



Семантическая сегментация геополей с использованием нейронных сетей на примере проблематики выделения зданий на космо- и аэрофотоснимках

Н.А. Бирюков¹✉, А.А. Майоров¹, М.П. Лапчинская¹

¹ Московский государственный университет геодезии и картографии,
Москва, Россия

✉ 4096414@gmail.com

ЦИТИРОВАНИЕ 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.

КЛЮЧЕВЫЕ СЛОВА геополь, семантическая сегментация, выделение зданий, нейронные сети, распознавание образов, машинное обучение, компьютерное зрение, предобработка данных

АННОТАЦИЯ Задача выделения зданий в контексте семантической сегментации геополей представляет собой задачу перехода от множества геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного излучения ко множеству геополей пространственного распределения собственного и отраженного объектами класса «Здания» электромагнитного излучения. Проблематика выделения зданий включает в себя следующие основные проблемы: ложноположительные и ложноотрицательные срабатывания при выделении зданий, выделение отдельных зданий как одного объекта и распознавание объектов класса «Здания» на снимках, имеющих разные яркостные характеристики. Данные проблемы, в рамках семантической сегментации геополей, рассматриваются с точки зрения снижения точности перехода от исходного множества геополей к результирующему множеству и относятся главным образом к исходному множеству. Посредством анализа указанных проблем определены основные концептуальные решения в архитектурах нейронных сетей, позволяющие уменьшить их влияние: механизмы внимания и обработка глобальных признаков. В результате эксперимента с архитектурами нейронных сетей STT, STEB-UNet, U-Net, MF-CNN и MSCFF, реализующими данные концептуальные решения, определены зависимости между временными затратами и конкретной реализацией концептуальных решений и наиболее эффективное с точки зрения обеспечиваемой точности решение — обработка глобальных признаков посредством мультимасштабного слияния признаков, реализуемое MSCFF.

1 Введение

Геополе — изменяющаяся в пространстве и во времени непрерывная переменная величина [1–3]. С математической точки зрения она описывается функцией $z = f(x, y, t)$, где z — значение геополя; x, y — пространственные координаты, а t — время¹ [2, 4].

Множество геополей, связанных предметной областью, может быть определено как кортеж вида $(z_1, z_2, \dots, z_n) = f(x, y, t)$, где z_n — числовое значение физической величины n -го геополя, $n \geq 1$; x, y — пространственные координаты, а t — время. Значение $n = 1$ отражает случай, когда берется одно геополе в рамках конкретной предметной области.

При этом временная координата t может быть опущена² [2], и тогда функция принимает вид $(z_1, z_2, \dots, z_n) = f(x, y)$.

В таком случае x, y будут представлять собой пространственные координаты для определенного фиксированного момента времени $t = \text{const}$.

В связи тем, что геополе есть изменяющаяся в пространстве и во времени физическая величина, как отмечалось выше, его корректно рассматривать как непрерывный сигнал³, задаваемый в общем случае одной из указанных ранее функций: $z = f(x, y, t)$ или $z = f(x, y)$, если опущена временная координата. Таким образом, геополе можно представить и в виде цифрового изображения, которое, как считают авторы данной статьи, необходимо рассматривать как результат дискретизации исходного непрерывного многомерного сигнала, соответствующего геополю, по текущей координате или времени и его квантования по уровню с постоянным шагом⁴.

Таким образом, с точки зрения работы с геополями семантическая сегментация представляет собой задачу перехода от одного множества геополей к другому и в указанном контексте в современном мире еще не рассматривалась.

Выделение зданий (*англ.* — «building extraction») как задача семантической сегментации геополей является сложной, но важной задачей, которая позволяет упростить проведение кадастровых работ [5], выявить незаконное использование территории или нелегальное строительство [5, 6]. Кроме того, выделение зданий применяется при городском планировании [7–9], оценке последствий от стихийных бедствий [8, 9], экологическом мониторинге [8].

Постоянное увеличение городской застройки [10–13] и наличие у зданий определенных особенностей, как с точки зрения их пространственного расположения и ориентирования на космо- и аэрофотоснимках [7, 9, 10, 14, 15], так и с точки зрения их яркостных характеристик [14–21], усложняет эффективное по точности выделение зданий, а в некоторых случаях приводит к невозможности его выполнения.

