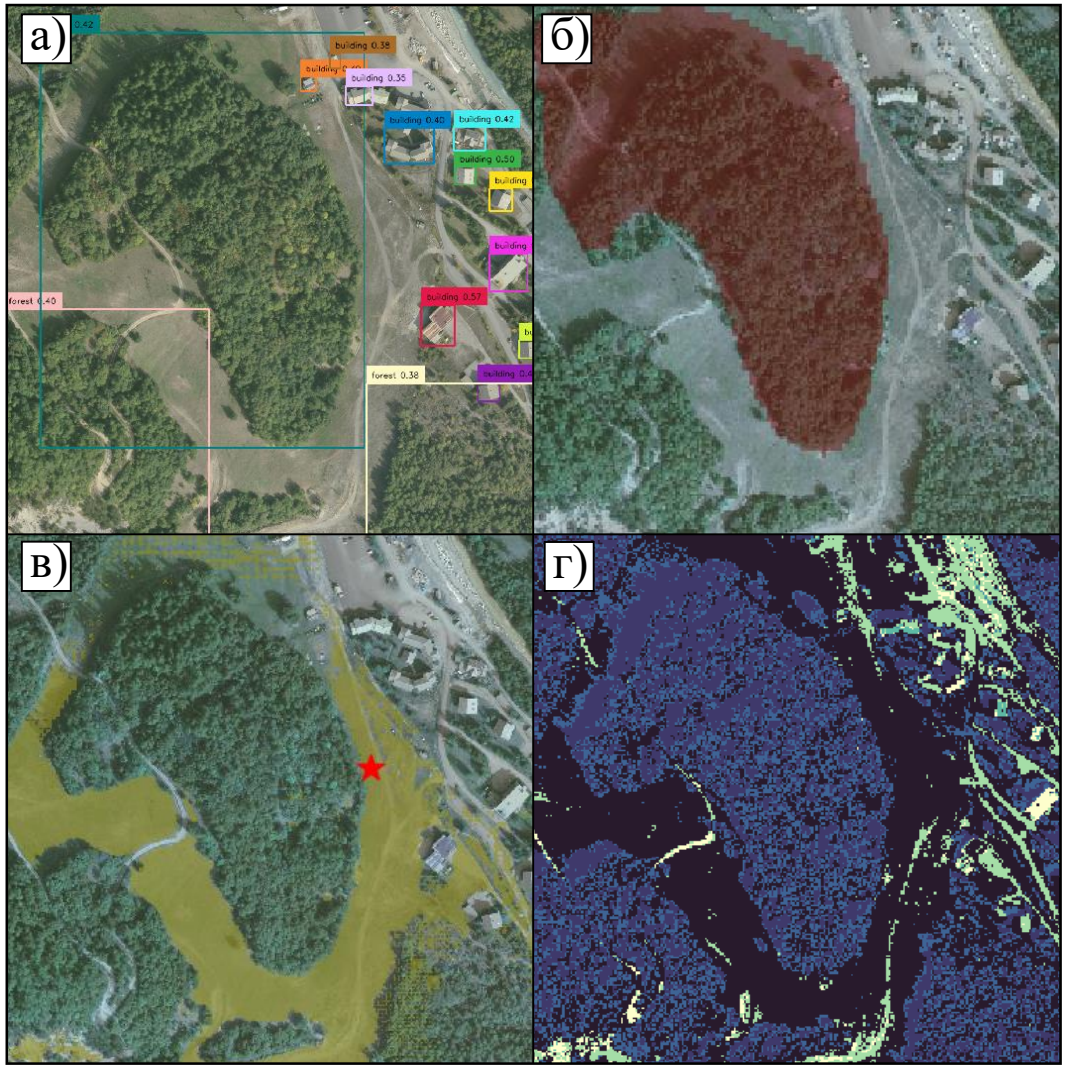




Доменный сдвиг в действии: *zero-shot* на *RGB* данных ДЗЗ



Исходный *RBG* снимок и разметка



а) - [Grounding DINO], запрос «*forest. building.*»
б) - *[PaliGemma 2]*, запрос «*segment forest*»
в) - [SAM], при выделении точки
г) - классический *[Mean-shift]*

Современные подходы к решению доменного сдвига [1]

Доменная адаптация (*ap*. целевой домен!)		Адаптация во время тестирования		Обобщение доменов	
Категория:	**Применимость к ДЗЗ:**	**Категория:**	**Применимость к ДЗЗ:**	**Категория:**	**Применимость к ДЗЗ:**
Состязательное выравнивание (*DANN*)	смешение классов	Пересчёт статистик *BatchNorm*	+	Аугментации, рандомизация стиля	+
Выравнивание признаков и предсказаний	+	Псевдометки	+, полное обновление весов	Синтетические домены	сложности с корректной генерацией
Спектральные индексы (*NDWI*) как регуляризатор	привязка к каналам	Вспомогательное самообучение на инференсе	+, полное обновление весов	Метаобучение	нужно несколько доменов
				Самообучение	+

[1] Zhou, K., Liu, Z., Qiao, Y., Xiang, Tao., & Loy, C. C. (2022). Domain generalization: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4), 4396-4415.

