



Влияние состава выборок аэрокосмических изображений ДЗЗ высокого и сверхвысокого пространственного разрешения на обучение и точность нейронных сетей при семантической сегментации геополей на примере распознавания различных классов земной поверхности

Н.А. Бирюков¹✉

АФФИЛИАЦИИ

¹ Московский государственный университет геодезии и картографии, Москва, Россия
✉ 4096414@gmail.com

ЦИТИРОВАНИЕ

Бирюков Н.А. Влияние состава выборок аэрокосмических изображений ДЗЗ высокого и сверхвысокого пространственного разрешения на обучение и точность нейронных сетей при семантической сегментации геополей на примере распознавания различных классов земной поверхности // Пространственные данные: наука и технологии. 2025. Т. 16, №2. С. 30–57. DOI:10.30533/scidata-2025-16-12.

КЛЮЧЕВЫЕ СЛОВА

геополе, множество геополей, семантическая сегментация, состав выборок данных, дисбаланс классов, нейронная сеть, точность распознавания

АННОТАЦИЯ

Выборки из аэрокосмических изображений и масок, используемые при решении задач распознавания различных классов земной поверхности, могут оказывать существенное влияние на обучаемость моделей нейронных сетей и получаемые в дальнейшем

с их помощью результаты распознавания. Состав выборок данных в большинстве случаев рассматривается не относительно самих выборок, а с точки зрения обработки данных нейронными сетями в целом в конкретной задаче. В контексте семантической сегментации геополей сформулированы общие для задач семантической сегментации объектов на аэрокосмических изображениях проблемы: разные яркостные характеристики снимков, тени, эквивалентность яркостных характеристик объектов целевого класса и других объектов сцен, некорректная разметка, граничные случаи дисбаланса классов. Все перечисленное рассматривается как проблемы представления исходного множества геополей в выборках данных. В результате эксперимента с нейронными сетями U Net, STT и MF-CNN определено, что включаемые в выборки граничные случаи дисбаланса классов и применение снимков с разрешением, при котором дисбаланс классов выше, чем при использовании частей снимков, существенно снижают обучаемость нейронных сетей и точность распознавания, а отбор данных на основе удаления граничных случаев дисбаланса классов при предобработке позволяет как повысить точность распознавания, так и снизить необходимые для обучения моделей временные затраты.

1 Введение

Выборки данных, используемые для обучения, валидации и тестирования, оказывают прямое влияние как на обучение моделей нейронных сетей, так и на получаемые в дальнейшем результаты распознавания различных классов объектов. Наиболее распространенными при решении задач семантической сегментации аэрокосмических изображений являются классы «Здания», «Дорожная сеть», «Водные объекты» и «Растительность». Влияние состава аэрокосмических изображений и масок для распознавания указанных классов, в который могут входить снимки с разными яркостными характеристиками [1–4], цифровые изображения с эквивалентностью яркостных характеристик объектов целевого класса другим объектам сцены [3–6] или аэрофото- и космоснимки с различием в яркостных характеристиках распознаваемых объектов [4, 5] и т. д., рассматривается главным образом на уровне включения концептуальных и прикладных решений в рамках архитектур нейронных сетей. Это, с одной стороны, приводит к развитию существующих решений, а с другой — затрудняет оценку получаемых результатов, т. к. даже с учетом развития концептуальных решений, появления новых архитектур нейронных сетей и включения разных модулей выборки данных оказывают влияние как на обучение моделей, которое определяет возможность дальнейшего использования, так и на количественные результаты метрик оценки точности распознавания при валидации и тестировании.

Известно, что при предобработке цифровых аэрокосмических изображений и масок в выборках производится деление аэрокосмических снимков

и соответствующих им эталонов на части [1, 7–9] и аугментация данных [1, 2, 10]. Однако предшествующий делению на части (или следующий за ним) отбор снимков и масок, который может приводить к изменению состава выборки данных, и само деление во многих случаях опускаются или не рассматриваются как необходимые, уступая свое место количеству данных в выборке, являющемуся важным фактором для успешного решения задач глубокого обучения. Это может приводить к ситуации, когда граничные случаи дисбаланса классов (т. е. случаи, когда объект целевого класса занимает все цифровое изображение или отсутствует на снимке) могут стать основой состава выборки данных, а упущенная некорректная разметка или снимки с разными яркостными характеристиками, соотношение которых в рамках состава выборки не сбалансировано, могут приводить к существенному снижению обучаемости и точности распознавания моделей нейронных сетей для семантической сегментации дорожной сети, зданий, водных объектов и растительности.

Целью данной работы является определение в рамках семантической сегментации геополей, объединяющей задачи семантической сегментации объектов на аэрофото- и космоснимках, влияния состава выборок данных на результаты обучения, валидации и тестирования моделей нейронных сетей на примере распознавания дорожной сети, зданий, растительности и водных объектов на аэрокосмических изображениях высокого и сверхвысокого пространственного разрешения.

2 Материалы и методы

Для формулирования общей задачи семантической сегментации геополей посредством нейронных сетей определим на базе изложенного в [11], а также на основе предпосылок в действующих государственных стандартах Российской Федерации (ст. 3.1.1, 3.1.20, 3.1.21 ГОСТ Р 59079-2020¹, ст. 23–25 ГОСТ Р 59753-2021², ст. 3.1.7 ГОСТ Р 58849-2020³ и ст. 3.1.12, 3.1.21 ГОСТ Р 59328-2021⁴) следующее:

- цифровой космический снимок — графическое представление для некоторого участка земной поверхности множества геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного

1 ГОСТ Р 59079-2020. Данные дистанционного зондирования Земли из космоса. Типы данных дистанционного зондирования Земли из космоса. М.: Стандартинформ, 2020. 12 с.

2 ГОСТ Р 59753-2021. Данные дистанционного зондирования Земли из космоса. Термины и определения. М.: Российский институт стандартизации, 2021. 20 с.

3 ГОСТ Р 58849-2020. Авиационная техника гражданского назначения. Порядок создания. Основные положения. М.: Стандартинформ, 2020. 62 с.

4 ГОСТ Р 59328-2021. Аэрофотосъемка топографическая. Технические требования. М.: Стандартинформ, 2021. 36 с.

излучения в определенных спектральных диапазонах длин волн, полученное с космического аппарата в единой геометрии съемки на одну дату и время;

- цифровой аэрофотоснимок — графическое представление для некоторого участка объекта аэрофотосъемки множества геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного излучения в определенных спектральных диапазонах длин волн, полученное с воздушного судна в единой геометрии съемки на одну дату и время;
- результат распознавания модели нейронной сети — цифровое изображение, являющееся графическим представлением для некоторого участка земной поверхности геополя пространственного распределения собственного и отраженного одним или несколькими классами объектов, которые могут являться элементами суши, океана и атмосферы Земли, электромагнитного излучения в определенных спектральных диапазонах длин волн.

Таким образом, семантическая сегментация геополей посредством нейронных сетей представляет собой переход от исходного множества геополей, представленного цифровым аэрокосмическим изображением, к результату распознавания, который является графическим представлением результирующего множества геополей, что запишем в общем виде как

1

$$G_{RGFS} = Q(G_{OGFS}),$$

где G_{RGFS} — результирующее множество геополей;

G_{OGFS} — исходное множество геополей;

Q — оператор связи, представленный сложной нелинейной функцией, реализуемой нейронной сетью.

2.1 Выборки данных, используемые для различных задач семантической сегментации классов земной поверхности

В зависимости от решаемой задачи характеристики цифровых изображений и масок могут различаться. Основные характеристики выборок данных, используемых для различных задач семантической сегментации аэрокосмических изображений, в числе которых пространственное разрешение (ПР) и радиометрическое разрешение (РР), представлены в **табл. 1**.

