

*VII Международная научно-практическая конференция «Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем» («РАДИОИНФОКОМ – 2023»)*

*РТУ МИРЭА, Москва, ноябрь 2023 г.*



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

«МИРЭА – Российский технологический университет»

РТУ МИРЭА

---

Институт радиоэлектроники и информатики

Кафедра Геоинформационных систем

---

# ГИС-ИНСТРУМЕНТЫ С ОТКРЫТЫМ КОДОМ ДЛЯ РАСПОЗНАВАНИЯ ДОРОВ

*Студент: Кокуркин Александр Дмитриевич*

*Руководитель: Подольская Екатерина Сергеевна, к. т. н.  
доцент кафедры геоинформационных систем ИРИ РТУ МИРЭА  
ведущий научный сотрудник ЦЭПЛ РАН*

# Актуальность

- Распознавание дорог по снимках разного типа продолжает оставаться актуальной темой в широком спектре тематических проектов.
- Для лесного хозяйства и регионального лесного транспортного моделирования наземного доступа к лесным пожарам и ресурсам леса особое значение имеет инфраструктура дорог, распознавание которых по космическим снимкам может стать основой для актуализации дорожного графа.
- Инструменты с открытым кодом и космические снимки открытого доступа являются важной основой проектов по транспортному моделированию.

# Цель и задачи

**Цель работы** состоит в изучении возможностей ГИС-инструментов с открытым кодом для распознавания дорог разных классов на примере тестовых территорий в Красноярском крае, одном из центров современного лесного хозяйства и лесной промышленности и обладающего значительной неравномерностью развития инфраструктуры дорог.

## Задачи:

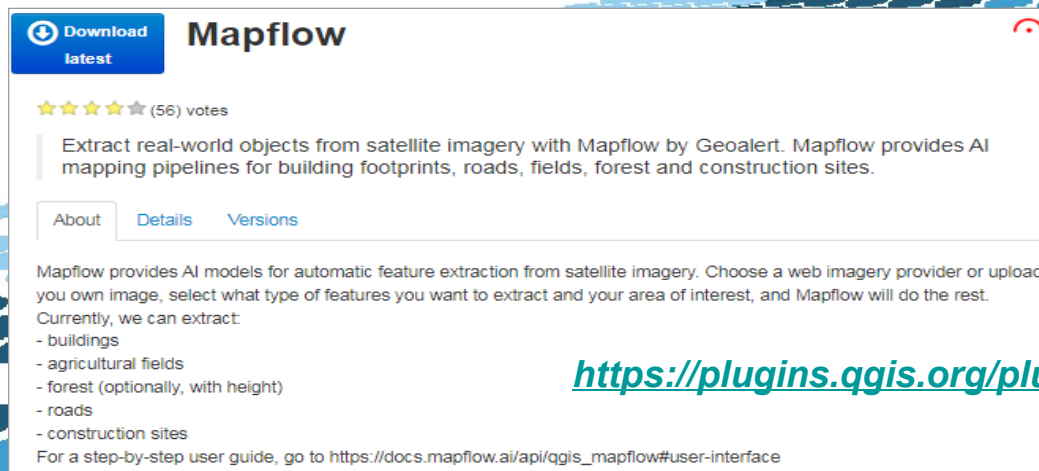
- Изучить доступные сервисы/приложения с открытым кодом для распознавания дорог;
- Оценить возможности современных нейросетей в области распознавания дорог;
- Протестировать применимость нейросетей, получить результаты распознавания дорог для тестовой территории в Красноярском крае.

# Плагин Open Source QGIS Mapflow

MapFlow для QGIS является одним из используемых и популярных по данным библиотеки «открытых» инструментов для геоинформатики (<https://plugins.qgis.org/>) плагинов тематики искусственного интеллекта.

Согласно документации, обучение модели сегментации дорог выполнено преимущественно для сельской местности по данным из космоса уровня пространственного разрешения порядка 0.3 - 0.5 м с использованием улучшения связности отдельных сегментов дороги ([https://ru.docs.mapflow.ai/userguides/models\\_changelog/index.html#id9](https://ru.docs.mapflow.ai/userguides/models_changelog/index.html#id9)).

Для тестов выбраны территории трех типичных (различных по типам дорог, внутренней планировке и наличию подъездов по дорогам) населенных пунктов Ермаковского лесничества Красноярского края. Была использована демо-версия плагина, имеющая ограничения по площади обработки территории.



**Download latest** **Mapflow**

★★★★☆ (56) votes

Extract real-world objects from satellite imagery with Mapflow by Geoalert. Mapflow provides AI mapping pipelines for building footprints, roads, fields, forest and construction sites.

About Details Versions

Mapflow provides AI models for automatic feature extraction from satellite imagery. Choose a web imagery provider or upload your own image, select what type of features you want to extract and your area of interest, and Mapflow will do the rest. Currently, we can extract:

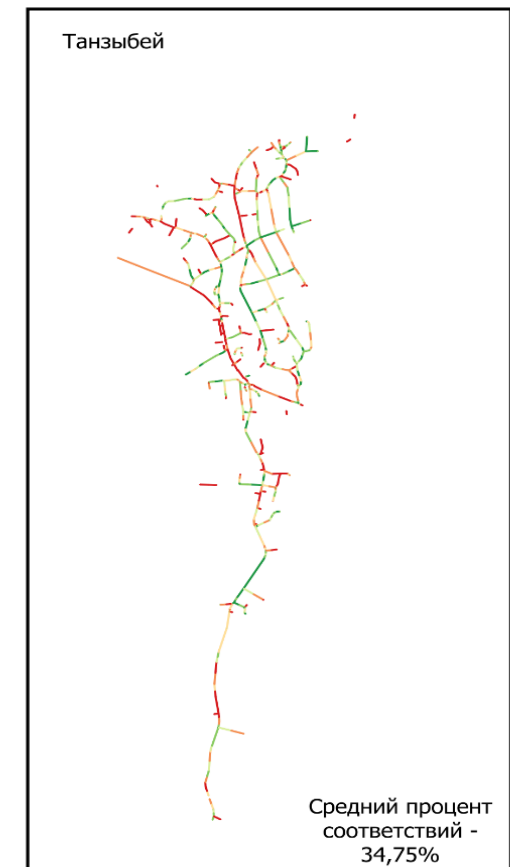
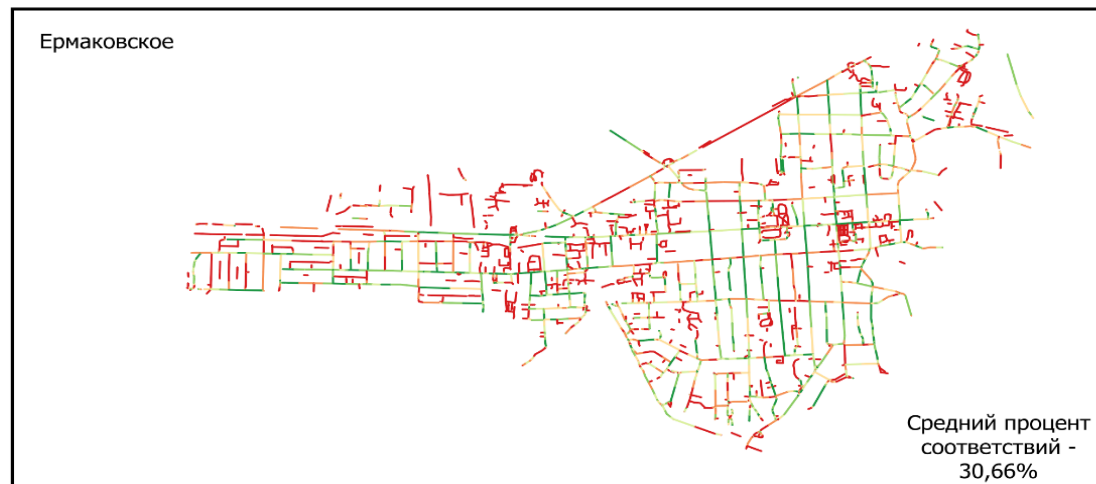
- buildings
- agricultural fields
- forest (optionally, with height)
- roads
- construction sites