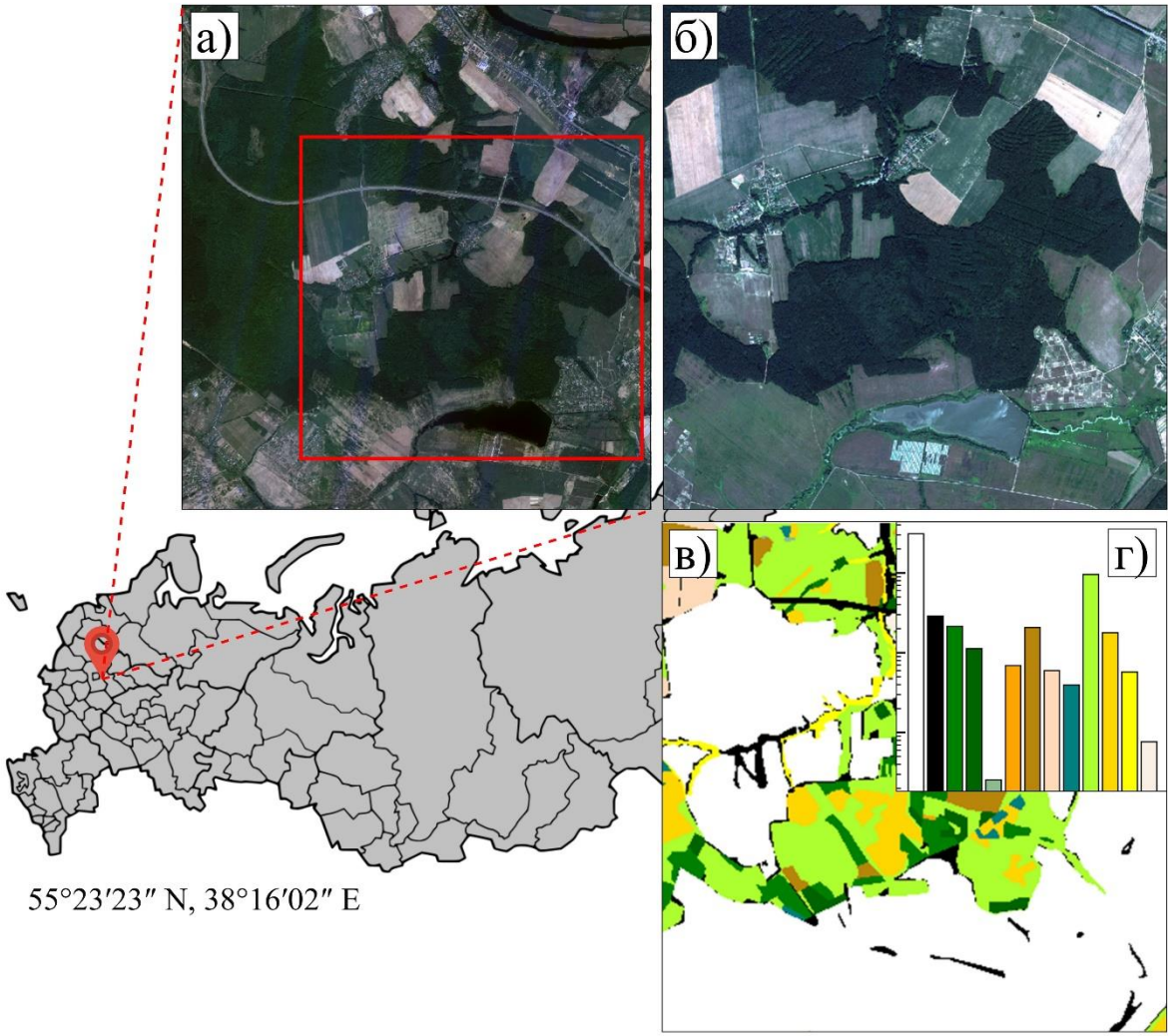
Цели и задачи работы

Цель: оценить эффективность различных стратегий аугментации данных и обучения легких *CNN* моделей как средств повышения обобщающей способности на примере семантической сегментации лесных территорий на малых выборках аэрокосмических снимков.

Задачи:

1. Сформировать датасеты на основе разметки видового состава древостоев и разновременных аэрокосмических снимков;
2. Провести сравнительный анализ аугментаций и определить их оптимальную комбинацию при обучении на одном домене;
3. Оценить вклад предварительного самообучения *CNN*;
4. Исследовать устойчивость моделей к доменному сдвигу при переносе между спутниковыми платформами («*WorldView-2*» → «*Ресурс-П*») без дообучения.

Основной тестовый полигон



- Бронницкое лесничество, Московская обл.;
- 2 мультиспектральных снимка со спутников «*WorldView-2*» (8 каналов, атмосферная коррекция) и «*Ресурс-П*» (4 канала, радиометрическая коррекция);
- Карта таксонов древесной растительности [2].

→ сформированы 2 датасета по **81** паре тайлов, размеры 256x256, разрешение ~ 2 м/п.

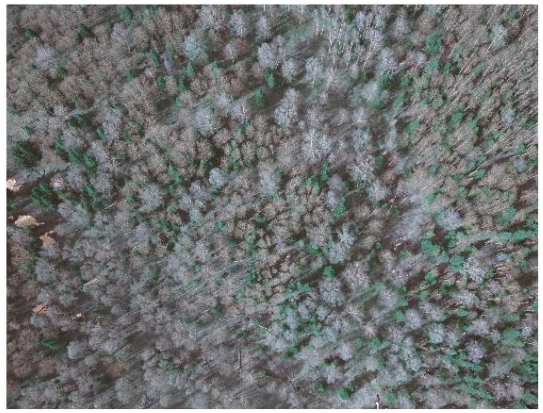
Не лес
Нет информации
Сосна обыкн.
Ель европейская
Пихта сибирск.
Листв. европ.
Листв. сибирск.
Листв. Сукачева
Дуб черешчатый
Береза
Осина
Ольха черная
Липа мелколистн.

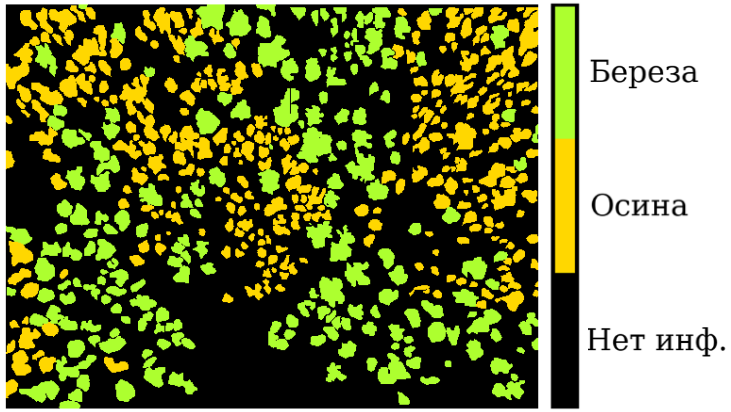
[2] Дмитриев Е. В., Кондранин Т. В. и др. Определение видового состава смешанного леса на основе совместной обработки публичных спутниковых карт и многовременных изображений Sentinel-2 // Современные проблемы дистанционного зондирования Земли из космоса. – 2023. – Т. 21. – № 1. – С. 31-50.