Целью работы является определение концептуальных решений в рамках архитектур нейронных сетей для выделения зданий на космо- и аэрофотоснимках как задачи семантической сегментации геополей.

¹ Жуковский О.И. Геоинформационные системы: учеб. пособие. Томск: Эль Контент, 2014. 130 с.

² Там же.

³ Журкин И.Г., Шавенько Н.К. Сигналы: учеб. пособие по курсу «Автоматизированная обработка аэрокосмической информации»: для студентов 4 и 5 курсов специальностей «Исследование природных ресурсов» и «Информационные системы в технике и технологиях». М.: МИИГАиК, 2007. 95 с.

⁴ Там же.

2 Материалы и методы

Задача выделения зданий на аэрокосмических снимках начала развиваться во второй половине XX века от формулирования общих принципов до сравнения разных подходов в научной периодике [10].

В настоящее время для выделения зданий используются выборки данных, состоящие из цифровых космо- и аэрофотоснимков высокого (1–10 м) [7, 22] и сверхвысокого (0,2–1 м) [22–27] пространственного разрешения и масок. В рамках выделения зданий как задачи семантической сегментации геополей выборка включает в себя:

- цифровые космо- и аэрофотоснимки, которые являются графическим представлением для некоторых участков земной поверхности множеств геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного излучения (ГОСТ Р 59079–2020, ГОСТ Р 59753–2021)⁵ [28] в определенных спектральных диапазонах длин волн;
- маски, которые представляют собой цифровые изображения, отражающие вид графического представления результирующего множества геополей, получаемого посредством семантической сегментации.

Результат распознавания модели нейронной сети при выделении зданий — цифровое изображение, являющееся графическим представлением множества геополей пространственного распределения собственного и отраженного объектами класса «Здания» электромагнитного излучения в определенных спектральных диапазонах длин волн.

Таким образом, задача выделения зданий с точки зрения семантической сегментации геополей представляет собой задачу перехода от множества геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного излучения ко множеству геополей пространственного распределения собственного и отраженного объектами класса «Здания» электромагнитного излучения.

Проблематика выделения зданий как задачи семантической сегментации геополей включает в себя проблемы, которые приводят к снижению точности перехода от исходного множества геополей к результирующему множеству. Представляется целесообразным определить концептуальные решения в архитектурах нейронных сетей, позволяющие уменьшить влияние существующих проблем на точность выделения зданий.

2.1 Ложноположительные и ложноотрицательные срабатывания при распознавании объектов класса «Здания»

Выделение зданий на космо- и аэрофотоснимках как задача, относящаяся к распознаванию объектов [29], имеет ряд проблем — ложноположительные и ложноотрицательные срабатывания [10, 23, 27], заключающиеся в ошибочном отнесении пикселей к распознаваемым классам.

В качестве основных причин, существенно влияющих на возможность появления указанных проблем, авторы данной статьи выделяют:

- 1) наличие у зданий теней [7, 18, 29], которые могут быть отнесены к объектам класса «Здания» [6, 7, 10, 18, 26, 30] и уменьшают контрастность объектов, «скрывают» их, что, в свою очередь, приводит к менее точному распознаванию [7, 18, 26, 31];

⁵ Прэтт У. Цифровая обработка изображений. Кн. 1; пер. с англ. М: Мир, 1982. 312 с.; Иконика. Цифровая обработка видеоинформации / Под ред. Ю.М. Штарькова. М.: Наука, 1989. 128 с.; Лебедев Д.С. Теория и методы первичной обработки видеоинформации: дис. ... д-ра физ.-мат. наук: 05.13.17 «Теоретические основы информатики». М., 1993. 78 с.

- 2) разные яркостные характеристики зданий [6, 7, 9, 15, 17, 18, 19], которые:
 - в случае разных яркостных характеристик для одного объекта класса «Здания» могут вести к некорректному его разделению на два отдельных объекта [31] или ошибочному отнесению части пикселей этого здания к классу «Фон» [7];
 - в случае их эквивалентности для нескольких зданий могут приводить к выделению близко расположенных зданий как одного объекта [15, 17, 18];
- 3) эквивалентность яркостных характеристик зданий и других объектов на космо- и аэрофотоснимках [15, 21, 23, 27], которая может быть причиной ложноположительных и ложноотрицательных срабатываний при выделении зданий на снимках крупных городов [29, 32].