For a step-by-step user guide, go to [https://docs.mapflow.ai/api/qgis\\_mapflow#user-interface](https://docs.mapflow.ai/api/qgis_mapflow#user-interface)

<https://plugins.qgis.org/plugins/mapflow/>

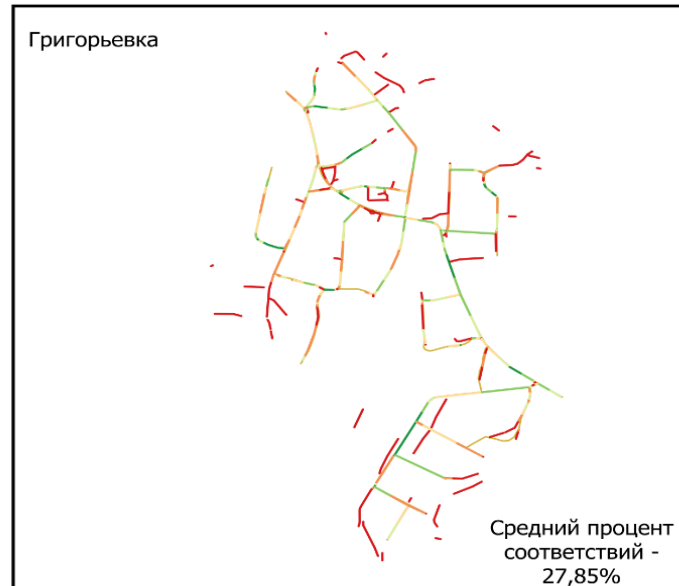
# Результаты применения Mapflow (1)

Совпадения дорог, сгенерированных MapFlow по снимкам Марбокс и дорог проекта OSM. Буфер слоя OSM 1.5 м



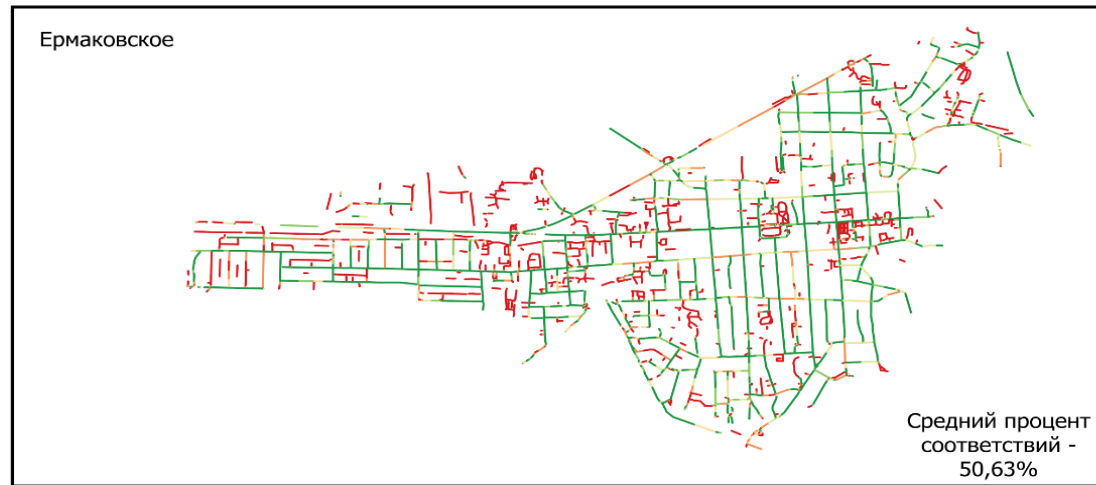
Процент соответствий

- 0
- 1- 20
- 21 - 40
- 41 - 60
- 61 - 80
- 81 - 100



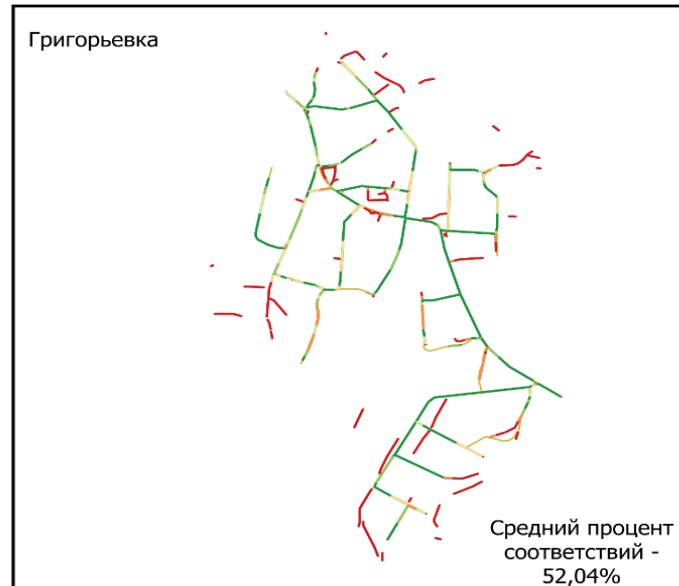
# Результаты применения Mapflow (2)

Совпадения дорог, сгенерированных MapFlow по снимкам Марбокс и дорог проекта OSM. Буфер слоя OSM 3 м



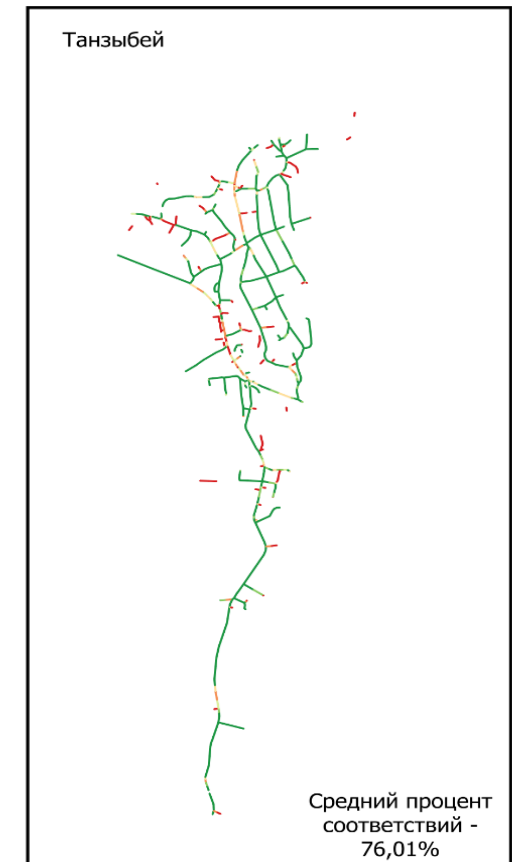
Процент соответствий

- 0
- 1- 20
- 21 - 40
- 41 - 60
- 61 - 80
- 81 - 100



# Результаты применения Mapflow (3)

Совпадения дорог, сгенерированных MapFlow по снимкам Mapbox и дорог проекта OSM. Буфер слоя OSM 4.5 м



Процент соответствий

- 0
- 1- 20
- 21 - 40
- 41 - 60
- 61 - 80
- 81 - 100



# Выводы по использованию плагина Mapflow

- Несмотря на открытый код самого плагина, у пользователя нет доступа к самим моделям нейросетей и к спутниковым изображениям сверхвысокого разрешения.
- Плагин функционирует на системе расходуемых “токенов” (эквивалентных реальным деньгам). Чем больше их у пользователя, тем большие площади ему доступны к дешифрованию.
- Такая популярная на сегодняшний день концепция условно бесплатного использования делает затруднительным применение таких данных для проектов на региональном уровнях.



# Что делать?

Для обучения нейросети необходимо два набора данных – спутниковые данные, с которых нейросеть будет считывать дешифровочные признаки дорог, и маски истины дорог, по которым будет производиться обучение.

Подходящими для этого данными могут быть:

- спутниковые снимки системы Sentinel 2 (10 м);
- и растрованная векторная основа дорог Open Street Map.