Дополнительные *RGB*-наборы данных









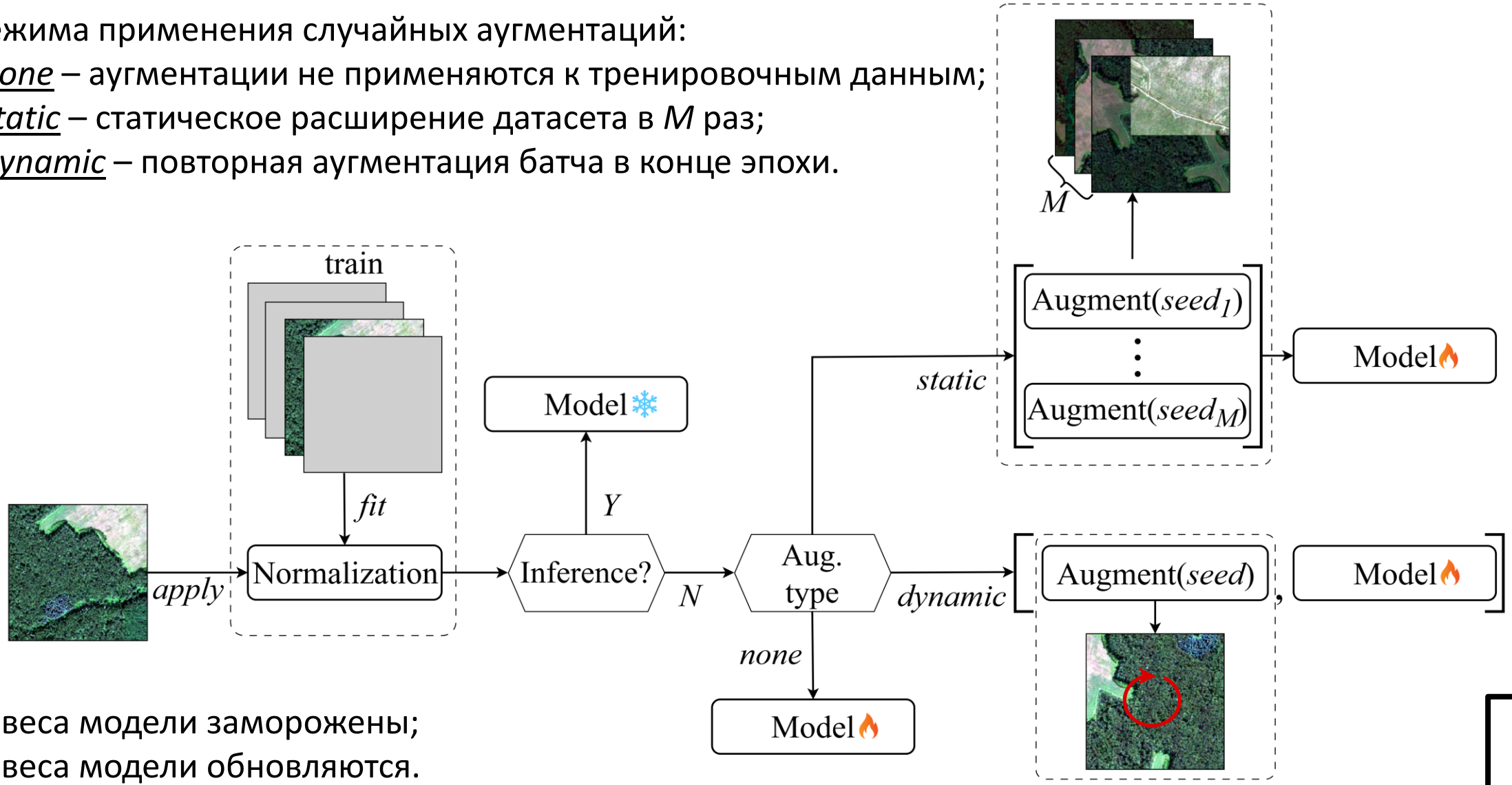
- *IGN Hautes-Alpes* [3], 800 *uint8* аэрофотоснимков территории Верхних Альп (Франция);
- Участок смешанного леса в районе п. Аламбай Алтайского края [4] – 1752 *uint8* аэрофотоснимка.

[3] Guérin E, et. al. Satellite image semantic segmentation. *arXiv*. 2021. https://doi.org/10.48550/arXiv.2110.05812
[4] Пестунов И.А., Калашников Р.А., Рылов С.А. Семантическая сегментация осиновых и березовых древостоев на RGB - изображениях с БПЛА при помощи сверточных нейронных сетей // Вычислительные технологии. 2024. Т. 29. № 5. С. 100-112

Схема эксперимента

3 режима применения случайных аугментаций:

- *none* – аугментации не применяются к тренировочным данным;
- *static* – статическое расширение датасета в *M* раз;
- *dynamic* – повторная аугментация батча в конце эпохи.



❄ – веса модели заморожены;