Таблица 1 Основные характеристики выборок данных в задачах распознавания различных классов земной поверхности посредством нейронных сетей

Table 1 Main datasets characteristics in different earth's surfaces classes recognition with neural networks tasks

Задача	ПР, м	PP, бит/пиксель	Разрешение	Композиция	Тип снимка
1 Выделение зданий	0,2–1 1–10	8–24	256×256 512×512	RGB	аэрофото-, космический
2 Выделение водных объектов	0,2–1 1–10 10–100 100–1000		256×256 512×512 640×640	RGB NIR-RGB	космический
3 Распознавание растительности	0,2–1 1–10 10–100 100–1000		224×224 256×256 512×512	NIR-RGB RGB	аэрофото-, космический
4 Распознавание дорожной сети	0,2–1 10–100		256×256 512×512 1024×1024 1500×1500	RGB	

Примечание. Жирным шрифтом выделены наиболее распространенные значения характеристик в контексте каждой задачи.

Исходя из представленных значений основных характеристик выборок данных, можно сделать следующий вывод: в настоящее время наиболее часто при решении с помощью нейронных сетей задач, указанных в **табл. 1**, используются цифровые изображения сверхвысокого пространственного разрешения и соответствующие им маски.

2.2 Проблемы состава выборок данных в задачах распознавания различных классов земной поверхности на аэрофото- и космоснимках

В задачах семантической сегментации аэрокосмических изображений на основе объектов целевого класса можно выделить проблемы, влияющие на точность распознавания моделей нейронных сетей и определяющие развитие концептуальных и прикладных решений в каждой семантической сегментации объектов на аэрокосмических изображениях [2, 3, 5, 7]. Эти проблемы относятся и к составу

выборки данных, т. к. на основе обеспечения и проверки робастности нейронных сетей к поступающим в них данным в состав выборок включают снимки, отражающие разные случаи в рамках генеральной совокупности конкретной задачи распознавания классов земной поверхности. Кроме того, проблемы, связанные с метрической составляющей распознаваемого объекта или с его яркостными характеристиками (запечатленным собственным и отраженным электромагнитным излучением), могут быть присущи всем снимкам и поэтому будут присутствовать в любой выборке.

В задаче выделения объектов класса «Здания» на аэрокосмических изображениях в качестве проблем в составе выборок данных выделены:

- 1) разные яркостные характеристики зданий [5, 12, 13];
- 2) эквивалентность яркостных характеристик зданий и других объектов сцен [5, 12];
- 3) наличие на снимках теней, скрывающих здания [13, 14];
- 4) высокая плотность застройки городских территорий [11];
- 5) разные яркостные характеристики снимков [1, 15], взятых для разных областей интереса или полученных с помощью съемочной аппаратуры, имеющей разное спектральное и радиометрическое разрешение [16].

В задаче выделения объектов класса «Водные объекты» на космических и аэрофотоснимках проблемами, относящимися к составу выборок данных, являются:

- 1) разные яркостные характеристики водных объектов [17];
- 2) эквивалентность яркостных характеристик объектов сцен и водных объектов [3, 17];
- 3) наличие на снимках лесной растительности и теней [17, 18];
- 4) разные погодные условия и времена года [19];
- 5) наличие водных объектов, занимающих малую площадь на сцене и имеющих нечеткие границы [20–22];
- 6) разные яркостные характеристики снимков [18, 22], снятых с помощью аппаратуры, имеющей разное спектральное и радиометрическое разрешение, или полученных для разных областей интереса [3, 23, 24].

К основным проблемам состава выборок данных в задаче распознавания объектов класса «Растительность» отнесены:

- 1) разные яркостные характеристики растительности [25, 26], в том числе из-за того, что на снимках могут быть запечатлены разные этапы цветения [2];
- 2) эквивалентность яркостных характеристик растительности и других объектов сцен [10, 25, 26];
- 3) городские территории, на которых наличие растительности носит фрагментарный характер [27, 28];
- 4) разные яркостные характеристики цифровых изображений [27, 29, 30], снятых в разное время года.

В задаче выделения объектов класса «Дорожная сеть» для состава выборок данных характерны такие проблемы, как:

- 1) разные яркостные характеристики дорог [31, 32];
- 2) эквивалентность яркостных характеристик дорожной сети и других объектов [7, 33, 34];
- 3) наличие на снимках теней, скрывающих участки дорожной сети [4, 35, 36];
- 4) дисбаланс классов [37–39].

Следует отметить, что некоторые проблемы в составе выборок данных, могущие влиять на результат обучения, валидации, тестирования и дальнейшего использования моделей нейронных сетей для семантической сегментации классов земной поверхности, в научных публикациях в большинстве случаев не рассматриваются. **Среди таких проблем можно выделить:**

- 1) некорректную разметку снимков в выборках данных, которая может напрямую влиять на обучаемость моделей и на получаемые при валидации и тестировании результаты количественной оценки точности распознавания по метрикам оценки точности;
- 2) граничные случаи дисбаланса классов (когда объекты целевого класса занимают весь снимок или отсутствуют на нем) после деления цифровых снимков и масок на части при предобработке данных или имеющиеся на исходных цифровых изображениях.

2.3 Проблемы состава выборок данных для семантической сегментации геополей

В контексте семантической сегментации геополей посредством нейронных сетей указанные проблемы в составах выборок для задач распознавания различных классов земной поверхности относятся к проблемам представления исходного множества геополей G_{OGFS} в выборках данных.

Таким образом, проблемы состава выборок аэрокосмических изображений и масок для семантической сегментации геополей включают в себя:

- 1) присутствие на снимках теней;
- 2) разные яркостные характеристики объектов целевого класса;
- 3) эквивалентность яркостных характеристик объектов целевого класса и других объектов;
- 4) дисбаланс классов;
- 5) разные яркостные характеристики снимков;
- 6) некорректная разметка снимков;
- 7) случаи, в которых объекты целевого класса занимают весь снимок (или его часть, полученную в рамках предобработки данных) либо отсутствуют на нем.

3 Результаты

В вычислительном эксперименте по определению влияния состава выборок данных на результаты обучения, валидации и тестирования моделей нейронных сетей семантической сегментации геополей **использованы следующие нейронные сети:**

1. U-Net с основой ResNet-50 [11] — базовая нейронная сеть для семантической сегментации цифровых изображений в настоящее время, архитектура которой используется в качестве основы при решении задач распознавания различных классов земной поверхности;
2. STT (англ. Sparse Token Transformer) [11] — архитектура на основе трансформера, являющаяся одним из вариантов реализации механизмов внимания;
3. MF-CNN (англ. Multiscale Features-Convolutional Neural Network) [11] — сверточная нейронная сеть, являющаяся одной из реализаций обработки глобальных признаков (в данном случае через мультимасштабный модуль в архитектуре сети).

Рассмотрение результатов для нескольких нейронных сетей обусловлено тем, что механизмы внимания и обработка глобальных признаков являются основными используемыми в настоящее время концептуальными решениями, на которые состав выборок данных может оказывать разное влияние с точки зрения как обучаемости, так и точности распознавания.

В рамках вычислительного эксперимента рассматривается четыре класса объектов, а именно: «Здания», «Дорожная сеть», «Водные объекты» и «Лес». Следует отметить, что при проводимой семантической сегментации берется случай, когда должны быть распознаны и отнесены к соответствующему классу в рамках всякой решаемой задачи семантической сегментации классов земной поверхности каждое здание (вне зависимости от его типа), каждый водный объект (река, болото, море, озеро) и т. д. Это связано с намерением рассмотреть один из сложных случаев. Выбор определенного типа зданий, конкретного водного объекта или дорожной сети с заранее указанным покрытием существенно упрощает задачу (из-за наличия типичных яркостных характеристик, меньшей схожести с другими объектами и т. д.) и снимает часть рассматриваемых проблем. Подобное упрощение, однако, может приводить к отсутствию робастности при обработке данных нейронной сетью, например в случае работы со снимками разных территорий или изображениями, полученными с помощью разной съемочной аппаратуры, снимками с нечеткими границами или разными яркостными характеристиками объектов целевого класса.