Указанные причины возникновения ложноположительных и ложноотрицательных срабатываний при выделении зданий с точки зрения семантической сегментации геополей связаны с особенностями исходного множества геополей электромагнитного излучения, представленного космо- и аэрофотоснимками в выборках данных для рассматриваемой задачи.

Таким образом, исходное множество геополей собственного и отраженного элементами суши, океана и атмосферы Земли электромагнитного излучения может оказывать влияние на точность перехода к результирующему геополю пространственного распределения собственного и отраженного объектами класса «Здания» электромагнитного излучения.

В качестве возможных концептуальных решений в рамках архитектур нейронных сетей, позволяющих уменьшить влияние приведенных выше причин возникновения ложноположительных и ложноотрицательных срабатываний, авторы данной статьи предлагают рассматривать:

- 1) обработку дополнительных представлений данных о местности, которыми в задаче выделения зданий являются цифровые модели местности (ЦММ), нормализованные цифровые модели местности (нЦММ) [30, 33] и проблемно-ориентированные спектральные индексы, например нормализованный дифференциальный вегетационный индекс (*англ.* — «Normalized Difference Vegetation Index», NDVI) [30, 32] в рамках архитектур нейронных сетей, как последовательную [30], так и параллельную [32];
- 2) включение в архитектуры нейронных сетей механизмов внимания [9, 21, 25, 34, 35];
- 3) обработку глобальных признаков в рамках архитектур нейронных сетей [6, 17, 34, 36].

Использование дополнительных представлений данных о местности может позволить при выделении зданий с помощью нейронных сетей учитывать высоту при обработке в рамках архитектур нейронных сетей ЦММ и нЦММ, данные о растительности при работе с NDVI, что может позволить уменьшить влияние указанной проблемы на результаты распознавания. Однако данное решение является наименее эффективным, так как эти данные, в особенности ЦММ и нЦММ, требуют постоянной актуализации и, в случае их отсутствия, подготовки, что увеличивает время на предобработку данных перед проведением выделения зданий.

Включение механизмов внимания и обработки глобальных признаков в архитектуры нейронных сетей является более эффективным с точки зрения временных затрат на предобработку данных концептуальными решениями, позволяющими уменьшить влияние указанных причин возникновения ложноположительных и ложноотрицательных срабатываний на результаты распознавания, так как они не требуют для своей работы дополнительных действий по предобработке данных.

Механизмы внимания позволяют обеспечить обработку глобальных взаимосвязей между признаками, что, в свою очередь, положительно сказывается на точности выделения зданий [6, 17, 25, 34, 36] и уменьшает влияние разных и эквивалентных яркостных характеристик на результаты распознавания. Однако включение механизмов внимания может приводить к появлению краевых аномалий.

Обработка глобальных признаков в рамках архитектур нейронных сетей также позволяет повысить точность выделения зданий и снизить влияние указанных выше причин на получаемые результаты [6, 36].

Следует отметить, что механизмы внимания и обработка глобальных признаков не являются взаимоисключающими концептуальными решениями и могут использоваться вместе [36].

Таким образом, при решении проблемы ложноположительных или ложноотрицательных срабатываний при выделении зданий авторы рекомендуют в качестве включаемых в архитектуры нейронных сетей концептуальных решений рассматривать механизмы внимания и обработку глобальных признаков. Однако сравнение данных концептуальных решений по отдельности для выбора наиболее эффективного с точки зрения обеспечиваемой их внедрением точности выделения зданий в настоящее время не проводилось.

2.2 Выделение отдельных зданий как одного объекта

Постоянное повышение плотности застройки как в крупных, так и в малых городах [10–13] повлекло за собой проблему распознавания разных зданий как одного объекта. Данная проблема в контексте семантической сегментации геополей относится к исходному множеству геополей и приводит к снижению точности перехода к результирующему геополю.

В качестве возможных решений данной проблемы предлагается:

- 1) представление задачи выделения зданий в виде нескольких подзадач и включение в архитектуру нейронной сети блоков, обеспечивающих их выполнение, как, например, в [9, 31, 37];
- 2) включение в архитектуры нейронных сетей механизмов внимания [9, 25, 34, 35].