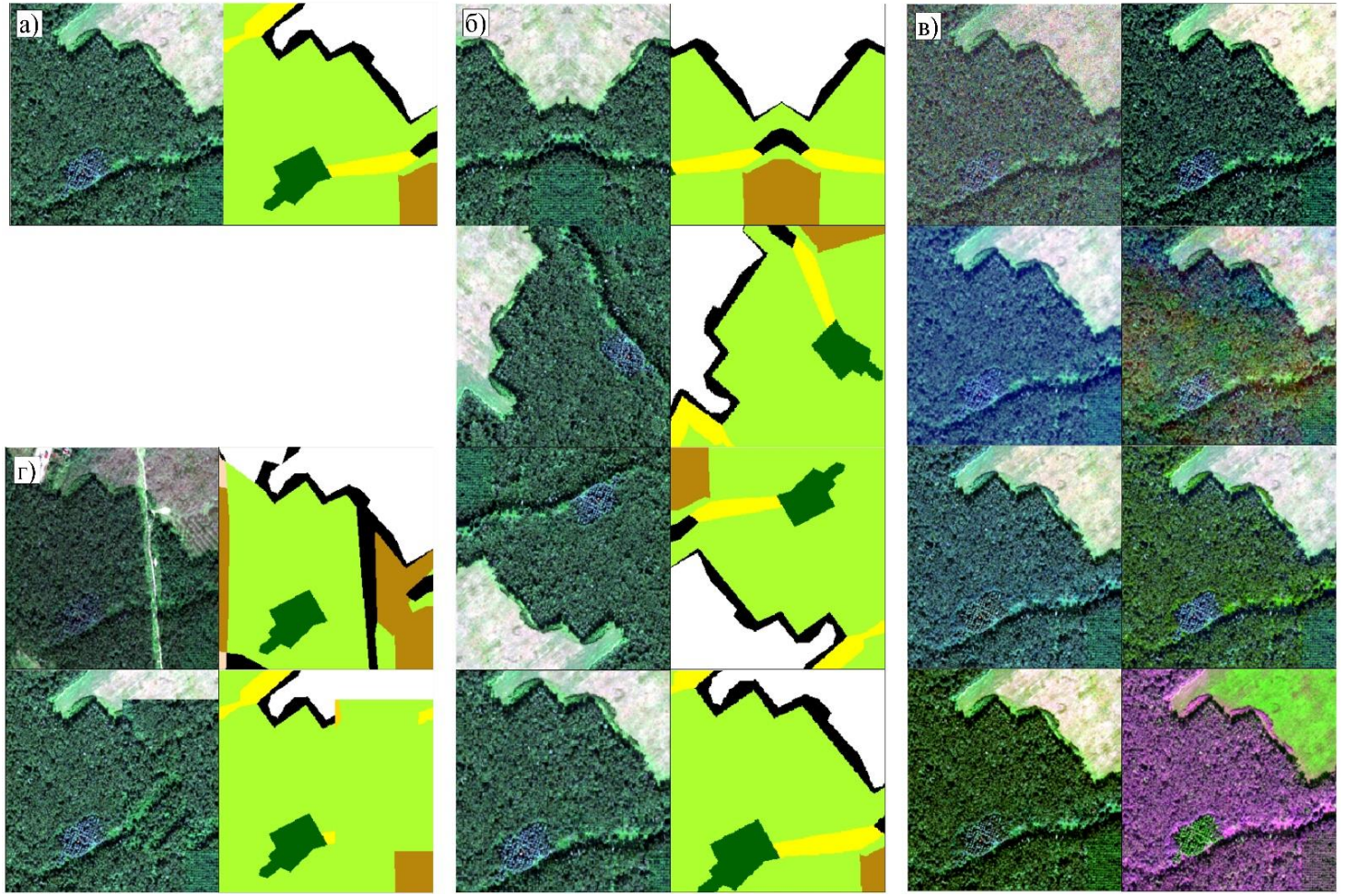
🔥 – веса модели обновляются.

Стратегии нормализации и аугментации

Рассмотрены 3 типа нормализации:
линейная, *стандартизация*, *процентное растяжение*.

Общий вид используемой поканальной независимой нормализации:

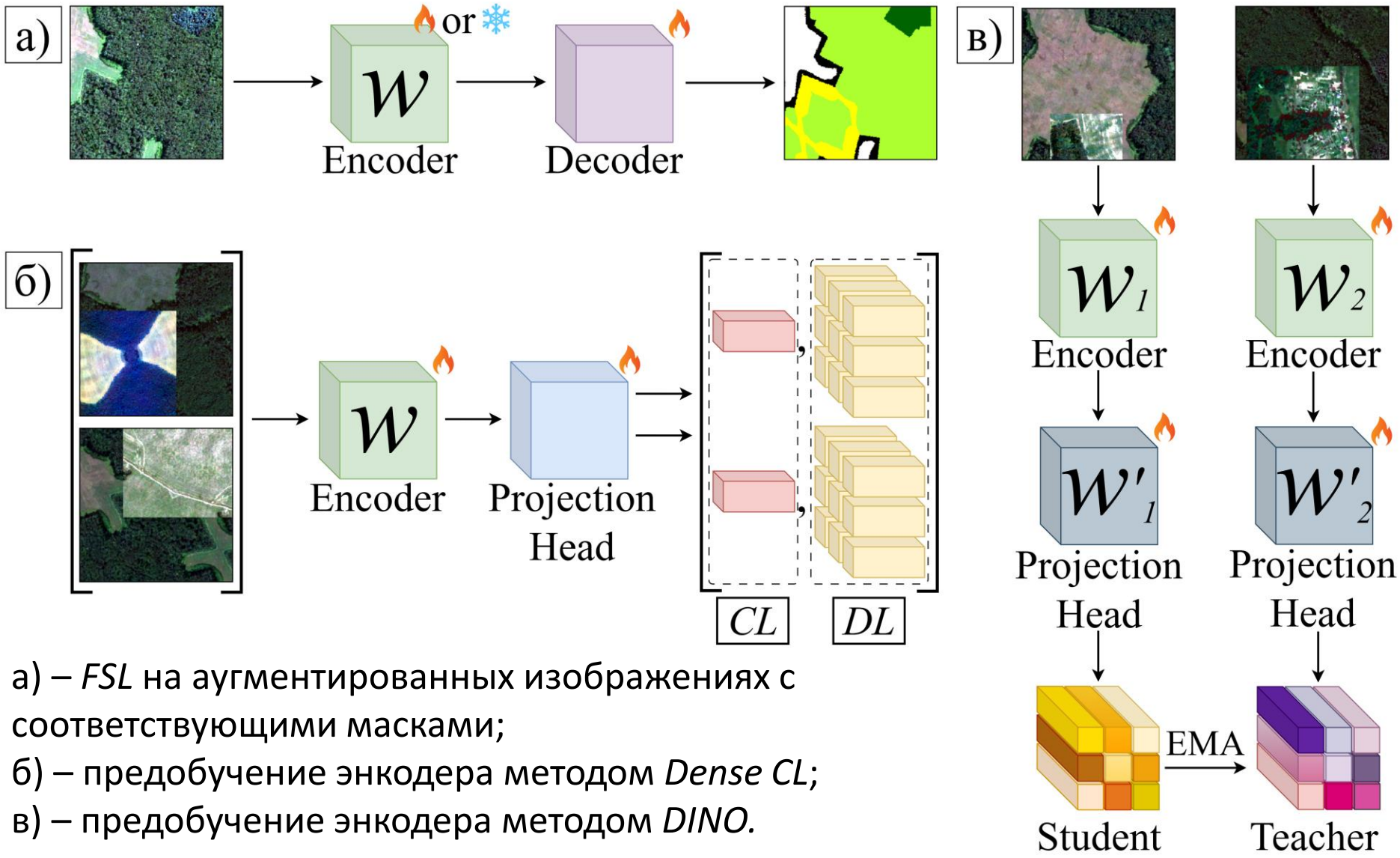
$$f(I_{hwc}, a_c, b_c, \alpha, \beta) = max\left(\alpha, min\left(\frac{I_{hwc} - a_c}{b_c}, \beta\right)\right)$$