Для выполнения эксперимента использованы три общедоступные выборки данных:

- WHU Buildings Dataset [1, 5, 6] (также известная как WHU-BE) — для определения влияния состава выборок данных на результаты обучения, валидации и тестирования нейронных сетей в задаче выделения объектов класса «Здания»;
- DeepGlobe Land Cover Classification Dataset [40] — для определения влияния состава выборок данных на обучаемость и точность распознавания моделей в задаче выделения объектов класса «Водные объекты» и распознавания объектов класса «Лес»;
- DeepGlobe Road Extraction Dataset [7, 33] — для определения влияния состава выборок данных на результаты, получаемые в задаче выделения объектов класса «Дорожная сеть».

В состав данных выборок входят снимки:

- с разными яркостными характеристиками объектов целевых классов;
- эквивалентностью яркостных характеристик объектов интереса и других объектов сцен;
- тенями;
- разными яркостными характеристиками (данные случаи в выборке WHU-BE отсутствуют, но присутствуют в выборках DeepGlobe);
- дисбалансом классов.

Таким образом, используемые общедоступные выборки данных позволяют рассмотреть большую часть проблематики в составе выборок данных для задач распознавания различных классов объектов на аэрокосмических изображениях. Характеристики указанных выборок представлены в **табл. 2**.

Таблица 2 Метаданные исходных выборок данных для выполнения вычислительного эксперимента

Table 2 Metadata of original datasets for semantic segmentation experiment

Выборка данных	Год	ПР, м	РР, бит/пикс.	Разрешение снимка	Тип снимка	Кол-во снимков в подвыборке		
						обуч.	валид.	тест.
WHU Building Dataset	2022	1	8	512×512	аэрофото	4736	1036	2416
DeepGlobe Road Extraction Dataset	2018	0,5	24	1024×1024	космический	6226	1243*	1101*
DeepGlobe Land Cover Classification Dataset				2448×2448		803	172*	171*

Примечание. Знаком * обозначено отсутствие разметки для данных.

На основе используемых общедоступных выборок данных составлены при пре-добработке, включающей деление на части, выборки с сохранением граничных случаев дисбаланса классов и выборки, в которых данные случаи сведены к минимуму (5–10 % от данных в выборке). В рамках выделения дорожной сети рассматривается влияние дисбаланса классов при разном разрешении данных в выборках на получаемые результаты обучения и распознавания, т.к. в случае дорожной сети дисбаланс классов в научных работах отмечается отдельно [37–39]. Кроме того, аугментация данных для уравнивания их количества в случае сохранения и удаления граничных случаев дисбаланса классов при предобработке не проводилась, поскольку она может исказить (повысить относительно сохранения граничных случаев) результаты обучения, валидации и тестирования на выборках при удалении граничных случаев дисбаланса классов. Поэтому было решено искусственно не повышать количество данных (несмотря на его снижение) после соответствующего отбора. Характеристики составленных выборок представлены в **табл. 3**.

Таблица 3 Составленные для вычислительного эксперимента выборки данных с сохранением и после удаления большей части граничных случаев дисбаланса классов (сохранение 5–10 % случаев в подвыборках)

Table 3 Datasets compiled for semantic segmentation experiment with borderline cases of class imbalance preservation and with part of class imbalance borderline cases deletion (preservation of 5–10% borderline cases in subsets)

Класс объектов	Исходная выборка	Используемое разрешение	Количество каналов	Количество снимков в подвыборке		
				обучающей	валидационной	тестовой
С сохранением граничных случаев дисбаланса классов						
Здания	WHU Building Dataset	512×512	3	4736	1036	2416
Водные объекты	DeepGlobe Land Cover Classification Dataset			4300	1375	750
Лес				3175	400	400
Дорожная сеть	DeepGlobe Road Extraction Dataset	1024×1024		4800	964	462
После удаления большей части граничных случаев дисбаланса классов						
Водные объекты	DeepGlobe Land Cover Classification Dataset	512×512	3	1980	463	175
Лес				1918	106	158
Дорожная сеть				DeepGlobe Road Extraction Dataset	2938	640

Обучение, валидация и тестирование моделей нейронных сетей проведены с использованием видеокарты NVIDIA GeForce RTX 3080 Ti. Параметры моделей нейронных сетей, используемые при работе с выборками аэрокосмических изображений и масок с сохранением граничных случаев и выборками, полученными в результате удаления последних, представлены в **табл. 4**.

Таблица 4 Параметры моделей для обучения, валидации и тестирования на выборках с сохранением и после удаления граничных случаев дисбаланса классов

Table 4 Model parameters for training, validation and testing on datasets with borderline cases of class imbalance preservation and when class imbalance borderline cases are deleted

Класс объектов	Количество эпох обучения	Скорость обучения	Количество поступающих образцов
С сохранением граничных случаев дисбаланса классов			
Здания	20	0,001	4
Водные объекты	5		4
Лес	20		4
Дорожная сеть	5	0,01	2
После удаления граничных случаев дисбаланса классов			
Водные объекты	20	0,001	4
Лес	20		4
Дорожная сеть	25		4

Результаты обучения моделей в задачах распознавания объектов классов «Дорожная сеть», «Здания», «Водные объекты» и «Лес» на указанных выборках данных представлены в **табл. 5–10**.

Таблица 5 Результат обучения моделей для выделения дорожной сети на подвыборке с разрешением снимков 1024×1024 (среднее значение бинарной кросс-энтропии)

Table 5 Model training result for roads extraction on subset with images resolution 1024×1024 (mean binary cross-entropy value)

Модель	Эпоха			Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	3-я	5-я	
U-Net (ResNet50)	0,0581	0,0358	0,0324	03:25:42
STT	0,0769	0,0464	0,0372	03:52:45
MF-CNN	0,0593	0,0412	0,0374	01:50:12

Таблица 6 Результат обучения моделей для выделения дорожной сети на подвыборке с разрешением частей 512×512 после удаления граничных случаев дисбаланса классов (среднее значение бинарной кросс-энтропии)

Table 6 Model training result for roads extraction on subset with crops resolution 512×512 after class imbalance borderline cases deletion (mean binary cross-entropy value)

Модель	Эпоха				Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	5-я	15-я	25-я	
U-Net (ResNet50)	0,0955	0,0568	0,0406	0,0312	02:52:42
STT	0,0658	0,0457	0,0387	0,0348	02:59:07
MF-CNN	0,0854	0,0575	0,0467	0,0412	01:18:08

Таблица 7 Результат обучения моделей для распознавания объектов класса «Лес» с сохранением и после удаления граничных случаев дисбаланса классов (среднее значение бинарной кросс-энтропии)

Table 7 Model training result for “Forest” class recognition with borderline cases of class imbalance preservation and when class imbalance borderline cases are deleted (mean binary cross-entropy value)

Модель	Эпоха			Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	10-я	20-я	
С сохранением граничных случаев дисбаланса классов				
U-Net (ResNet50)	0,1271	0,0772	0,0667	02:23:24
STT	0,1359	0,0800	0,0760	02:26:10
MF-CNN	0,1152	0,0746	0,0683	01:20:10
После удаления граничных случаев дисбаланса классов				
U-Net (ResNet50)	0,1248	0,0852	0,0767	01:28:38
STT	0,1215	0,0996	0,0784	01:20:07
MF-CNN	0,1153	0,0869	0,0820	00:51:24

Таблица 8 Результат обучения моделей для выделения объектов класса «Водные объекты» при сохранении граничных случаев дисбаланса классов (среднее значение бинарной кросс-энтропии)

Table 8 Model training result for “Water bodies” class extraction with borderline cases of class imbalance preservation (mean binary cross-entropy value)

Модель	Эпоха			Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	3-я	5-я	
U-Net (ResNet50)	0,0554	0,0335	0,0234	00:52:26
STT	0,0674	0,0425	0,0718	00:58:18
MF-CNN	0,0506	0,0315	0,0255	00:42:26