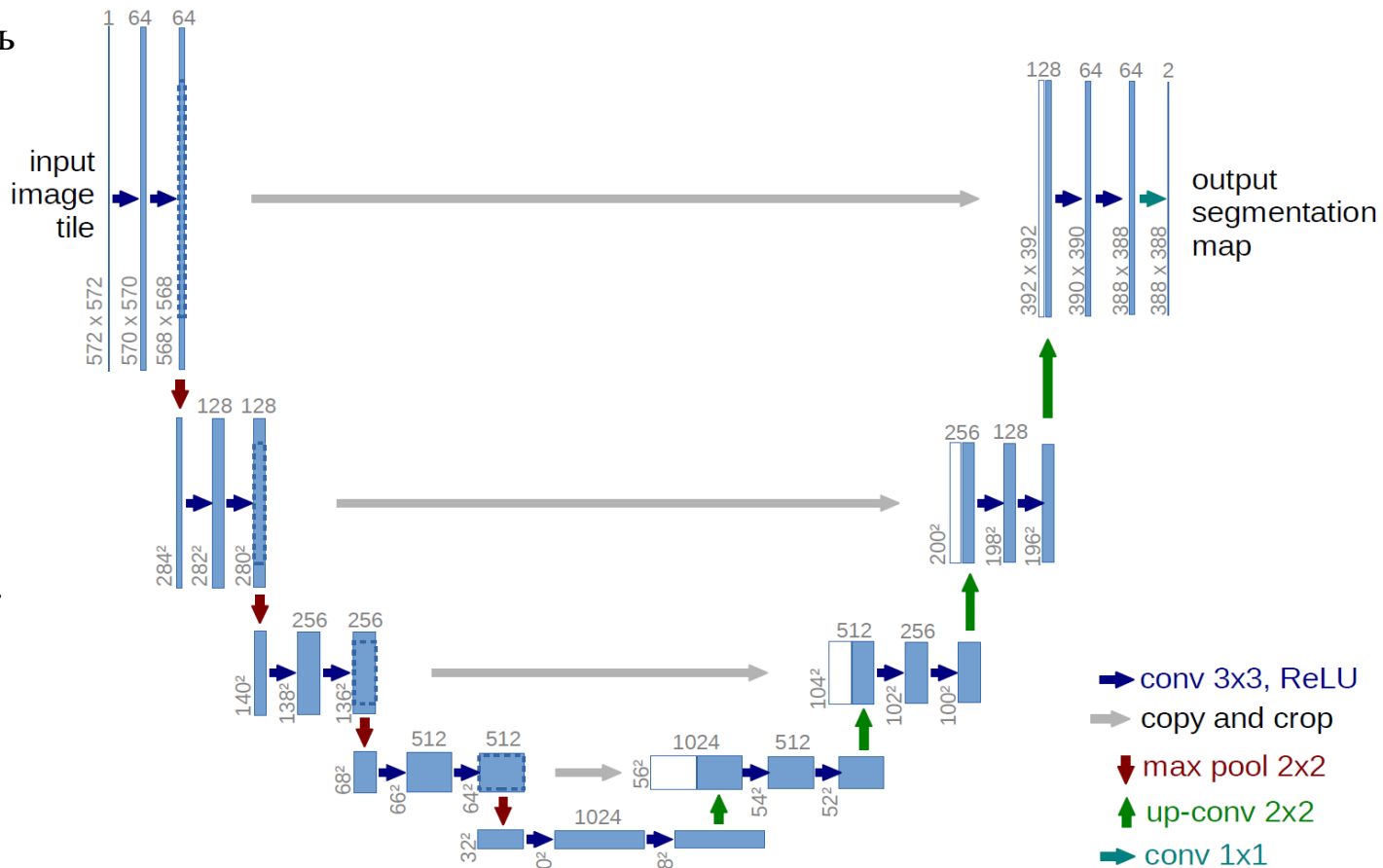
**OpenStreetMap**

# Архитектура сверточной сети U-Net

Необходимо сформировать датасет для обучения сверточной нейросети из:

- серии спутниковых изображений;
- серии масок истины.

Каждое изображение на входе имеет размер  $A \times B$ , где  $A=B$  ( $128 \times 128$ ), ( $256 \times 256$ ) и т.д.

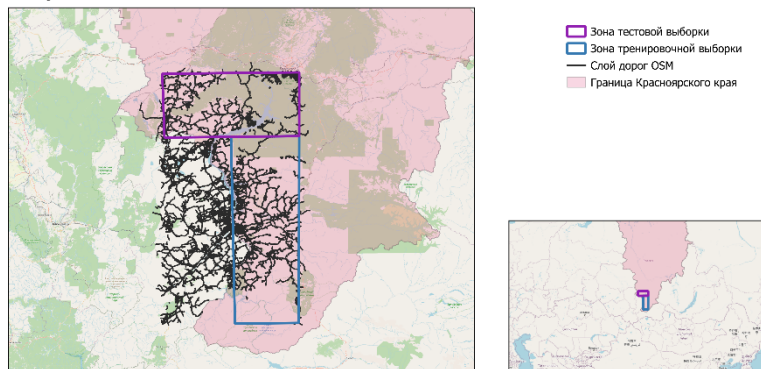


U-Net - это архитектура сверточной нейронной сети, которая была разработана в 2015 году Олафом Роннебергером, Филиппом Фишером, Томасом Броксом для сегментации клеток на микроскопических снимках.

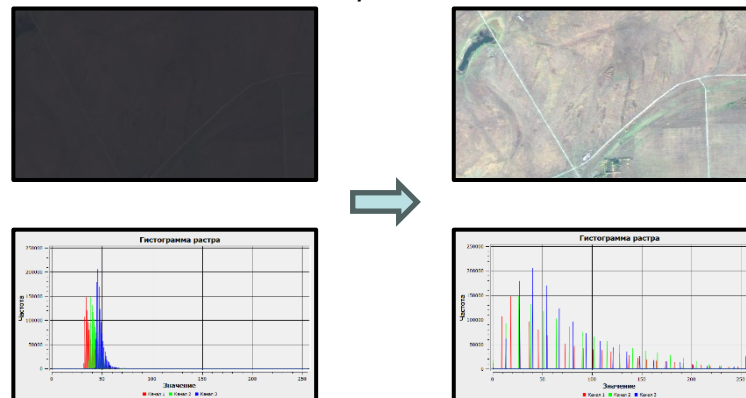
<https://cyberleninka.ru/article/n/primeneniye-nei-ronnoi-seti-arhitektury-u-net-dlya-segmentatsii-stm-izobrazhenii>

# Подготовка данных для обучения

## 1) Определение зон тренировочной и тестовой выборки



## 2) Контрастирование спутниковых изображений

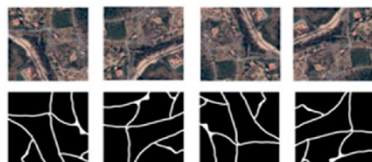


## 4) Аугментация данных

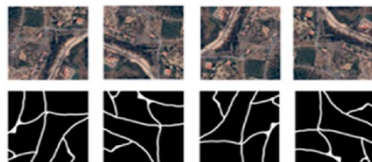
(1) 486 Image tiles of 1024x1024  
(Aerial Images, Masks)



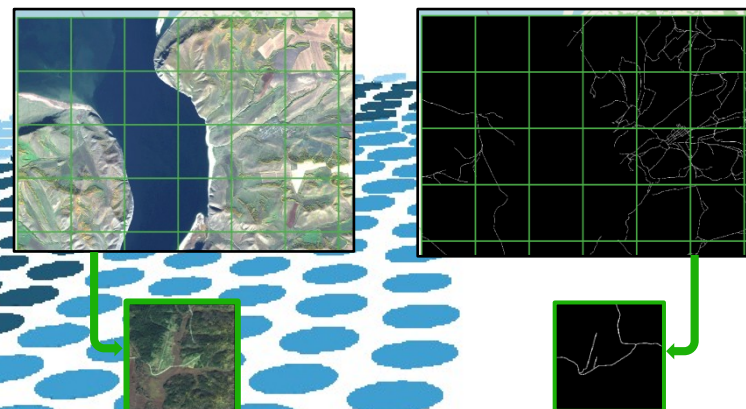
(2) 4x Rotation of 90°  
(1944 Aerial Images)  
(1944 Masks)



(3) 1x Mirror horizontal  
(3888 Aerial Images)  
(3888 Masks)