а) – исходные пара изображение-маска;
б-г) результаты применения аугментаций:
геометрических, *спектральных* и *батч-смешивающих*

геометр. (сверху вниз): трансляция, поворот, зеркальное отражение, масштаб.; спектр. (слева направо): гауссов шум, изменение контраста, резкости, маск. Фурье-спектра, HSV-подобное преобр., перестановка каналов; батч-смеш. (сверху вниз): MixUp, CutMix

Стратегии обучения моделей

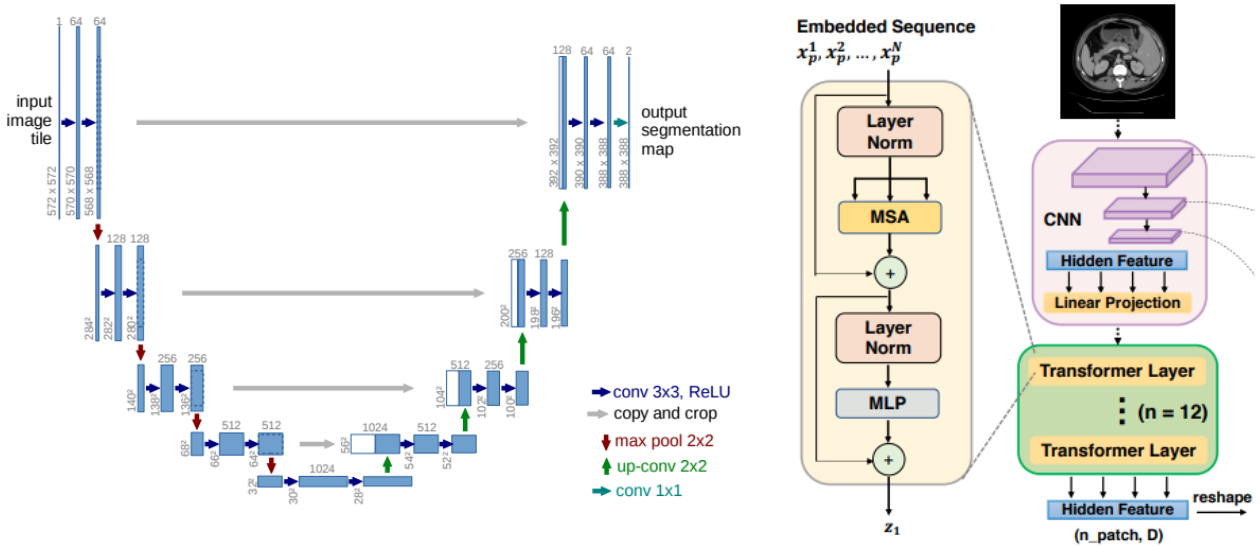


а) – *FSL* на аугментированных изображениях с соответствующими масками;

б) – предобучение энкодера методом *Dense CL*;

в) – предобучение энкодера методом *DINO.*

Используемые модели и метрики



$$micro\ \text{acc} = \frac{\sum_i TP_i}{\sum_i TP_i + \sum_i FP_i}$$

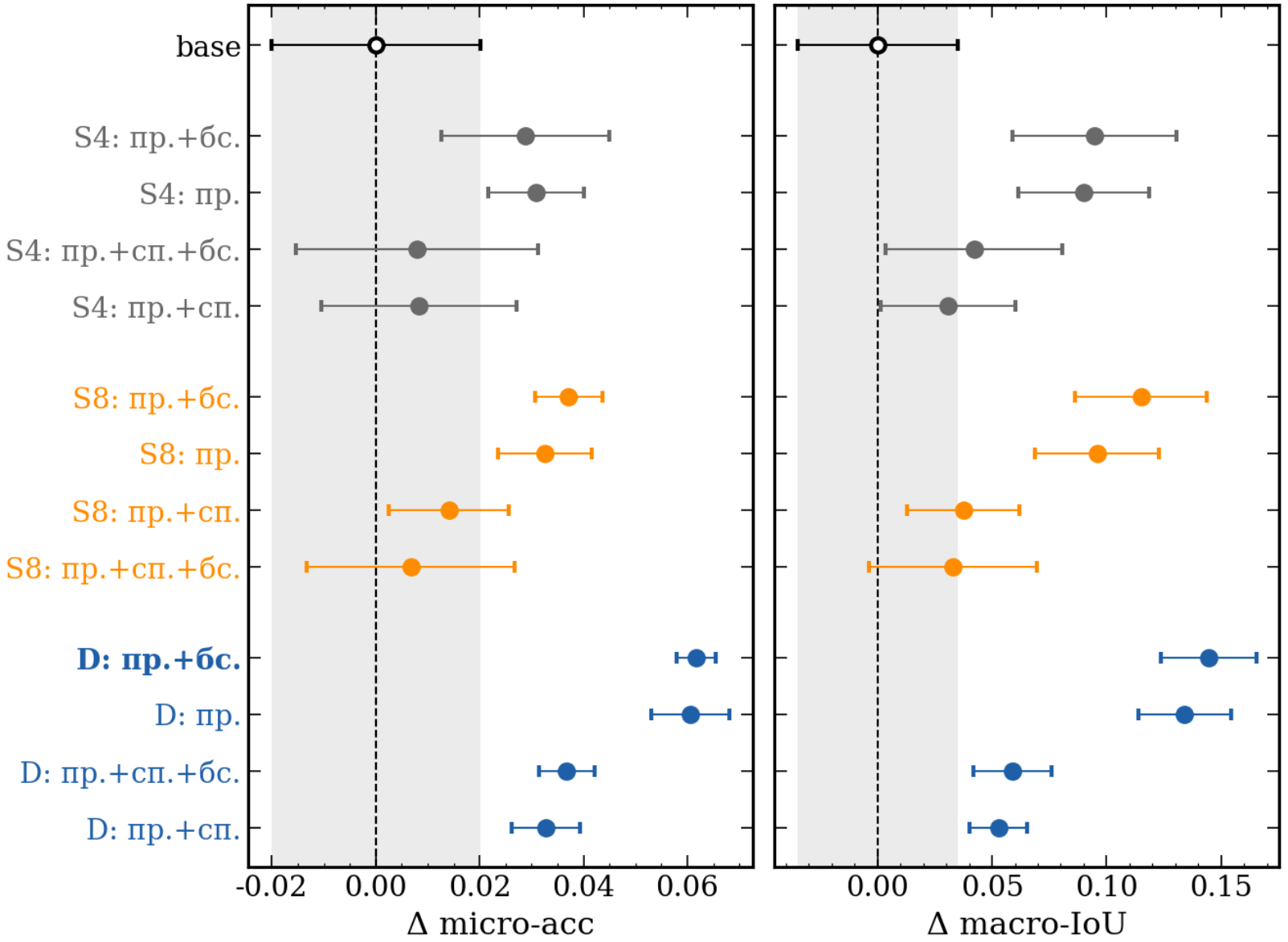
$$macro\ IoU = \frac{1}{N}\sum_{i=0}^{N-1}\frac{TP_i}{TP_i + FP_i + FN_i}$$