Таблица 9 Результат обучения моделей для выделения водных объектов после удаления граничных случаев дисбаланса классов (среднее значение бинарной кросс-энтропии)

Table 9 Model training result for water bodies extraction with class imbalance borderline cases deletion (mean binary cross-entropy value)

Модель	Эпоха				Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	5-я	10-я	20-я	
U-Net (ResNet50)	0,0913	0,0487	0,0362	0,0264	01:37:37
STT	0,1254	0,0882	0,0524	0,0316	01:40:09
MF-CNN	0,0837	0,0496	0,0386	0,0268	00:56:04

Таблица 10 Результат обучения моделей для выделения зданий (среднее значение бинарной кросс-энтропии)

Table 10 Model training result for building extraction (mean binary cross-entropy value)

Модель	Эпоха					Затраченное время (с учетом проведения валидации), чч:мм:сс
	1-я	5-я	10-я	15-я	20-я	
U-Net (ResNet50)	0,0593	0,0228	0,0181	0,0153	0,0134	03:35:08
STT	0,1685	0,0758	0,0656	0,0647	0,0595	03:46:46
MF-CNN	0,0478	0,0251	0,0208	0,0185	0,0171	01:59:23

Результаты валидации для соответствующих эпох обучения представлены в табл. 11–16.

Таблица 11 Результаты валидации для выделения класса «Дорожная сеть» на подвыборке с разрешением снимков 1024×1024 (метрика IoU)

Table 11 Validation results for “Roads” class extraction on subset with images resolution 1024×1024 (IoU metric)

Модель	Эпоха		
	1-я	3-я	5-я
U-Net (ResNet50)	0,4260	0,3850	0,5145
STT	0,5139	0,5521	0,5642
MF-CNN	0,2939	0,5526	0,4470

Таблица 12 Результаты валидации для выделения класса «Дорожная сеть» на подвыборке с разрешением частей 512×512 после удаления граничных случаев дисбаланса классов (метрика IoU)

Table 12 Validation results for “Roads” class extraction on subset with crops resolution 512×512 after class imbalance borderline cases deletion (IoU metric)

Модель	Эпоха			
	1-я	5-я	15-я	25-я
U-Net (ResNet50)	0,0181	0,4349	0,5359	0,5622
STT	0,4983	0,5521	0,5973	0,6416
MF-CNN	0,2490	0,3963	0,4883	0,5311

Таблица 13 Результаты валидации для распознавания класса «Лес» при сохранении и после удаления граничных случаев дисбаланса классов (метрика IoU)

Table 13 Validation results for “Forest” class recognition with borderline cases of class imbalance preservation and when class imbalance borderline cases are deleted (IoU metric)

Модель	Эпоха		
	1-я	10-я	20-я
С сохранением граничных случаев дисбаланса классов			
U-Net (ResNet50)	0,4097	0,4435	0,5112
STT	0,4282	0,5159	0,5688
MF-CNN	0,4199	0,5323	0,4603
После удаления граничных случаев дисбаланса классов			
U-Net (ResNet50)	0,4667	0,5767	0,6370
STT	0,6211	0,6377	0,6734
MF-CNN	0,5792	0,6366	0,6069

Таблица 14 Результаты валидации для выделения класса «Водные объекты» при сохранении граничных случаев дисбаланса классов (метрика IoU)

Table 14 Validation results for “Water bodies” class extraction with borderline cases of class imbalance preservation (IoU metric)

Модель	Эпоха		
	1-я	3-я	5-я
U-Net (ResNet50)	0,4671	0,5881	0,6185
STT	0	0	0
MF-CNN	0,4629	0,6206	0,6694

Таблица 15 Результаты валидации для выделения класса «Водные объекты» после удаления граничных случаев дисбаланса классов (метрика IoU)

Table 15 Validation results for “Water bodies” class extraction after class imbalance borderline cases deletion (IoU metric)

Модель	Эпоха			
	1-я	5-я	10-я	20-я
U-Net (ResNet50)	0,6027	0,7412	0,7722	0,7976
STT	0,0627	0,3534	0,6331	0,7825
MF-CNN	0,6286	0,6897	0,7704	0,7977

Таблица 16 Результаты валидации для выделения класса «Здания» (метрика IoU)

Table 16 Validation results for “Buildings” class extraction (IoU metric)

Модель	Эпоха				
	1-я	5-я	10-я	15-я	20-я
U-Net (ResNet50)	0,6521	0,7895	0,8366	0,8470	0,8602
STT	0,8484	0,8763	0,8852	0,9006	0,9045
MF-CNN	0,5651	0,8104	0,7973	0,7851	0,8537

Количественные результаты тестирования на выборках с сохранением граничных случаев дисбаланса классов и при их удалении представлены в **табл. 17–20**, где $F1$ — F -мера, OA (англ. Overall Accuracy) — попиксельная точность.

Таблица 17 Результаты тестирования для класса «Дорожная сеть» на подвыборке с разрешением снимков 1024×1024 и подвыборке с разрешением снимков 512×512 после удаления граничных случаев дисбаланса классов

Table 17 Testing results for class “Roads” on subset with images resolution 1024×1024 and subset with images resolution 512×512 after class imbalance borderline cases deletion

Модель	Количество эпох обучения	VCE (об.)	IoU (вал.)	IoU	F1	OA
На подвыборке с разрешением 1024×1024						
U-Net (ResNet50)	5	0,0324	0,5145	0,5330	0,6954	0,9770
STT	5	0,0372	0,5642	0,5847	0,7380	0,9812
MF-CNN	5	0,0375	0,4470	0,6200	0,8740	0,9734
На подвыборке с разрешением 512×512 после удаления граничных случаев дисбаланса классов						
U-Net (ResNet50)	22	0,0319	0,5702	0,5743	0,7296	0,9517
STT	25	0,0348	0,6416	0,6496	0,7876	0,9622
MF-CNN	24	0,0414	0,5301	0,5389	0,7004	0,9494

Таблица 18 Результаты тестирования для класса «Лес» при сохранении и после удаления граничных случаев дисбаланса классов

Table 18 Testing results for “Forest” class with borderline cases of class imbalance preservation and when class imbalance borderline cases are deleted

Модель	Количество эпох обучения	ВСЕ (об.)	IoU (вал.)	IoU	F1	OA
С сохранением граничных случаев дисбаланса классов						
U-Net (ResNet50)	16	0,0701	0,5944	0,7194	0,8368	0,8752
STT	17	0,0738	0,5911	0,6770	0,8074	0,8578
MF-CNN	13	0,0711	0,5689	0,6961	0,8208	0,8695
После удаления граничных случаев дисбаланса классов						
U-Net (ResNet50)	18	0,0791	0,6652	0,7851	0,8796	0,8816
STT	18	0,0808	0,6859	0,7973	0,8872	0,8853
MF-CNN	19	0,0806	0,6724	0,7973	0,8872	0,8840

Таблица 19 Результаты тестирования для класса «Водные объекты» при сохранении и после удаления граничных случаев дисбаланса классов

Table 19 Testing results for “Water bodies” class with borderline cases of class imbalance preservation and when class imbalance borderline cases are deleted

Модель	Количество эпох обучения	ВСЕ (об.)	IoU (вал.)	IoU	F1	OA
С сохранением граничных случаев дисбаланса классов						
U-Net (ResNet50)	5	0,0233	0,6185	0,6913	0,8175	0,9745
STT	3	0,0425	0	0	0	0,9373
MF-CNN	5	0,0255	0,6694	0,7762	0,8740	0,9855
После удаления граничных случаев дисбаланса классов						
U-Net (ResNet50)	15	0,0299	0,8054	0,8932	0,9436	0,9707
STT	19	0,0459	0,7906	0,8961	0,9452	0,9712
MF-CNN	18	0,0272	0,8063	0,8832	0,9380	0,9669