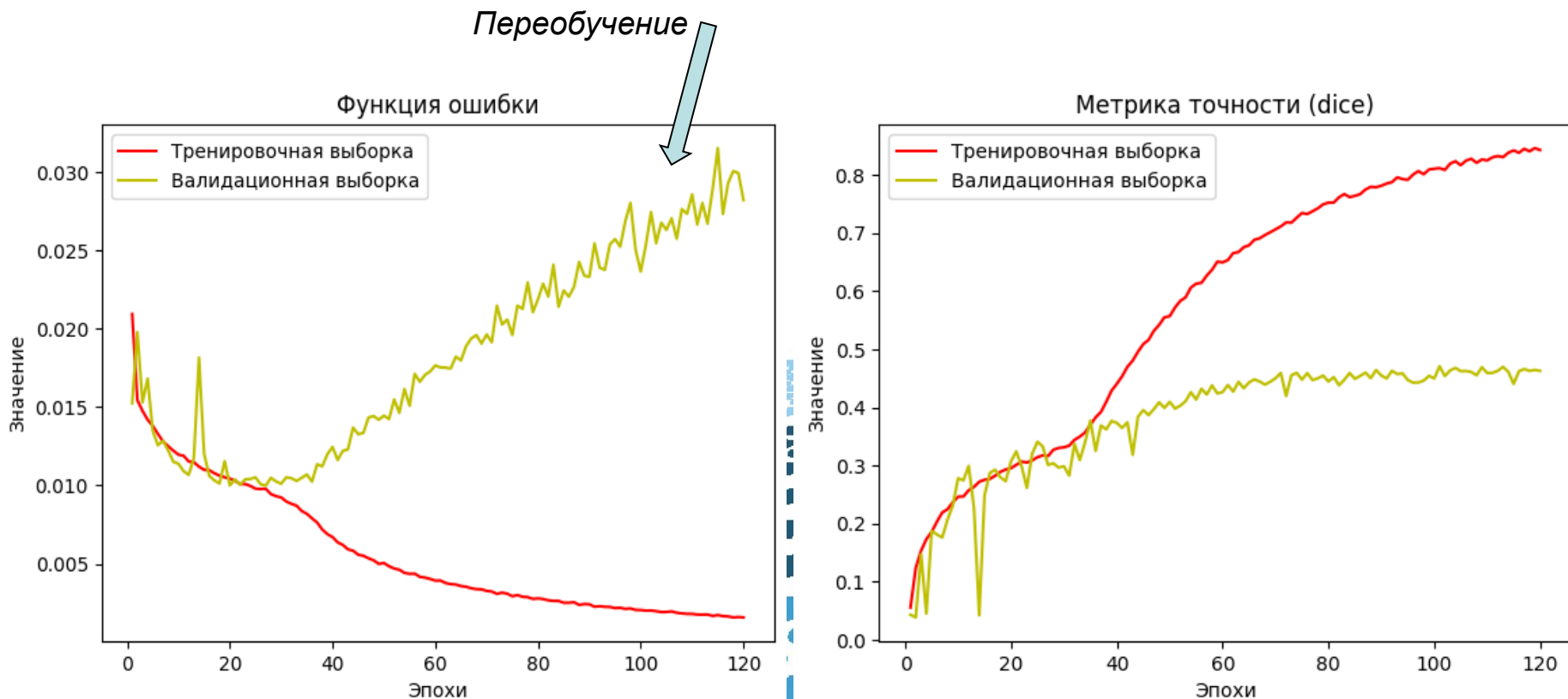
## 3) Разбиение изображений на тайлы 256x256



# Параметры обучения

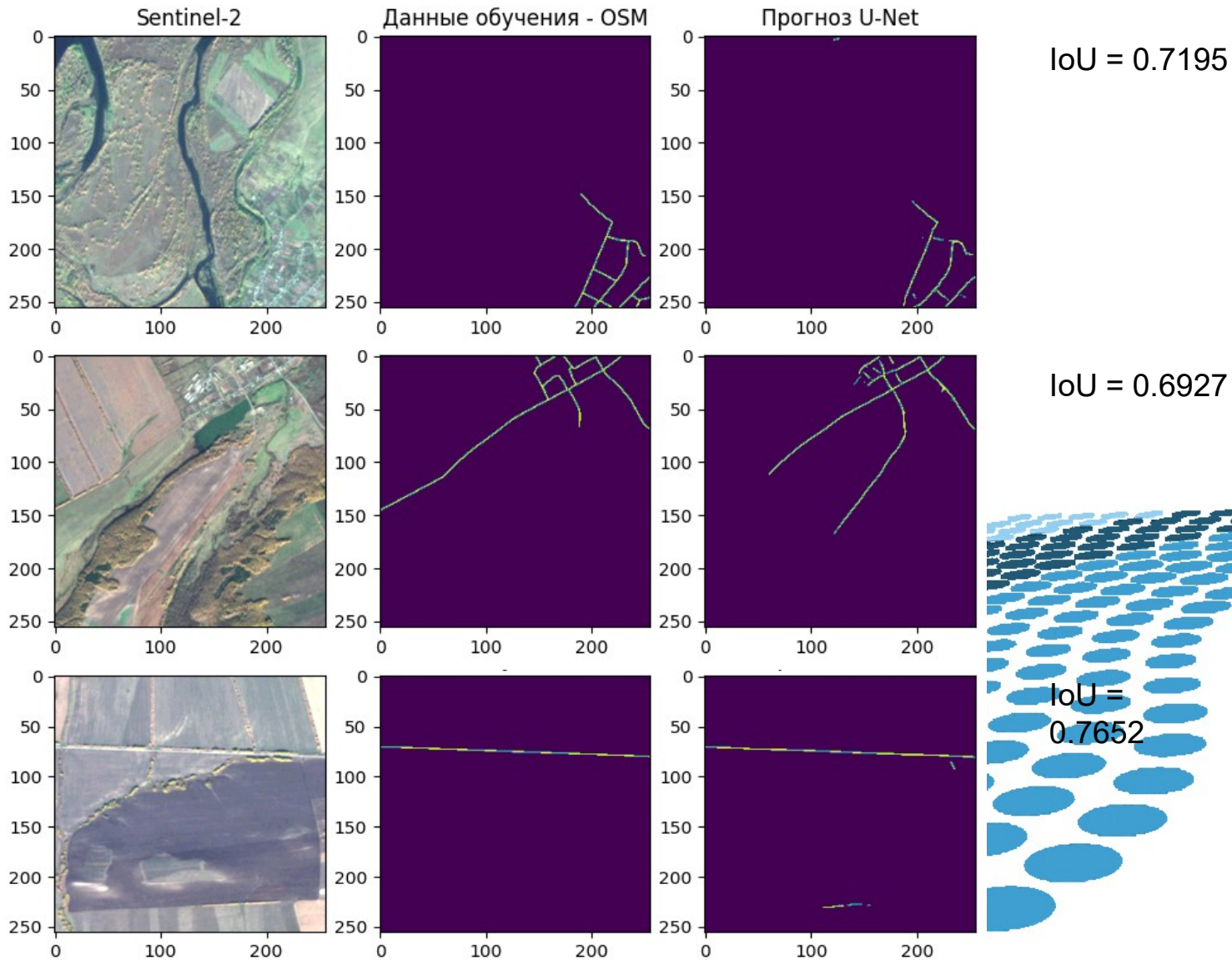
- Нейросеть была обучена на локальной машине на GPU NVidia Geforce Gtx 1080 Ti 11Gb. Оптимизатор градиентного спуска – Adam ( $lr=0.001$ ), функция ошибки – бинарная кросс-энтропия, метрика точности – коэффициент dice. Количество эпох = 120.