Стандартный вид архитектуры *U-Net* и энкодер на базе трансформера

- Фреймфорки: *Keras, Keras U-net Collection*, *Segmentation Models*;
- Модели: *FPN*, *Link-Net*, *U-Net*, *U-Net++*, *Att. U-Net*, *Res U-Net*, *U2-Net*, *Trans U-Net* энкодер – *ResNet-50*;
- Оптимайзер: *AdamW*;
- *FSL* функция потерь*: Dice + Focal* [5].

[5] Шатилов Д., Рылов. С. Семантическая сегментация лесных массивов на спутниковых снимках высокого разрешения с помощью U-Net-подобных моделей // Вычислительные технологии. – 2025. – Т. 30, № 3. – С. 111-126.

Роль аугментаций при сегментации в рамках одного домена



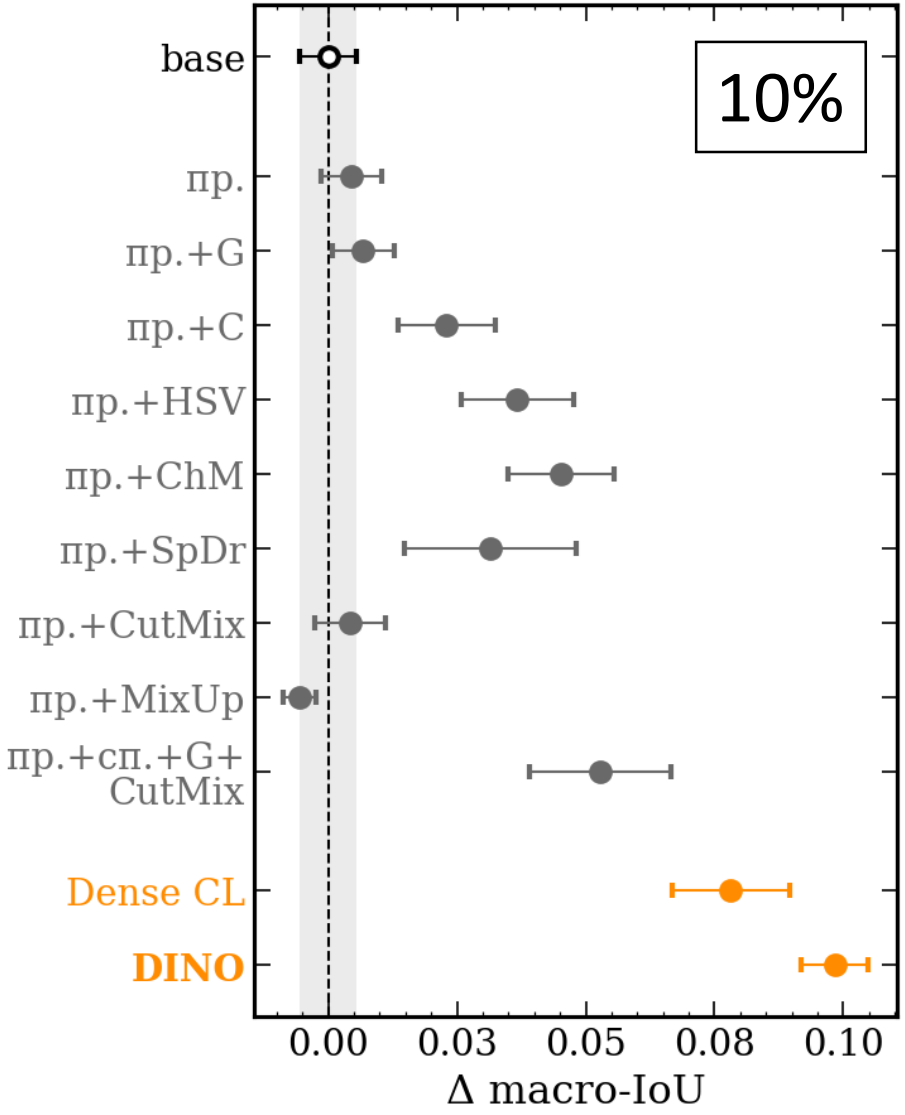
Прирост метрик для «*WV-2*», усредненный по моделям, *batch-size=[8, 16]* и кросс-валидации, относительно базового варианта.