Таблица 20 Результаты тестирования для класса «Здания»

Table 20 Testing results for “Buildings” class

Модель	Количество эпох обучения	ВСЕ (об.)	IoU (вал.)	IoU	F1	OA
U-Net (ResNet50)	20	0,0134	0,8602	0,8642	0,9271	0,9840
STT	20	0,0595	0,9045	0,8808	0,9366	0,9860
MF-CNN	20	0,0171	0,8537	0,8496	0,9187	0,9821

4 Обсуждение

По полученным результатам обучения можно заключить следующее:

1. Удаление граничных случаев дисбаланса классов оказывает положительное влияние на обучаемость (уменьшение значений бинарной кросс-энтропии) моделей нейронных сетей на выборках аэрокосмических изображений и масок в рамках каждой из рассматриваемых задач и позволяет с большей достоверностью (применительно к сопоставлению с результатами валидации) отразить ход обучения.
2. Снимки с разными яркостными характеристиками в составе выборки для задачи выделения дорожной сети, распознавания объектов класса «Лес» и выделения водных объектов не оказали негативного влияния на возможность обучения моделей нейронных сетей для решения указанных задач.
3. Удаление граничных случаев дисбаланса классов (а значит, уменьшение размеров выборки) при решении каждой из рассматриваемых задач привело к снижению временных затрат на обучение и валидацию использованных моделей нейронных сетей с повышением эффективности проводимого обучения. Это выразилось в уменьшении значений бинарной кросс-энтропии в сравнении с обучением моделей на выборках с сохранением всех граничных случаев дисбаланса классов.
4. Снимки, на которых имеются разные или эквивалентные другим объектам яркостные характеристики объектов целевого класса, в контексте решения задач выделения всех рассматриваемых классов объектов не привели к снижению обучаемости моделей на использованных выборках.

По результатам валидации в рамках проведенного вычислительного эксперимента сделаны следующие выводы:

1. При удалении большей части граничных случаев дисбаланса классов в подвыборке для обучения точность распознавания объектов целевого класса используемыми нейронными сетями стала выше, что указывает на положительный эффект от этого удаления как части предобработки данных.
2. Как удаление, так и сохранение снимков, на которых объекты целевого класса отсутствуют или занимают всю сцену, не влияет на оценку точности распознавания объектов целевых классов моделями нейронных сетей по метрике IoU.
3. Включенные в валидационную подвыборку снимки, отражающие разные возможные случаи в контексте генеральных совокупностей для рассматриваемых задач, позволяют учитывать при валидации робастность моделей к данным, отражающим эти случаи, что, в свою очередь, положительно сказывается на дальнейшем отборе моделей для тестирования.

4. Сохранение граничных случаев дисбаланса классов может приводить к низкой (вплоть до полного отсутствия) обучаемости моделей нейронных сетей (**табл. 14**).

Исходя из результатов тестирования, можно заключить следующее:

1. Попиксельная точность (OA) как метрика оценки точности распознавания во многих случаях может некорректно отражать точность семантической сегментации целевого класса (**табл. 19**) моделями нейронных сетей на аэрокосмических изображениях.
2. Использование снимков с разрешением, при котором дисбаланс классов выше, чем при использовании частей снимков, приводит к снижению точности распознавания и ухудшает результаты семантической сегментации объектов целевого класса.
3. Обучение, валидация и тестирование моделей с использованием снимков с разными яркостными характеристиками, эквивалентностью яркостных характеристик объектов целевых классов и снимков с разными яркостными характеристиками зданий, дорожной сети, растительности и водных объектов не оказывают негативного влияния на обучаемость и точность распознавания моделей, а позволяют обеспечить более полное представление генеральной совокупности в контексте задач семантической сегментации объектов на аэрокосмических изображениях с помощью нейронных сетей.
4. Во всех решаемых в вычислительном эксперименте задачах удаление граничных случаев дисбаланса классов, где оно проводилось, привело к повышению точности распознавания при тестировании использованных моделей нейронных сетей.

Результаты вычислительного эксперимента по семантической сегментации геополей посредством нейронных сетей свидетельствуют о следующем:

1. Отбор данных на основе удаления граничных случаев дисбаланса классов в рамках предобработки выборок аэрокосмических изображений и масок и на начальном этапе, и после деления снимков на части оказывает положительное влияние на этапы работы с моделями нейронных сетей при семантической сегментации геополей на разных этапах глубокого обучения (обучение, валидация и тестирование).
2. Использование аэрокосмических снимков, отражающих проблемы представления исходного множества геополей, в составе выборок данных при обучении, валидации и тестировании позволяет учесть подобные случаи, лучше отразить генеральную совокупность и, таким образом,

повысить робастность моделей нейронных сетей для семантической сегментации геополей к поступающим в них данным.

3. Применение снимков с разрешением, при котором дисбаланс классов выше, чем при использовании частей снимков, оказывает негативное влияние на обучаемость и точность распознавания моделей семантической сегментации геополей.

5 Выводы

Сформулирована общая задача семантической сегментации геополей. Определены проблемы состава выборок данных и его влияние на обучение и точность нейронных сетей в контексте выделения водных объектов, зданий, дорожной сети, растительности на аэрокосмических изображениях высокого и сверхвысокого пространственного разрешения.

Проблемы состава выборок данных для семантической сегментации геополей связаны главным образом с представлением исходного множества геополей. К таким проблемам относятся: наличие на снимках теней, скрывающих объекты целевого класса; разные или эквивалентные другим объектам сцен яркостные характеристики; дисбаланс классов; разные яркостные характеристики снимков; некорректная разметка снимков; случаи, при которых объекты целевого класса занимают весь снимок (либо его часть, полученную в рамках предобработки данных) или полностью отсутствуют на нем.

Экспериментальные результаты указывают на то, что при составлении выборок данных снимки с разными яркостными характеристиками и снимки, на которых объекты целевого класса имеют разные или эквивалентные другим объектам, запечатленным на цифровых изображениях, яркостные характеристики, оказывают положительное влияние на учет указанных причин возникновения ложноположительных и ложноотрицательных срабатываний, что впоследствии приводит к снижению числа ложных срабатываний.

На основе анализа результатов вычислительного эксперимента выявлено, что включение граничных случаев дисбаланса классов в подвыборку для обучения снижает качество обучения моделей нейронных сетей и точность семантической сегментации геополей. Кроме того, использование снимков с разрешением, при котором дисбаланс классов выше, чем при использовании частей снимков, приводит к снижению обучаемости и точности моделей нейронных сетей.

Обеспечение отбора данных при делении цифровых изображений и масок на части при предобработке данных позволяет снизить временные затраты на обучение моделей и повысить точность распознавания.

Таким образом, количество данных в выборках является приоритетным для решения задач глубокого обучения применительно к автоматизированному дешифрированию аэрокосмических изображений, однако качество данных (выраженное в наличии или отсутствии граничных случаев, в корректности разметки и включении снимков, отражающих разные случаи в рамках взятой для конкретной задачи генеральной совокупности) оказывает существенное влияние на возможность обучения и последующего использования моделей нейронных сетей.