Указанные концептуальные решения не являются взаимоисключающими и могут использоваться вместе [17, 34]. Однако представление задачи выделения зданий в виде нескольких подзадач и включение решающих их блоков в архитектуры нейронных сетей может приводить к необходимости последовательного устранения в рамках архитектур проблем, вызванных с самими блоками [9]. Следовательно, включение большого количества блоков из-за дробления задачи выделения зданий на подзадачи может приводить к снижению точности распознавания моделей.

Таким образом, основным концептуальным решением, по мнению авторов данной статьи, является включение в архитектуры нейронных сетей механизмов внимания.

2.3 Выделение зданий на снимках, имеющих разные яркостные характеристики

Данная проблема заключается:

- в различающихся яркостных характеристиках на снимках разных областей интереса, что, в свою очередь, связано с их территориальным положением [34, 38], характеризующимся разными типами земель, дорожных покрытий и др., различными материалами и типами сооружений [38];
- в разных радиометрических разрешениях и спектральных диапазонах съемочной аппаратуры [34].

Таким образом, в контексте семантической сегментации геополей данная проблема связана как с особенностями исходного множества геополей с точки зрения его пространственной локализации, так и с особенностями аппаратуры, с помощью которой получают цифровые снимки, графически отражающие это множество.

В настоящее время метода обучения нейронных сетей, позволяющего выполнять задачу выделения зданий с учетом разных яркостных характеристик различных территорий и съемочных систем, не существует.

Эта проблема, по мнению авторов, в современном мире может быть решена различными путями, каждый из которых имеет свои особенности, влияющие на его реализуемость и результативность.

В качестве первого решения предлагаются разработка и обучение нейронных сетей для каждой области интереса отдельно, с учетом особенностей каждой конкретной местности. Указанное решение, по мнению авторов данной статьи, возможно, если в рамках обучения нейронных сетей будет учтена изменчивость областей интереса, в противном случае применение его будет затруднительно и потребует постоянной работы по актуализации возможностей распознавания зданий.

В качестве второго решения предлагается составление выборки данных на основе нескольких общедоступных наборов данных, чтобы обеспечить вариативность яркостных характеристик как при разных областях интереса, так и при использовании разной съемочной аппаратуры [24, 39]. Однако данное концептуальное решение не рекомендуется авторами, поскольку в случае использования снимков разного пространственного разрешения, получаемых при работе с разной съемочной аппаратурой, одни проблемы в рамках проблематики выделения зданий могут становиться более значимыми, а другие не оказывать никакого влияния. Кроме того, возникает проблема с отбором как исходных выборок, так и репрезентативных снимков из каждой выборки, их количеством и соотношением, которая может приводить в случае своей неразрешенности к невозможности корректного обучения моделей нейронных сетей.

В качестве третьего решения предлагается аугментация данных, заключающаяся в формировании цифровых изображений с различными поворотами, увеличенными и уменьшенными по отдельности контрастом, яркостью и насыщенностью по исходным снимкам [39, 40]. Расширение выборки путем аугментации является альтернативой добавлению существующих данных, позволяющей в рамках рассматриваемой задачи учесть большую часть возможных изменений яркостных характеристик и работать с определенным пространственным разрешением.

Следует учитывать, что подобная аугментация как метод расширения выборки для улучшения обучаемости модели и качества распознавания имеет ту же особенность, что и любое расширение массива данных: увеличение времени на его обработку.

Тем не менее, по мнению авторов данной статьи, аугментация оптимальна для рассмотрения в качестве решения проблемы выделения зданий на снимках с разными яркостными характеристиками.

3 Результаты

По результатам проведенного анализа для сравнения отобраны следующие архитектуры нейронных сетей, реализующие концептуальные решения:

- 1) U-Net с основой ResNet50 — сверточная сеть с короткими путями (*англ.* — «skip connections», «shortcut connections»), обеспечиваемыми ResNet, описанная в [41] и представляющая собой базовый вариант архитектуры сверточной сети без применения рассмотренных ранее концептуальных решений;
- 2) Sparse Token Transformer (STT) — архитектура, реализующая механизм внимания на основе трансформера [42];
- 3) Swin-Transformer-Based Encoding Booster Integrated in U-Shaped Network (STEB-UNet) — архитектура, реализующая механизм внимания посредством использования трансформера со скользящим окном [43];

- 4) Multiscale Features-Convolutional Neural Network (MF-CNN) — сверточная нейронная сеть [44], в которой работа с глобальными признаками обеспечивается с помощью мультимасштабного модуля;
- 5) Multi-Scale Convolutional Feature Fusion (MSCFF) — сверточная сеть, обеспечивающая работу с глобальными признаками через мультимасштабное слияние признаков посредством свертки [45].