- Спутниковые данные тренировочной выборки были разделены на собственно тренировочную выборку и валидационную выборку в соотношении 90 % - 10%

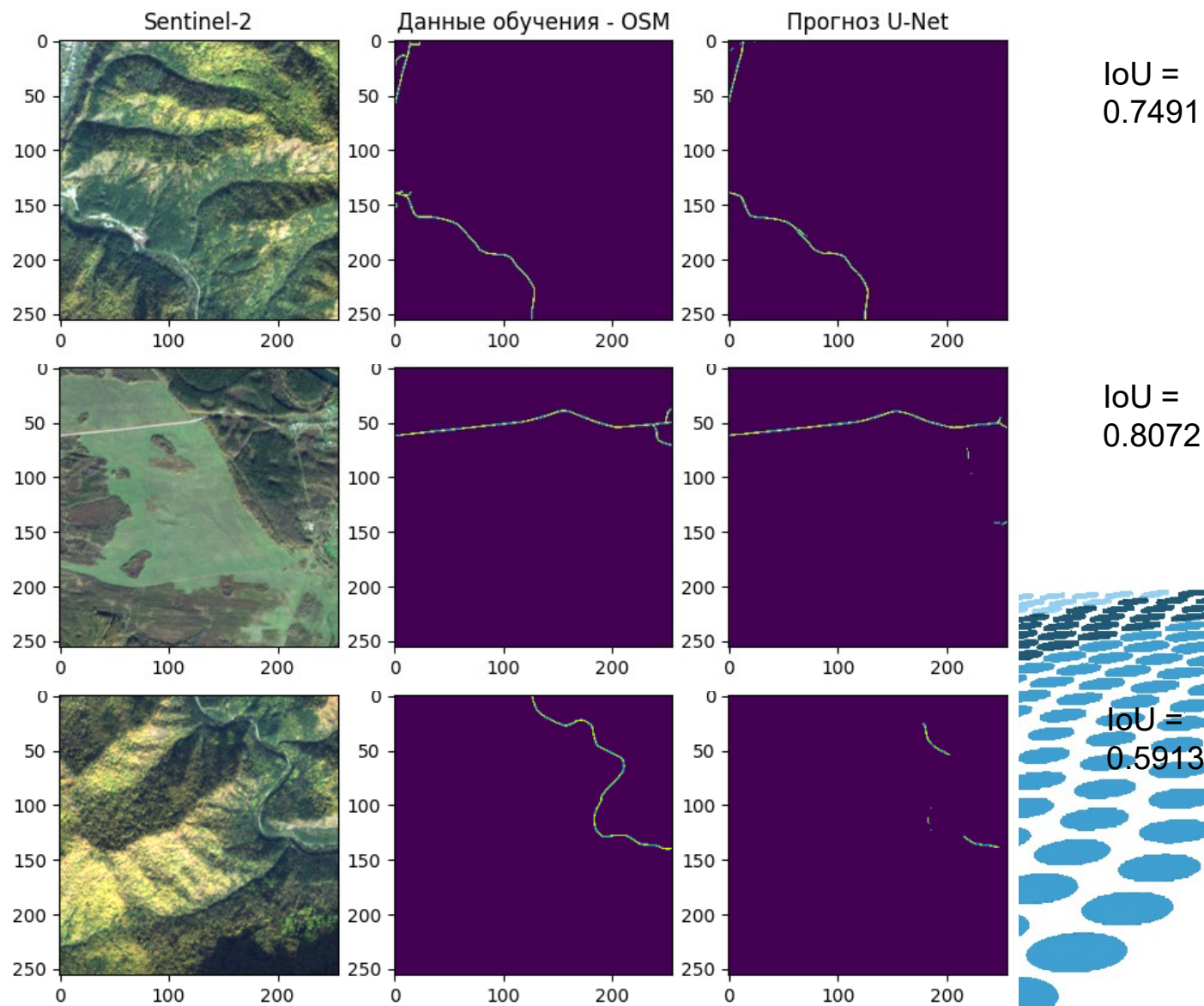


# Результаты: тестовая выборка (1)

Средние значения для тестовой выборки:  $IoU = 0.6228$ ;  $F1(dice) = 0.5725$ .