БИБЛИОГРАФИЯ

1. Xu R., Mao R., Zhuang Z., et al. Building Extraction from Remote Sensing Images Based on Multi-Scale Attention Gate and Enhanced Positional Information // PeerJ Computer Science. 2025. Vol. 11. P. e2826. DOI:10.7717/peerj-cs.2826.
2. Chen Y., Xie Y., Yao W., et al. U-MGA: a Multi-Module Unet Optimized with Multi-Scale Global Attention Mechanisms for Fine-Grained Segmentation of Cultivated Areas // Remote Sensing. 2025. Vol. 17. Iss. 5. P. 760. DOI:10.3390/rs17050760.
3. Cai J., Tao L., Li Y. CM-UNet++: A Multi-Level Information Optimized Network for Urban Water Body Extraction from High-Resolution Remote Sensing Imagery // Remote Sensing. 2025. Vol. 17. Iss. 6. P. 980. DOI:10.3390/rs17060980.
4. Gui L., Gu X., Huang F., et al. Road Extraction from Remote Sensing Images Using a Skip-Connected Parallel CNN-Transformer Encoder-Decoder Model // Applied Sciences. 2025. Vol. 15. Iss. 3. P. 1427. DOI:10.3390/app15031427.
5. Liu Y., Duan Y., Zhang X., et al. FEPA-Net: A Building Extraction Network Based on Fusing the Feature Extraction and Position Attention Module // Applied Sciences. 2025. Vol. 15. Iss. 8. P. 4432. DOI:10.3390/app15084432.
6. Zhu B., Yu D., Xiao X., et al. AP-Pointrend: An Improved Network for Building Extraction via High-Resolution Remote Sensing Images // Remote Sensing. 2025. Vol. 17. Iss. 9. P. 1481. DOI:10.3390/rs17091481.
7. Li X., Yang S., Meng F., et al. LCMorph: Exploiting Frequency Cues and Morphological Perception for Low-Contrast Road Extraction in Remote Sensing Images // Remote Sensing. 2025. Vol. 17. Iss. 2. P. 257. DOI:10.3390/rs17020257.
8. Weng Z., Li Q., Zheng Z., et al. SCR-Net: A Dual-Channel Water Body Extraction Model Based on Multi-Spectral Remote Sensing Imagery – A Case Study of Daihai Lake, China // Sensors. 2025. Vol. 25. Iss. 3. P. 763. DOI: 10.3390/s25030763.
9. Wu Q., Chen M., Shi H., et al. Algorithm for Detecting Trees Affected by Pine Wilt Disease in Complex Scenes Based on CNN-Transformer // Forests. 2025. Vol. 16. Iss. 4. P. 596. DOI:10.3390/f16040596.
10. Wang L., Gao Y., Liu Y., et al. Monitoring Pine Shoot Beetle Damage Using UAV Imagery and Deep Learning Semantic Segmentation Under Different Forest Backgrounds // Forests. 2025. Vol. 16. Iss. 4. P. 668. DOI:10.3390/f16040668.

11. Бирюков Н.А., Майоров А.А., Лапчинская М.П. Семантическая сегментация геополей с использованием нейронных сетей на примере проблематики выделения зданий на космо- и аэрофотоснимках // Известия вузов «Геодезия и аэрофотосъемка». 2024. Т. 68, № 1. С. 44–61. DOI:10.30533/GiA-2024-004.
12. Li J., He W., Cao W., et al. UANet: An Uncertainty-Aware Network for Building Extraction from Remote Sensing Images // IEEE Transactions on Geoscience and Remote Sensing. 2024. Vol. 62. P. 5608513. DOI:10.1109/TGRS.2024.3361211.
13. Han R., Fan X., Liu J. EUNet: Edge-Unet for Accurate Building Extraction and Edge Emphasis in Gaofen-7 Images // Remote Sensing. 2024. Vol. 16. Iss. 13. P. 2397. DOI:10.3390/rs16132397.
14. Fenglei W., Xin G., Zongze Z., et al. A Boundary-Enhanced Semantic Segmentation Model for Buildings // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2025. Vol. 18. P. 5733–5748. DOI:10.1109/JSTARS.2025.3529456.
15. Chen P., Huang H., Ye F., et al. A Benchmark Gaofen-7 Dataset for Building Extraction from Satellite Images // Scientific Data. 2024. Vol. 11. P. 187. DOI:10.1038/s41597-024-03009-5.
16. Li S., Bao T., Liu H., et al. Multilevel Feature Aggregated Network with Instance Contrastive Learning Constraint for Building Extraction // Remote Sensing. 2023. Vol. 15. Iss. 10. P. 2585. DOI:10.3390/rs15102585.
17. Lyu X., Jiang W., Li X., et al. MSAFNet: Multiscale Successive Attention Fusion Network for Water Body Extraction of Remote Sensing Images // Remote Sensing. 2023. Vol. 15. Iss. 12. P. 3121. DOI:10.3390/rs15123121.
18. Li M., Hong L., Guo J., et al. Automated Extraction of Lake Water Bodies in Complex Geographical Environments by Fusing Sentinel-1/2 Data // Water. 2022. Vol. 14. Iss. 1. P. 30. DOI:10.3390/w14010030.
19. Wang Y., Li S., Lin Y., et al. Lightweight Deep Neural Network Method for Water Body Extraction from High-Resolution Remote Sensing Images with Multisensors // Sensors. 2021. Vol. 21. Iss. 21. 7397. DOI:10.3390/s21217397.
20. Wang B., Chen Z., Wu L., et al. SADA-Net: A Shape Feature Optimization and Multiscale Context Information-Based Water Body Extraction Method for High-Resolution Remote Sensing Images // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2022. Vol. 15. P. 1744–1759. DOI:10.1109/JSTARS.2022.3146275.
21. Weng Y., Li Z., Tang G., et al. OCNet-Based Water Body Extraction from Remote Sensing Images // Water. 2023. Vol. 15. Iss. 20. P. 3557. DOI:10.3390/w15203557.
22. Yu J., Cai Y., Lyu X., et al. Boundary-Guided Semantic Context Network for Water Body Extraction from Remote Sensing Images // Remote Sensing. 2023. Vol. 15. Iss. 17. P. 4325. DOI:10.3390/rs15174325.
23. Li M., Wu P., Wang B., et al. A Deep Learning Method of Water Body Extraction from High Resolution Remote Sensing Images with Multisensors // IEEE Journal

- of Selected Topics in Applied Earth Observations and Remote Sensing. 2021. Vol. 14. P. 3120–3132. DOI:10.1109/JSTARS.2021.3060769.
24. Zhao Z., Yang J., Wang M., et al. The PCA-NDWI Urban Water Extraction Model Based on Hyperspectral Remote Sensing // *Water*. 2024. Vol. 16. Iss. 7. P. 963. DOI:10.3390/w16070963.
25. Sandum H.N., Ørka H.O., Tomic O., et al. Semantic Segmentation of Forest Stands Using Deep Learning // Preprint arXiv.org, 2025. [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/2504.02471> (дата обращения: 09.06.2025).
26. Lin N., Quan H., He J., et al. Urban Vegetation Extraction from High-Resolution Remote Sensing Imagery on SD-Unet and Vegetation Spectral Features // *Remote Sensing*. 2023. Vol. 15. Iss. 18. P. 4488. DOI:10.3390/rs15184488.
27. Chen P., Li X., Peng Y., et al. WSSGCN: Hyperspectral Forest Image Classification via Watershed Superpixel Segmentation and Sparse Graph Convolutional Networks // *Forests*. 2025. Vol. 16. Iss. 5. P. 827. DOI:10.3390/f16050827.
28. Chen S., Zhang M., Lei F. Mapping Vegetation Types by Different Fully Convolutional Neural Network Structures with Inadequate Training Labels in Complex Landscape Urban Areas // *Forests*. 2023. Vol. 14. Iss. 9. P. 1768. DOI:10.3390/f14091788.
29. Li Y., Min S., Song B., et al. Multisource High-Resolution Remote Sensing Image Vegetation Extraction with Comprehensive Multifeature Perception // *Remote Sensing*. 2024. Vol. 16. Iss. 4. P. 712. DOI:10.3390/rs16040712.
30. Wang B., Yao Y. Mountain Vegetation Classification Method Based on Multi-Channel Semantic Segmentation Model // *Remote Sensing*. 2024. Vol. 16. Iss. 2. P. 256. DOI:10.3390/rs16020256.
31. Tao J., Chen Z., Sun Z., et al. Seg-Road: A Segmentation Network for Road Extraction Based on Transformer and CNN with Connectivity Structures // *Remote Sensing*. 2023. Vol. 15. Iss. 6. P. 1602. DOI:10.3390/rs15061602.
32. Lin S., Yao X., Liu X., et al. MS-AGAN: Road Extraction via Multi-Scale Information Fusion and Asymmetric Generative Adversarial Networks from High-Resolution Remote Sensing Images under Complex Backgrounds // *Remote Sensing*. 2023. Vol. 15. Iss. 13. P. 3367. DOI:10.3390/rs15133367.
33. Zhong B., Dan H., Liu M., et al. FERDNet: High-Resolution Remote Sensing Road Extraction Network Based on Feature Enhancement of Road Directionality // *Remote Sensing*. 2025. Vol. 17. Iss. 3. P. 376. DOI:10.3390/rs17030376.
34. Chen J., Yang L., Wang H., et al. Road Extraction from High-Resolution Remote Sensing Images via Local and Global Context Reasoning // *Remote Sensing*. 2023. Vol. 5. Iss. 17. P. 4177. DOI:10.3390/rs15174177.
35. Mahara A., Khan M.R.K., Deng L., et al. Automated Road Extraction from Satellite Imagery Integrating Dense Depthwise Dilated Separable Spatial Pyramid Pooling with DeepLabV3+ // *Applied Sciences*. 2025. Vol. 15. Iss. 3. P. 1027. DOI:10.3390/app15031027.