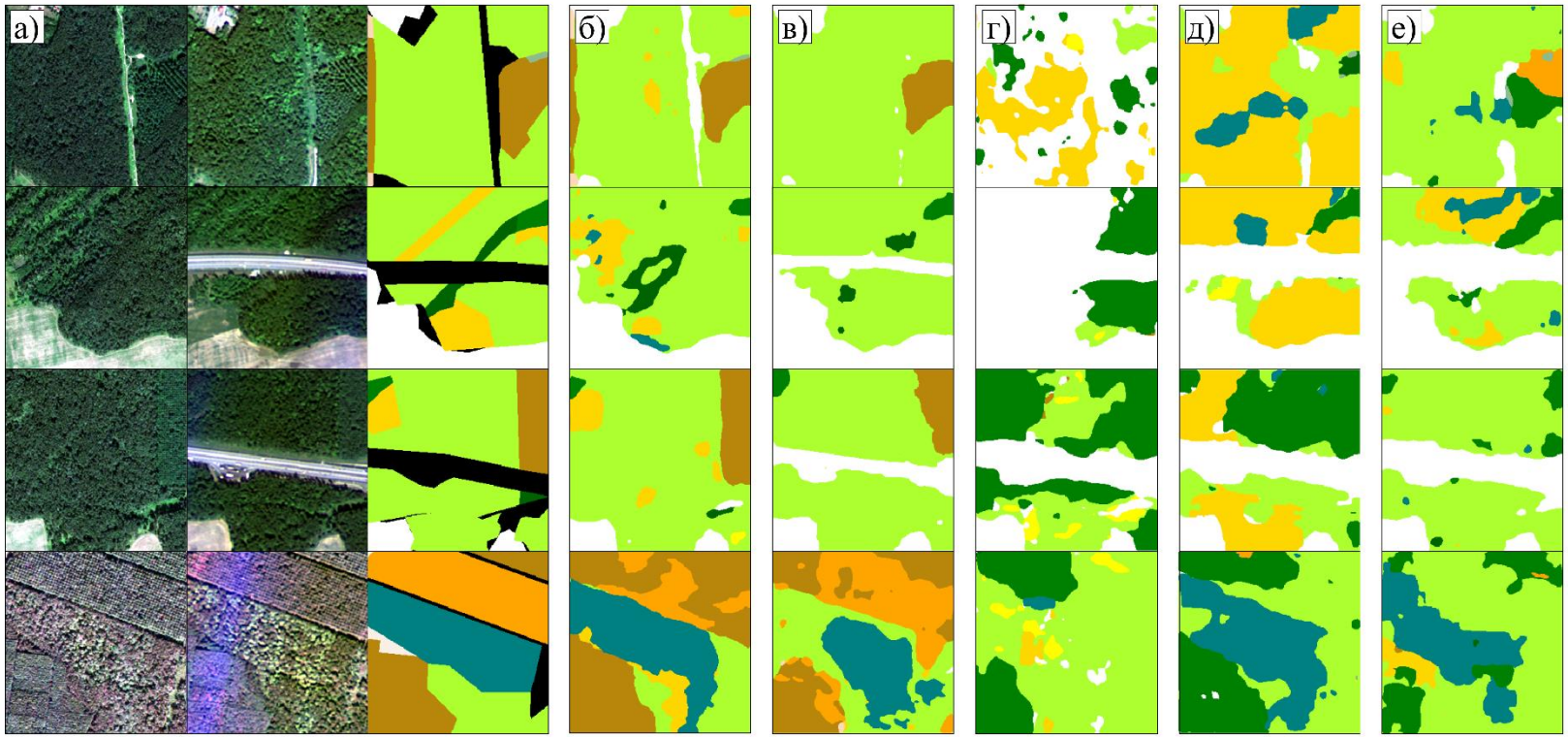
base – нет аугментаций;
S4/8 – статическое расширение в 4/8 раз;
D – динамические аугментации.

Увеличение *macro-IoU*: для мультиспектральных датасетов – **15%** (*WV-2)* и **10%** (*Ресурс-П*);

для *RGB* датасетов – **4%** (*IGN*) и **3%** (*Аламбай*).

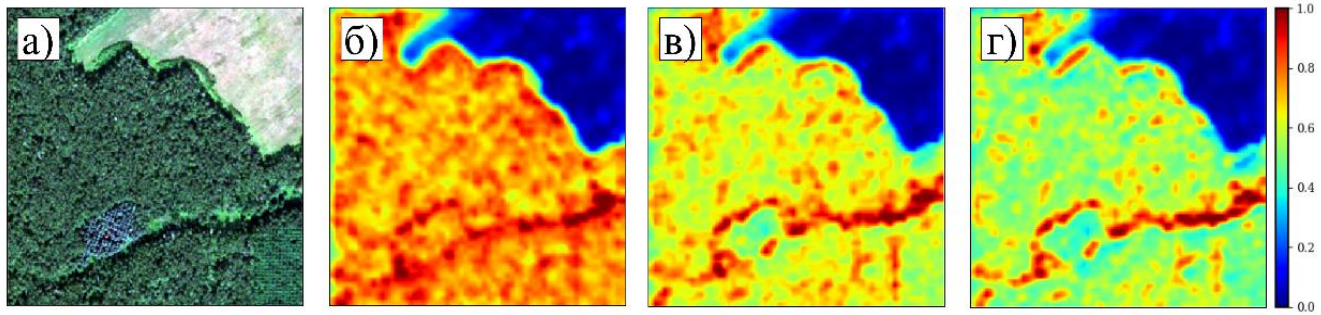
Перенос моделей между доменами: «*WorldView-2*» → «*Ресурс-П*»