Для проведения обучения, валидации и тестирования используется выборка Уханьского университета для выделения зданий «WHU Building Dataset»⁶ [17], которая считается наиболее широко используемой в рамках задачи выделения зданий и включает в себя 8189 частей аэрофотоснимков в ложной цветопередаче, покрывающих 450 км² города Крайстчерч в Новой Зеландии, и соответствующие им маски.

Метаданные указанной выборки:

- пространственное разрешение — 0,3 м [17, 40];
- радиометрическое разрешение — 8 бит/пиксель;
- разрешение — 512×512 пикселей.

Из основных проблем, которые могут влиять на точность выделения зданий как задачи семантической сегментации геополей, присутствуют разные яркостные характеристики зданий, тени, отбрасываемые зданиями, и узкие границы между зданиями.

Выборка Уханьского университета для выделения зданий была поделена ее составителем на обучающую, валидационную и тестовую подвыборки следующим образом:

- 4736 снимков и масок для проведения обучения;
- 1036 образцов и эталонов для валидации;
- 2416 цифровых изображений и масок для выполнения тестирования.

Для проведения обучения выбраны следующие параметры:

- количество эпох обучения — 20;
- скорость обучения $lr = 0,001$;
- количество образцов, поступающих в рамках одной итерации на эпохе обучения, — 4–8.

При обучении, валидации и тестировании моделей использовалась видеокарта NVIDIA GeForce RTX 3080 Ti. Результаты обучения и валидации отобранных методов нейронных сетей в задаче выделения зданий отражены в табл. 1, 2.

Таблица 1

Результаты обучения моделей на выборке WHU-BE, представленные с шагом в 5 эпох (среднее значение бинарной кросс-энтропии)

Table 1

The results of training models on the WHU-BE dataset, presented in increments of 5 epochs (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	
STT	0,1685	0,0758	0,0656	0,0647	0,0595	03:46:46
STEB-UNet	0,2386	0,1735	0,1424	0,1310	0,1103	04:45:22
U-Net (ResNet50)	0,0593	0,0228	0,0181	0,0153	0,0134	03:35:08
MF-CNN	0,0478	0,0251	0,0208	0,0185	0,0171	01:59:23
MSCFF	0,0446	0,0204	0,0166	0,0142	0,0118	14:56:36

Таблица 2

Результаты валидации моделей на выборке WHU-BE, представленные с шагом в 5 эпох (метрика IoU)

Table 2

The results of model validation on the WHU-BE dataset, presented in increments of 5 epochs (IoU metric)

Модель	Эпоха				
	1-я	5-я	10-я	15-я	20-я
STT	0,8484	0,8763	0,8852	0,9006	0,9045
STEB-UNet	Валидация не предусмотрена				
U-Net (ResNet50)	0,6521	0,7895	0,8366	0,8470	0,8602
MF-CNN	0,5651	0,8104	0,7973	0,7851	0,8537
MSCFF	0,7643	0,7962	0,8498	0,8539	0,8790

Примечание. Жирным шрифтом выделены лучшие результаты.

6 WHU building dataset. Kaggle.com, 2023. [Электронный ресурс]. Режим доступа: <https://www.kaggle.com/datasets/xiaoqian970429/whu-building-dataset> (дата обращения: 27.11.2023).

После обучения и валидации выполнено тестирование. При тестировании в качестве основных метрик выбраны IoU и F_1 , представленные выражениями (1) и (2). Это связано с тем, что с помощью IoU можно оценить точность распознавания объектов класса «Здания», поскольку данная метрика является наиболее строгой для оценки точности. F_1 используется для оценки качества моделей, так как представляет собой гармоническое среднее между Precision и Recall: отражает, насколько модель избегает ошибочного отнесения пикселя класса «Фон» к классу «Здания» и насколько модель избегает пропуска пикселей, относящихся к классу «Здания». Для тестирования выбраны модели, полученные в результате проведения 20 эпох обучения в связи с отсутствием переобученности и лучшими показанными в ходе обучения и валидации результатами. Результаты тестирования представлены в табл. 3.