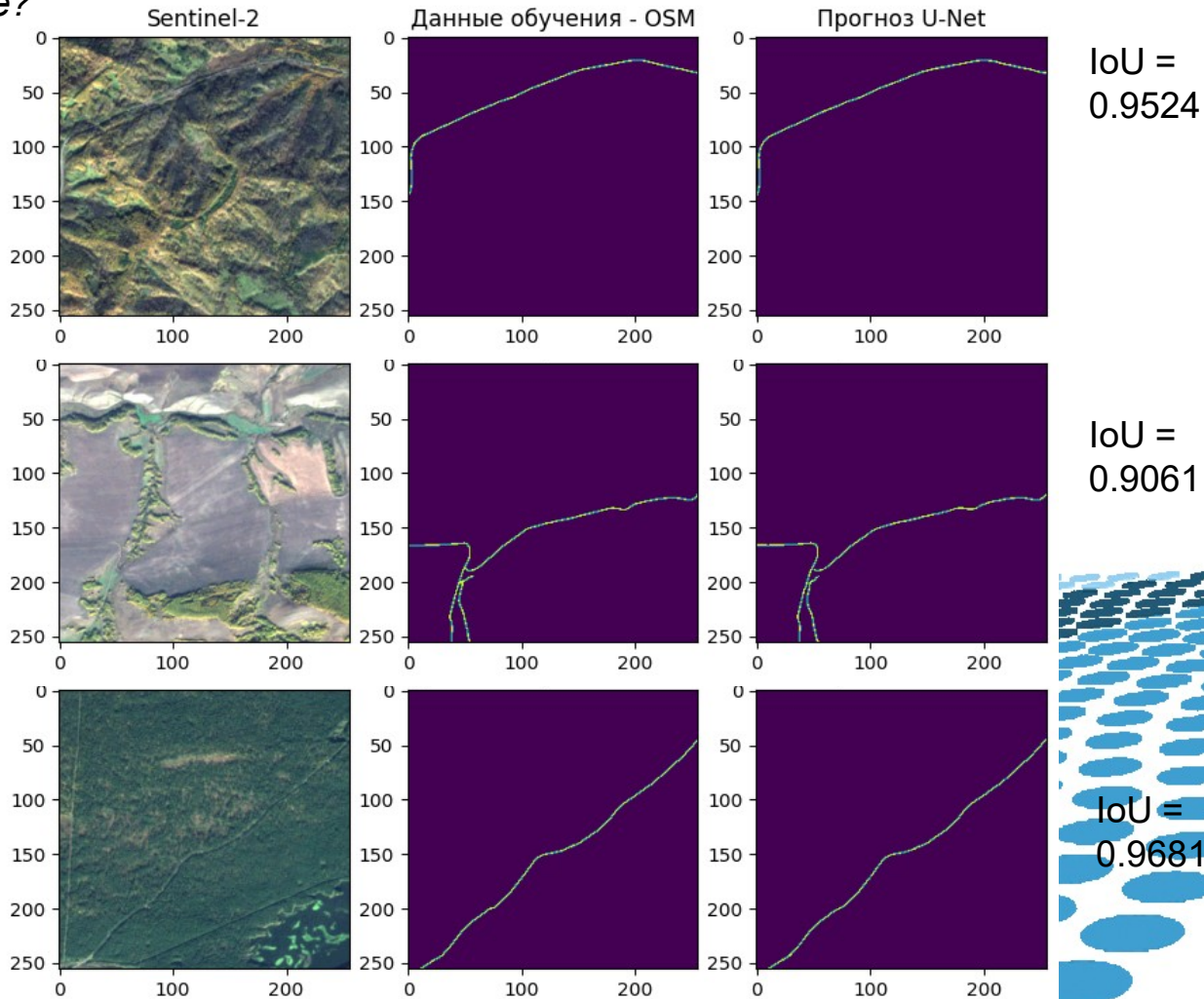


# Результаты: тестовая выборка (2)



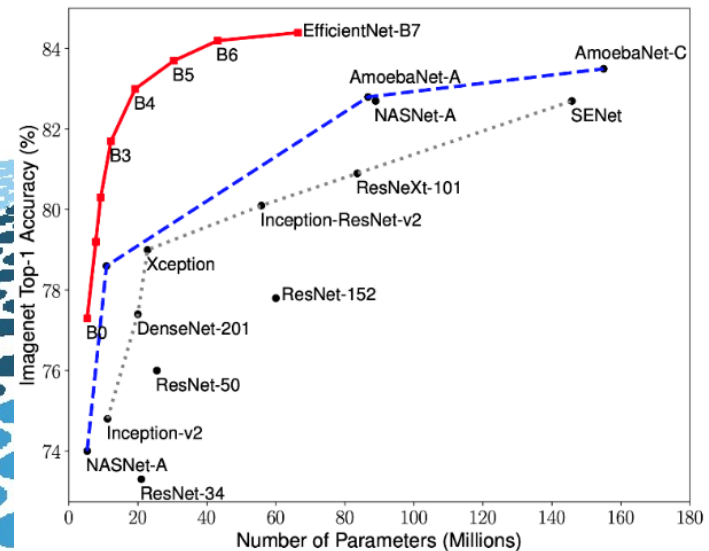
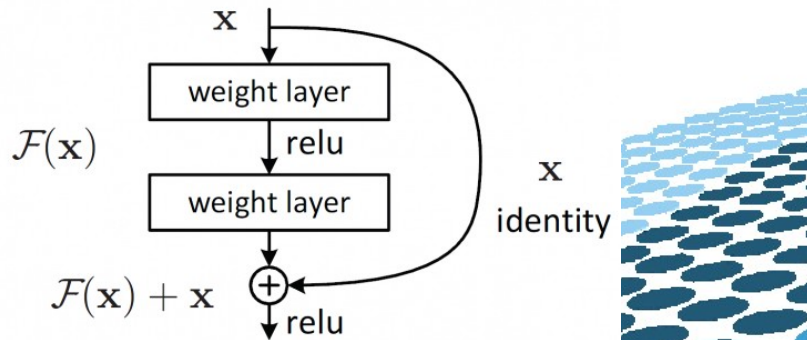
# Результаты: тренировочная выборка

Средние значения для тренировочной выборки:  $IoU = 0.9025$ ;  $F1(dice) = 0.9054$ . Целесообразно ли измерять точность на тренировочной выборке?



# Предобученные энкодеры

- Частым вариантом дополнения классической архитектуры U-Net является замена оригинального энкодера на какую-либо другую сверточную сеть. В пакете Keras (<https://keras.io/>) доступно множество таких энкодеров, в дополнение к этому они заранее предобучены на датасете ImageNet. Теоретически использование предобученных энкодеров поможет быстрее по времени достигнуть минимума функции ошибки и максимума метрики точности.
- Для сравнения с оригинальной архитектурой были выбраны энкодеры ResNet50, EfficientNet и EfficientNetV2.
- ResNet50 увеличивает точность и глубину обучения за счет соединений быстрого доступа (shortcut connections)
- Модели же EfficientNet построены по принципу compound scaling (балансирования между собой глубины и количества каналов сети, а также разрешения изображений в сети), тем самым одновременно сокращая количество расчетов и увеличивая точность.



K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

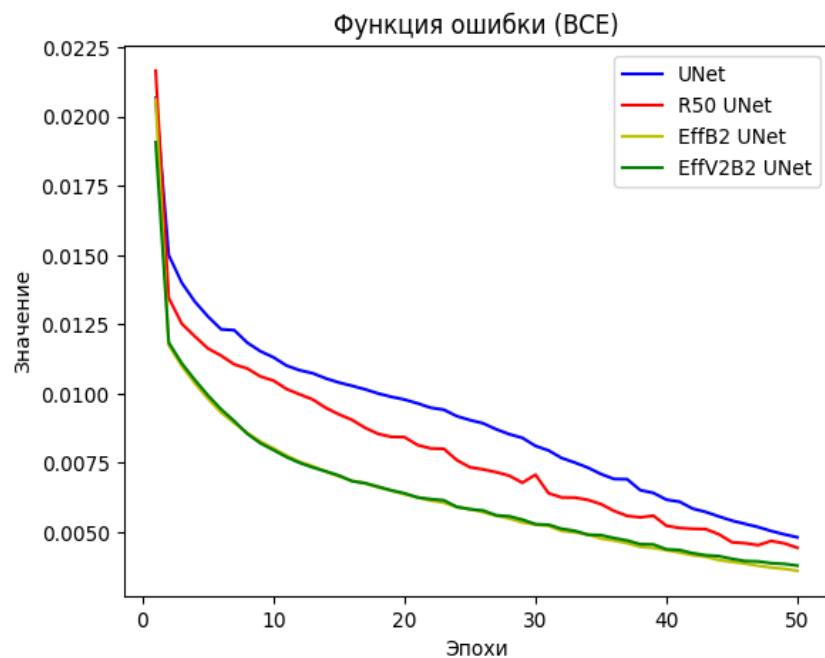
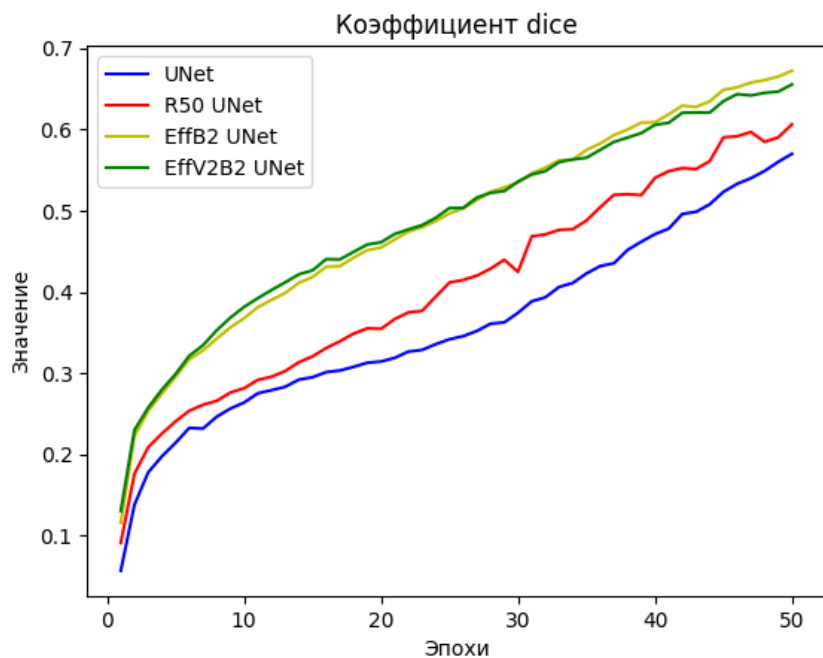
Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.



# Предобученные энкодеры: обучение

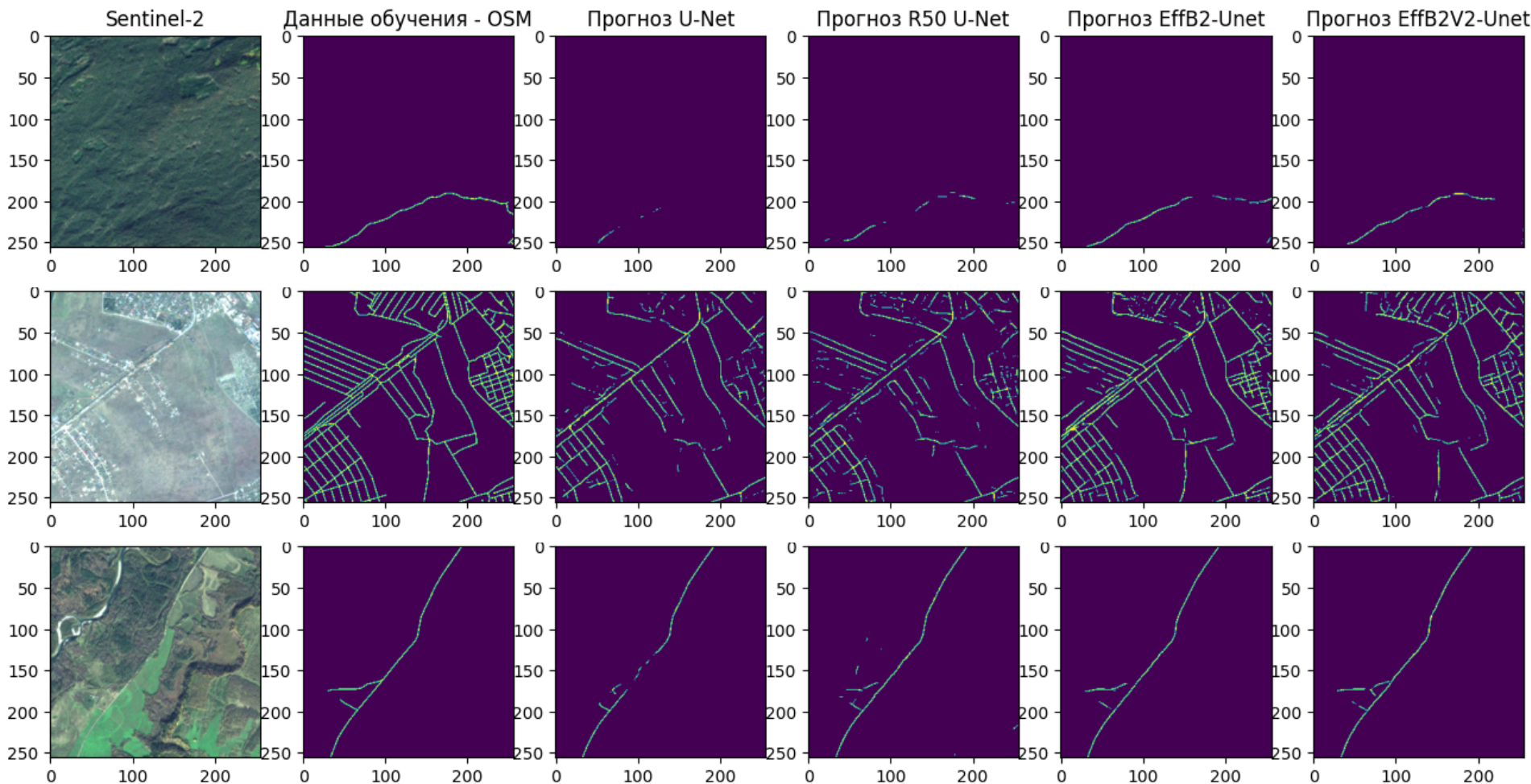
## Время обучения (часы)