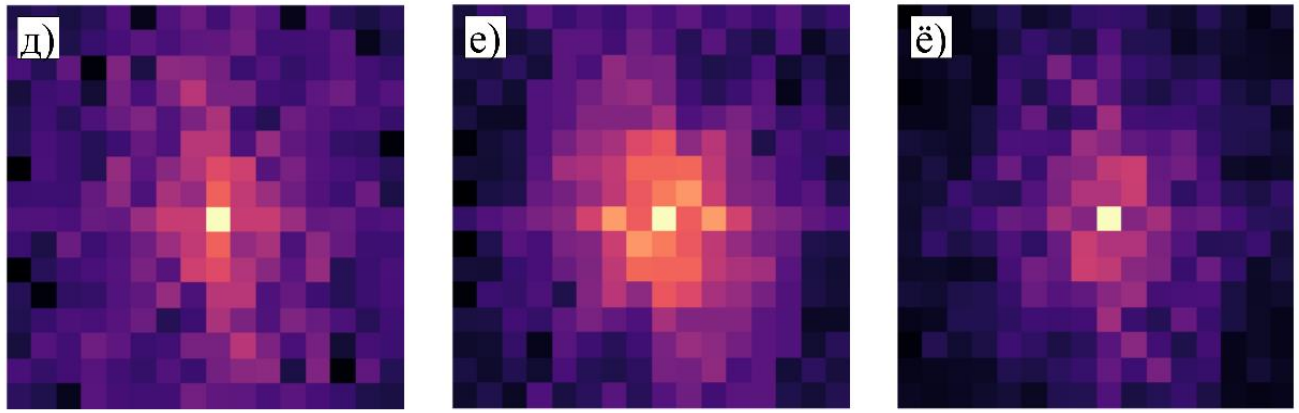


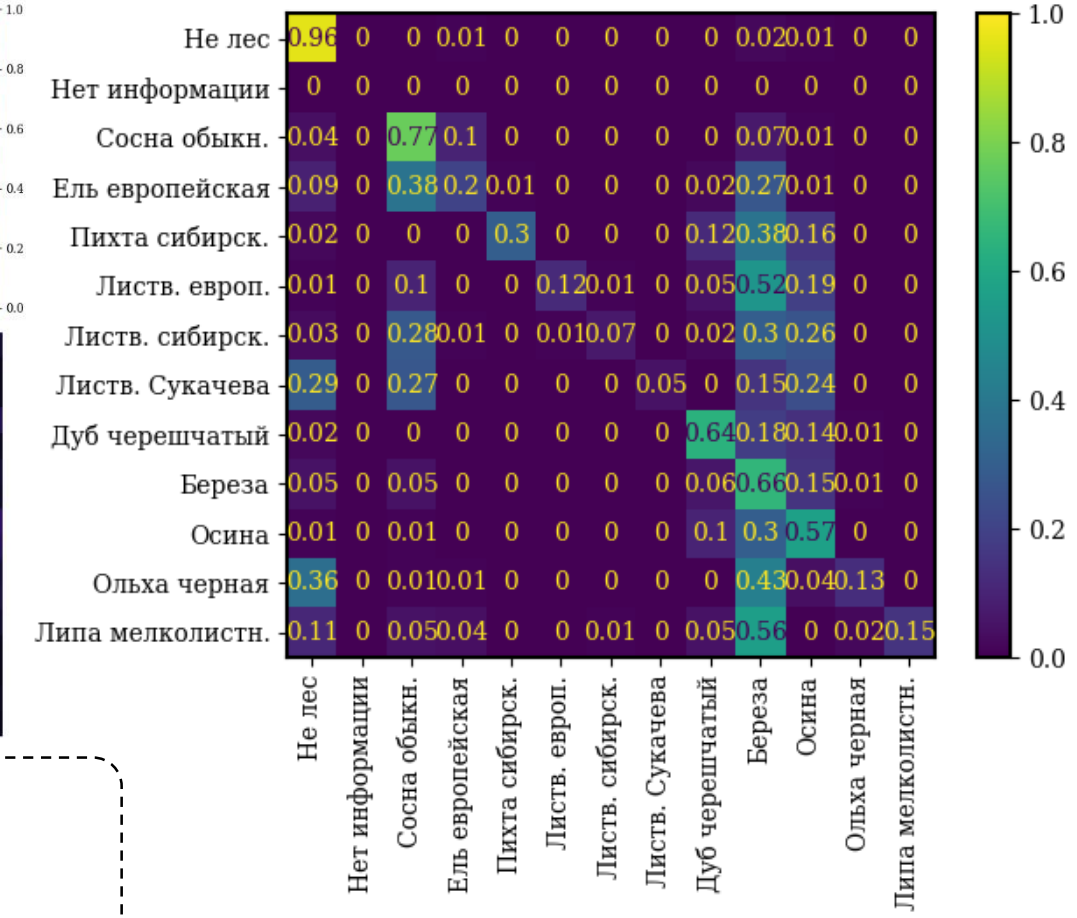


а) исходные данные («*WV-2*», «*Ресурс-П*», разметка);
б) сегментация при обучении на «*WV-2*»;
в) сегментация при обучении на «*Ресурс-П*»;
г) сегментация при переносе без применения аугментаций;
д) сегм. при переносе с агрессивными спектральными аугм.;
е) сегментация при переносе весов с *SSL* предобучением.

Анализ моделей при доменном сдвиге







Матрица ошибок при переносе *Dense CL*-предобученной модели на данные «*Ресурс-П*»

а) – исходное изображение;
б)-г) – визуализация карт активаций через первую главную компоненту (*PCA*): б) – *FSL* без аугментаций,
в) – *FSL* с динамическими аугментациями,
г) – предобучение методом *DINO*;
д)-ё) – спектры мощности *FFT* для соответ. конфигураций.

Заключение

Данные: 2 крупномасштабных *RGB* датасета и 2 экстремально малых мультиспектральных наборв («*WorldView-2*»/ «*Ресурс-П*») по 81 тайл 256×256.

Аугментация: +15/10% macro-IoU относительно обучения без аугментаций для мультиспектра (+4/3% *RGB*). Основной вклад дают пространственные преобразования; спектральные внутри одного домена могут снижать точность.

Перенос между сенсорами без дообучения: поканальная стандартизация вместе с агрессивными аугментациями (*HSV*-подобные, *CutMix*) сокращает деградацию на 4-6% macro-IoU.

Самообучение (*Dense CL, DINO*): дополнительно +5% *macro-IoU*. Энкодер ориентируется на текстурные признаки и глобальный контекст (анализ *PCA* и *FFT*).

Итог: предлагаемая методика – это комбинация известных техник поверх стандартного пайплайна → её легко внедрить и настроить под новые данные и сенсоры.

Спасибо за внимание!

Данил Андреевич Шатилов

d.shatilov.off@gmail.com

Авторы выражают признательность к.ф.-м.н. Егору Владимировичу Дмитриеву за предоставленные данные разметки