36. Li B., Tang X., Xiao R., et al. Dual Convolutional Network Based on Hypergraph and Multilevel Feature Fusion for Road Extraction from High-Resolution Remote Sensing Images // International Journal of Digital Earth. 2024. Vol. 17. No. 1. P. 2303354. DOI:10.1080/17538947.2024.2303354.
37. Zhao S., Feng Z., Chen L., et al. DANet: A Semantic Segmentation Network for Remote Sensing of Roads Based on Dual-ASPP Structure // Electronics. 2023. Vol. 12. Iss. 15. P. 3243. DOI:10.3390/electronics12153243.
38. Zhao L., Zhang J., Meng X., et al. Road Extraction Method of Remote Sensing Image Based on Deformable Attention Transformer // Symmetry. 2024. Vol. 16. Iss. 4. P. 468. DOI:10.3390/sym16040468.
39. Zhang Y., Zhang L., Wang Y., et al. AGF-Net: Adaptive Global Feature Fusion Network for Road Extraction from Remote-Sensing Images // Complex & Intelligent Systems. 2024. Vol. 10. P. 4311–4328. DOI:10.1007/s40747-024-01364-9.
40. Ma D., Jiang L., Li J., et al. Water Index and Swin Transformer (WISTE) for Water Body Extraction from Multispectral Remote Sensing Images // GIScience & Remote Sensing. 2023. Vol. 60. No. 1. P. 2251704. DOI:10.1080/15481603.2023.2251704.

АВТОРЫ

Бирюков Никита Андреевич

ФГБОУ ВО «Московский государственный университет геодезии и картографии»
(МИИГАиК), Москва, Россия
кафедра информационно-измерительных систем, факультет геоинформатики
и информационной безопасности

 0000-0001-9856-4696

Поступила 14.06.2025. Принята к публикации 22.08.2025. Опубликовано 29.08.2025.

UDC 528.8:004.62

DOI:10.30533/scidata-2025-16-12



The Impact of High and Super Spatial Resolution Remote Sensing Images Datasets Composition on Training and Accuracy of Geofields Semantic Segmentation Neural Networks on Example of a Different Earth's Surface Classes Recognition

Nikita A. Biryukov¹

AFFILIATIONS

¹ Moscow State University of Geodesy and Cartography, Moscow, Russia

4096414@gmail.com

CITATION

Biryukov NA. The Impact of High and Super Spatial Resolution Remote Sensing Images Datasets Composition on Training and Accuracy of Geofields Semantic Segmentation Neural Networks on Example of a Different Earth's Surface Classes Recognition. *Spatial Data: Science, Research and Technology*. 2025;16(2): 30–57. DOI:10.30533/scidata-2025-16-12.

KEYWORDS

geofield, geofields set, semantic segmentation, dataset composition, class imbalance, neural network, recognition accuracy

ABSTRACT

The remote sensing images and masks datasets that are used for different earth's surface classes recognition tasks can significantly impact neural network models learnability and semantic segmentation results that are received after model training. Datasets composition problematic

usually are not investigated from these datasets point of view as it is viewed as neural network data processing problematic in general in each specific remote sensing data semantic segmentation task. In geofields semantic segmentation context general problematic for object semantic segmentation on aerial and satellite data that includes such main problems as images with different spectral characteristics, images with shadows, images with “different object, same spectrum”, images with incorrect annotation and images with class imbalance borderline cases is determined. Mentioned problems are considered as problems of original geofields set representation in datasets. As result of a different earth’s surface classes semantic segmentation experiment with U-Net, STT and MF-CNN it was determined that class imbalance borderline cases and using images with resolution in which class imbalance is higher than using their crops reduce learning ability and recognition accuracy of neural networks and class imbalance borderline cases deletion based data selection in data preprocessing process leads to accuracy increase and models training time decrease.

REFERENCES

1. Xu R, Mao R, Zhuang F, et al. Building Extraction from Remote Sensing Images Based on Multi-Scale Attention Gate and Enhanced Positional Information. *PeerJ Computer Science*. 2025;11: e2826. DOI:10.7717/peerj-cs.2826.
2. Chen Y, Xie Y, Yao W, et al. U-MGA: a Multi-Module Unet Optimized with Multi-Scale Global Attention Mechanisms for Fine-Grained Segmentation of Cultivated Areas. *Remote Sensing*. 2025;17(5): 760. DOI:10.3390/rs17050760.
3. Cai J, Tao L, Li Y. CM-UNet++: A Multi-Level Information Optimized Network for Urban Water Body Extraction from High-Resolution Remote Sensing Imagery. *Remote Sensing*. 2025;17(6): 980. DOI:10.3390/rs17060980.
4. Gui L, Gu X, Huang F, et al. Road Extraction from Remote Sensing Images Using a Skip-Connected Parallel CNN-Transformer Encoder-Decoder Model. *Applied Sciences*. 2025;15(3): 1427. DOI:10.3390/app15031427.
5. Liu Y, Duan Y, Zhang X, et al. FEPA-Net: A Building Extraction Network Based on Fusing the Feature Extraction and Position Attention Module. *Applied Sciences*. 2025;15(8): 4432. DOI:10.3390/app15084432.
6. Zhu B., Yu D., Xiao X., et al. AP-PointRend: An Improved Network for Building Extraction via High-Resolution Remote Sensing Images. *Remote Sensing*. 2025;17(9): 1481. DOI:10.3390/rs17091481.
7. Li X, Yang S, Meng F, et al. LCMorph: Exploiting Frequency Cues and Morphological Perception for Low-Contrast Road Extraction in Remote Sensing Images. *Remote Sensing*. 2025;17(2): 257. DOI:10.3390/rs17020257.
8. Weng Z, Li Q, Zheng Z, et al. SCR-Net: A Dual-Channel Water Body Extraction Model Based on Multi-Spectral Remote Sensing Imagery – A Case Study of Daihai Lake, China. *Sensors*. 2025;25(3): 763. DOI: 10.3390/s25030763.