UNet	R-50 UNet	EffB2 UNet	EffV2B2 UNet
22,63	20,67	20,53	16,16



# Предобученные энкодеры: результаты

В данном случае для сравнения используется тренировочная выборка, так как цель – сравнение точности относительно референц-датасета

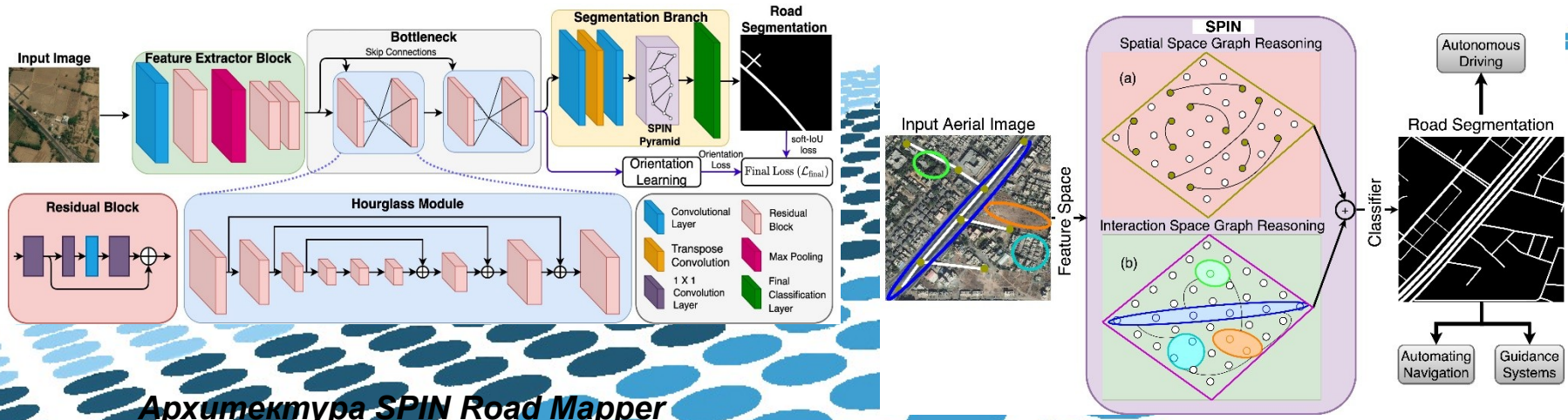


IoU

	U-Net	R50 U-Net	EffB2 U-Net	EffB2V2 U-Net
	0.5218	0.5400	0.6781	0.6902
	0.6294	0.6322	0.7315	0.7198
	0.6912	0.7623	0.8239	0.8223

# Поиск архитектур

- Архитектура UNet была представлена в 2015 году. Учитывая бурное развитие технологий в сфере Deep Learning, встает вопрос применения более современных и точных архитектур нейронных сетей.
- Самыми широко используемыми архитектурами сверточных нейросетей, помимо UNet, являются DeepLabV3+ и FPN, кроме того существует большое количество модификаций самой UNet, такие как UNet++, AttentionUnet, ResUnet, TransUnet, SwinUnet (последние две комбинируют элементы сверточных сетей и vision transformers)
- На сегодняшний день уже существуют архитектуры, предложенные целенаправленно для распознавания дорог, к примеру SPIN Road Mapper и D-LinkNet.
- Комбинируя элементы базовых архитектур, таких как UNet, и узконаправленных, например, SPIN Road Mapper, можно будет построить наиболее оптимальную архитектуру для такой узкой задачи.



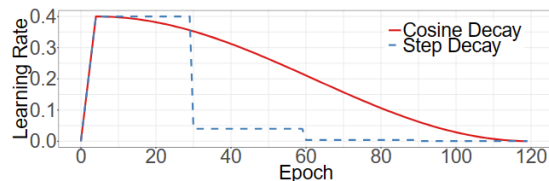
# Настройка гиперпараметров нейросети

## Learning Rate Schedule:

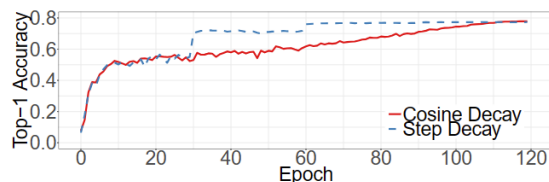
- Linear scaling Learning Rate (линейное увеличение learning rate по размеру батча)
- Learning rate warmup
- Cosine Learning Rate Decay

## Regularization:

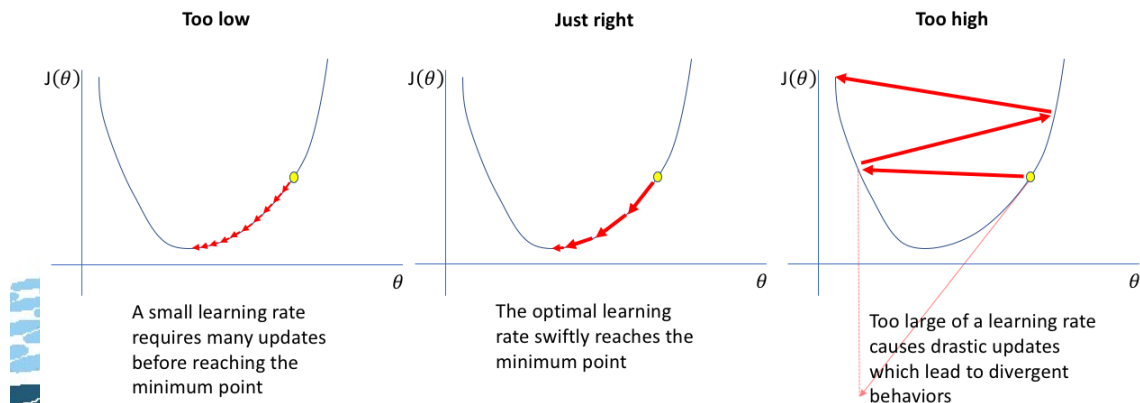
- Zero  $\gamma$  in Batch Normalization
- No bias decay
- Label Smoothing
- Mixup Data Augmentation



(a) Learning Rate Schedule



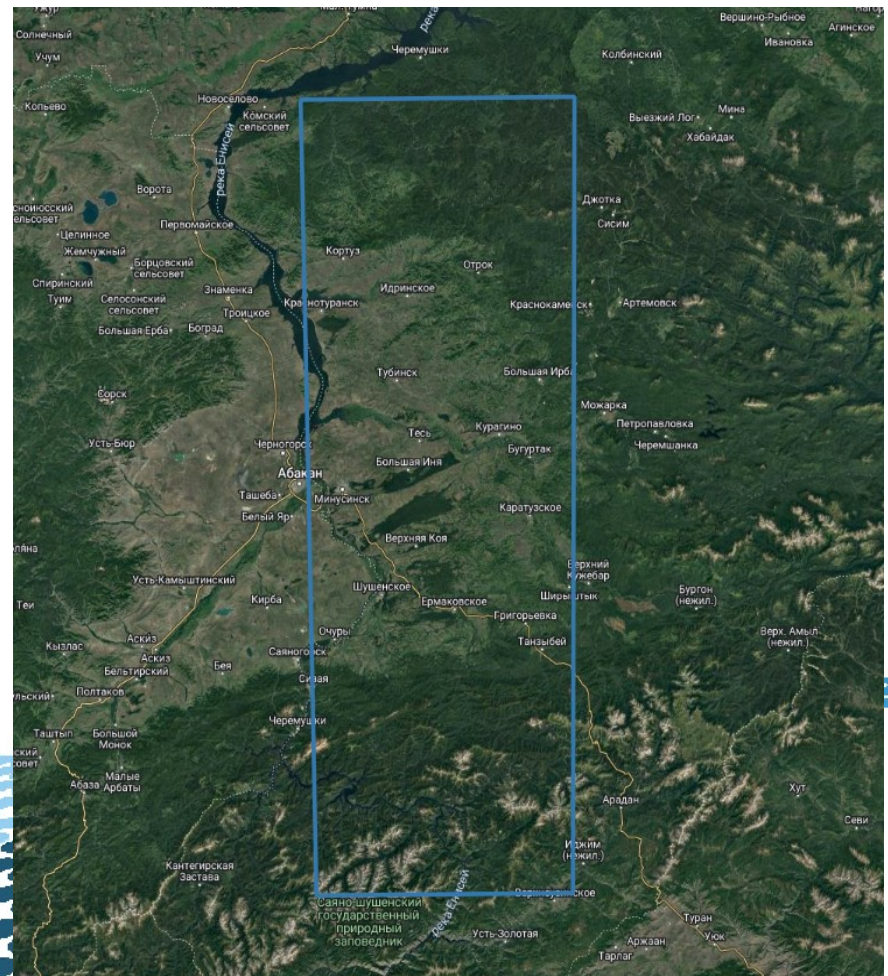
(b) Validation Accuracy



He T. et al. Bag of tricks for image classification with convolutional neural networks //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2019. – С. 558-567.

# Подготовка спутниковых данных (1)

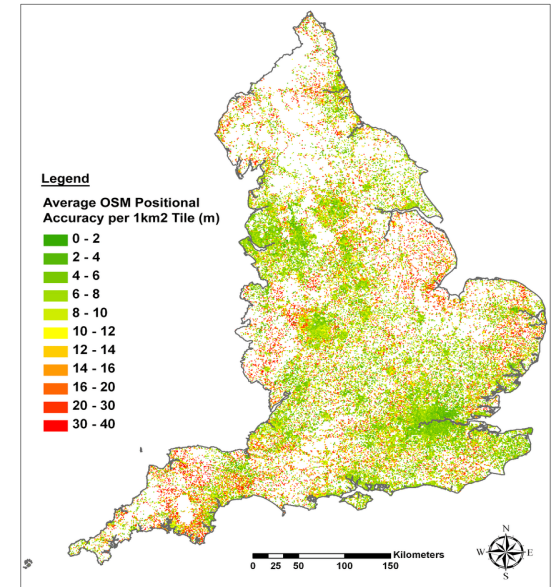
- Важно понимать, насколько узконаправленная/обобщающая нейросеть нам нужна.
- От этого будет зависеть разнообразие спутниковых данных, поступающих в нейросеть для обучения.
- В данном случае, тематика применения достаточно узка – распознавание дорог в конкретном регионе. Ввиду этого выборка возможных данных для обучения существенно ограничивается.



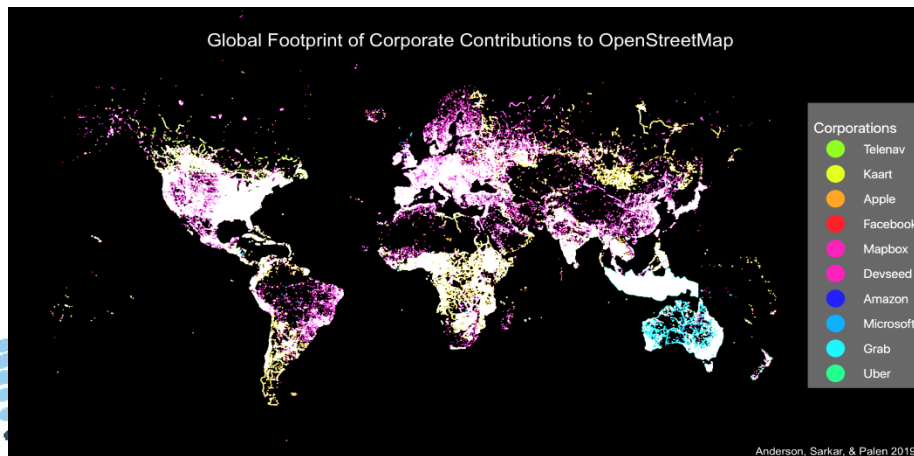
*Google Satellite изображение на исследуемую территорию*

# Подготовка спутниковых данных (2)

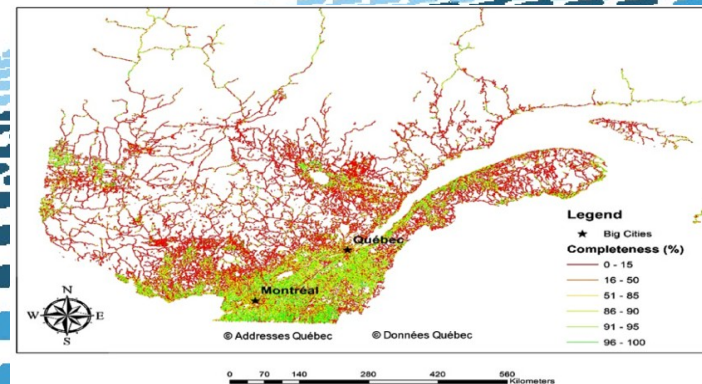
- Одной из ключевых проблем является отсутствие “истинных” данных о проселочных и грунтовых дорогах на территорию исследования в открытом доступе.
- Потенциальным выходом из такой ситуации может быть использование спутниковых данных и данных OSM для обучения нейросети на окрестностях крупных городов России, Канады, скандинавских стран находящихся в тех же природных и климатических зонах.
- Существует множество научных работ, показывающих, что в урбанизированных регионах данные OSM являются наиболее полными.
- Главный недостаток такого метода – ручной подбор территории и спутниковых данных



Antoniou, Vyron. (2011). User generated spatial content: an analysis of the phenomenon and its challenges for mapping agencies.



Jennings, Sarkar, and Palen, “Corporate Editors in the Evolving Landscape of OpenStreetMap.” *International Journal of Geo-Information*, April 2019.



Moradi, Milad & Roche, Stéphane & Mostafavi, Mir Abolfazl. (2022). Exploring five indicators for the quality of OpenStreetMap road networks: a case study of Québec, Canada. *Geomatica*. 00. 1-31. 10.1139/geomat-2021-0012.

# Итоги и направления продолжения работы

- Использование нейросетей для распознавания дорог является перспективным направлением.
- Архитектура UNet и представленные ее модификации неплохо показали себе в задаче распознавания дорог на разрешении 10м. Однако, одним из факторов такой сравнительно высокой точности может являться специфика датасета обучения OSM. Из-за этого, модели лучше всего “подстроились” под распознавание асфальтированных дорог регионального и местного значения.
- Некоторые манипуляции со спутниковыми данными, тонкой настройкой параметров обучения нейросетей потенциально могут увеличить точность распознавания дорог.
- Дальнейшие исследования будут посвящены, главным образом, подбору наиболее оптимальной архитектуры (с точки зрения точности и количества вычислений) нейронной сети для такой задачи.

*Работа выполнена в рамках темы Государственного задания «Методические подходы к оценке структурной организации и функционирования лесных экосистем», регистрационный номер № 121121600118-8.*