9. Wu Q, Chen M, Shi H, et al. Algorithm for Detecting Trees Affected by Pine Wilt Disease in Complex Scenes Based on CNN-Transformer. *Forests*. 2025;16(4): 596. DOI:10.3390/f16040596.
10. Wang L, Gao Y, Liu Y, et al. Monitoring Pine Shoot Beetle Damage Using UAV Imagery and Deep Learning Semantic Segmentation under Different Forest Backgrounds. *Forests*. 2025;16(4): 668. DOI:10.3390/f16040668.
11. Biryukov NA, Maiorov AA, Lapchinskaya MP. Semanticheskaya segmentatsiya geopolei s ispol'zovaniem neironnykh setei na primere problematiki vydeleniya zdanii na kosmo- i aerofotosnimkakh [Geofields Semantic Segmentation by Neural Networks on Example of Building Extraction on Satellite and Aerial Images Problematic]. *Izvestia Vuzov. Geodesy and Aerophotosurveying*. 2024;68(1): 44–61. (In Russian). DOI:10.30533/GiA-2024-004.
12. Li J, He W, Cao W, et al. UANet: An Uncertainty-Aware Network for Building Extraction from Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*. 2024;62: 5608513. DOI:10.1109/TGRS.2024.3361211.
13. Han R, Fan X, Liu J. EUNet: Edge-Unet for Accurate Building Extraction and Edge Emphasis in Gaofen-7 Images. *Remote Sensing*. 2024;16(13): 2397. DOI:10.3390/rs16132397.
14. Fenglei W, Xin G, Zongze Z, et al. A Boundary-Enhanced Semantic Segmentation Model for Buildings. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2025;18: 5733–5748. DOI:10.1109/JSTARS.2025.3529456.
15. Chen P, Huang H, Ye F, et al. A Benchmark Gaofen-7 Dataset for Building Extraction from Satellite Images. *Scientific Data*. 2024;11: 187. DOI:10.1038/s41597-024-03009-5.
16. Li S, Bao T, Liu H, et al. Multilevel Feature Aggregated Network with Instance Contrastive Learning Constraint for Building Extraction. *Remote Sensing*. 2023;15(10): 2585. DOI:10.3390/rs15102585.
17. Lyu X, Jiang W, Li X, et al. MSAFNet: Multiscale Successive Attention Fusion Network for Water Body Extraction of Remote Sensing Images. *Remote Sensing*. 2023;15(12): 3121. DOI:10.3390/rs15123121.
18. Li M, Hong L, Guo J, et al. Automated Extraction of Lake Water Bodies in Complex Geographical Environments by Fusing Sentinel-1/2 Data. *Water*. 2022;14(1): 30. DOI:10.3390/w14010030.
19. Wang Y, Li S, Lin Y, et al. Lightweight Deep Neural Network Method for Water Body Extraction from High-Resolution Remote Sensing Images with Multisensors. *Sensors*. 2021;21(21): 7397. DOI:10.3390/s21217397.
20. Wang B, Chen Z, Wu L, et al. SADA-Net: A Shape Feature Optimization and Multiscale Context Information-Based Water Body Extraction Method for High-Resolution Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022;15: 1744–1759. DOI:10.1109/JSTARS.2022.3146275.

21. Weng Y, Li Z, Tang G, et al. OCNet-Based Water Body Extraction from Remote Sensing Images. *Water*. 2023;15(20): 3557. DOI:10.3390/w15203557.
22. Yu J, Cai Y, Lyu X, et al. Boundary-Guided Semantic Context Network for Water Body Extraction from Remote Sensing Images. *Remote Sensing*. 2023;15(17): 4325. DOI:10.3390/rs15174325.
23. Li M, Wu P, Wang B, et al. A Deep Learning Method of Water Body Extraction from High Resolution Remote Sensing Images with Multisensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2021;14: 3120–3132. DOI:10.1109/JSTARS.2021.3060769.
24. Zhao Z, Yang J, Wang M, et al. The PCA-NDWI Urban Water Extraction Model Based on Hyperspectral Remote Sensing. *Water*. 2024;16(7): 963. DOI:10.3390/w16070963.
25. Sandum HN, Ørka HO, Tomic O, et al. Semantic Segmentation of Forest Stands Using Deep Learning. *Preprint arXiv.org*, 2025. Available from: <https://arxiv.org/pdf/2504.02471> (Accessed 9 June 2025).
26. Lin N, Quan H, He J, et al. Urban Vegetation Extraction from High-Resolution Remote Sensing Imagery on SD-Unet and Vegetation Spectral Features. *Remote Sensing*. 2023;15(18): 4488. DOI:10.3390/rs15184488.
27. Chen P, Li X, Peng Y, et al. WSSGCN: Hyperspectral Forest Image Classification via Watershed Superpixel Segmentation and Sparse Graph Convolutional Networks. *Forests*. 2025;16(5): 827. DOI:10.3390/f16050827.
28. Chen S, Zhang M, Lei F. Mapping Vegetation Types by Different Fully Convolutional Neural Network Structures with Inadequate Training Labels in Complex Landscape Urban Areas. *Forests*. 2023;14(9): 1768. DOI:10.3390/f14091788.
29. Li Y, Min S, Song B, et al. Multisource High-Resolution Remote Sensing Image Vegetation Extraction with Comprehensive Multifeature Perception. *Remote Sensing*. 2024;16(4): 712. DOI:10.3390/rs16040712.
30. Wang B, Yao Y. Mountain Vegetation Classification Method Based on Multi-Channel Semantic Segmentation Model. *Remote Sensing*. 2024;16(2): 256. DOI:10.3390/rs16020256.
31. Tao J, Chen Z, Sun Z, et al. Seg-Road: A Segmentation Network for Road Extraction Based on Transformer and CNN with Connectivity Structures. *Remote Sensing*. 2023;15(6): 1602. DOI:10.3390/rs15061602.
32. Lin S, Yao X, Liu X., et al. MS-AGAN: Road Extraction via Multi-Scale Information Fusion and Asymmetric Generative Adversarial Networks from High-Resolution Remote Sensing Images under Complex Backgrounds. *Remote Sensing*. 2023;15(13): 3367. DOI:10.3390/rs15133367.
33. Zhong B, Dan H, Liu M, et al. FERDNet: High-Resolution Remote Sensing Road Extraction Network Based on Feature Enhancement of Road Directionality. *Remote Sensing*. 2025;17(3): 376. DOI:10.3390/rs17030376.

34. Chen J, Yang L, Wang H, et al. Road Extraction from High-Resolution Remote Sensing Images via Local and Global Context Reasoning. *Remote Sensing*. 2023;15(17): 4177. DOI:10.3390/rs15174177.
35. Mahara A, Khan MRK, Deng L, et al. Automated Road Extraction from Satellite Imagery Integrating Dense Depthwise Dilated Separable Spatial Pyramid Pooling with DeepLabV3+. *Applied Sciences*. 2025;15(3): 1027. DOI:10.3390/app15031027.
36. Li BW, Tang XH, Xiao R, et al. Dual Convolutional Network Based on Hypergraph and Multilevel Feature Fusion for Road Extraction from High-Resolution Remote Sensing Images. *International Journal of Digital Earth*. 2024;17(1): 2303354. DOI:10.1080/17538947.2024.2303354.
37. Zhao S, Feng Z, Chen L, et al. DANet: A Semantic Segmentation Network for Remote Sensing of Roads Based on Dual-ASPP Structure. *Electronics*. 2023;12(15): 3243. DOI:10.3390/electronics12153243.
38. Zhao L, Zhang J, Meng X, et al. Road Extraction Method of Remote Sensing Image Based on Deformable Attention Transformer. *Symmetry*. 2024;16(4): 468. DOI:10.3390/sym16040468.
39. Zhang Y, Zhang L, Wang Y, et al. AGF-Net: Adaptive Global Feature Fusion Network for Road Extraction from Remote-Sensing Images. *Complex & Intelligent Systems*. 2024;10: 4311–4328. DOI:10.1007/s40747-024-01364-9.
40. Ma D, Jiang L, Li J, et al. Water Index and Swin Transformer (WISTE) for Water Body Extraction from Multispectral Remote Sensing Images. *GIScience & Remote Sensing*. 2023;60(1): 2251704. DOI:10.1080/15481603.2023.2251704.

AUTHORS

Nikita A. Biryukov

Moscow State University of Geodesy and Cartography, Moscow, Russia

Department of Data Acquisition Systems, Faculty of Geoinformatics and Information Security

 0000-0001-9856-4696

Submitted: June 14, 2025. Accepted: August 22, 2025. Published: August 29, 2025.