1 ➤

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

2 ➤

$$F_1 = 2 \frac{Precision \times Recall}{Precision + Recall}.$$

Таблица 3 ➤

Результаты тестирования моделей на выборке WHU-BE

Table 3

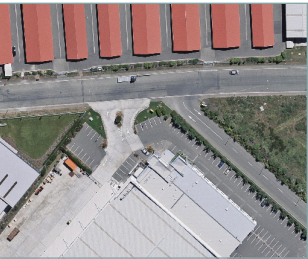
Results of testing models on the WHU-BE dataset

Модель	IoU	F_1	OA	Затраченное время (без учета загрузки модели), чч:мм:сс
STT	0,8808	0,9366	0,9860	00:12:52
STEB-UNet	0,4073	0,5788	0,9305	00:19:12
U-Net (ResNet50)	0,8642	0,9271	0,9840	00:06:11
MF-CNN	0,8496	0,9187	0,9821	00:05:36
MSCFF	0,8729	0,9321	0,9848	00:16:08

Для визуального представления результатов выполненного тестирования и обеспечения возможности оценки качества моделей оптическим способом выбраны:

- 1) производственная застройка, сливающаяся с фоном (рис. 1);
- 2) территория города Крайстчерч с частными домами, имеющими разные яркостные характеристики крыш (рис. 2);
- 3) территория, на которой отсутствует застройка, представляющая собой пустырь (рис. 3);
- 4) территория без зданий, представленная сельскохозяйственными землями (рис. 4).

Снимок / Image



Эталон / Ground truth



STT / STT



STEB-UNet / STEB-UNet



Рис. 1 ➤

Пример результатов распознавания класса «Здания» на территории с производственной застройкой

Fig. 1

Example of class “Buildings” recognition results on industrial area

U-Net (ResNet50) / U-Net (ResNet50)

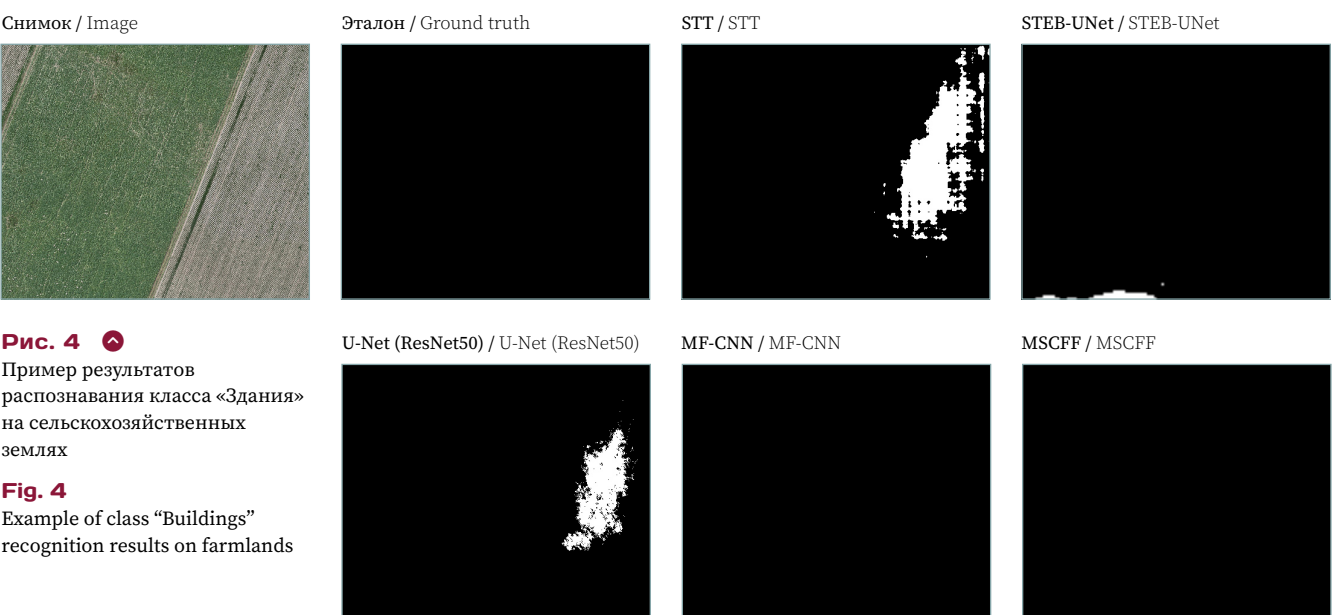
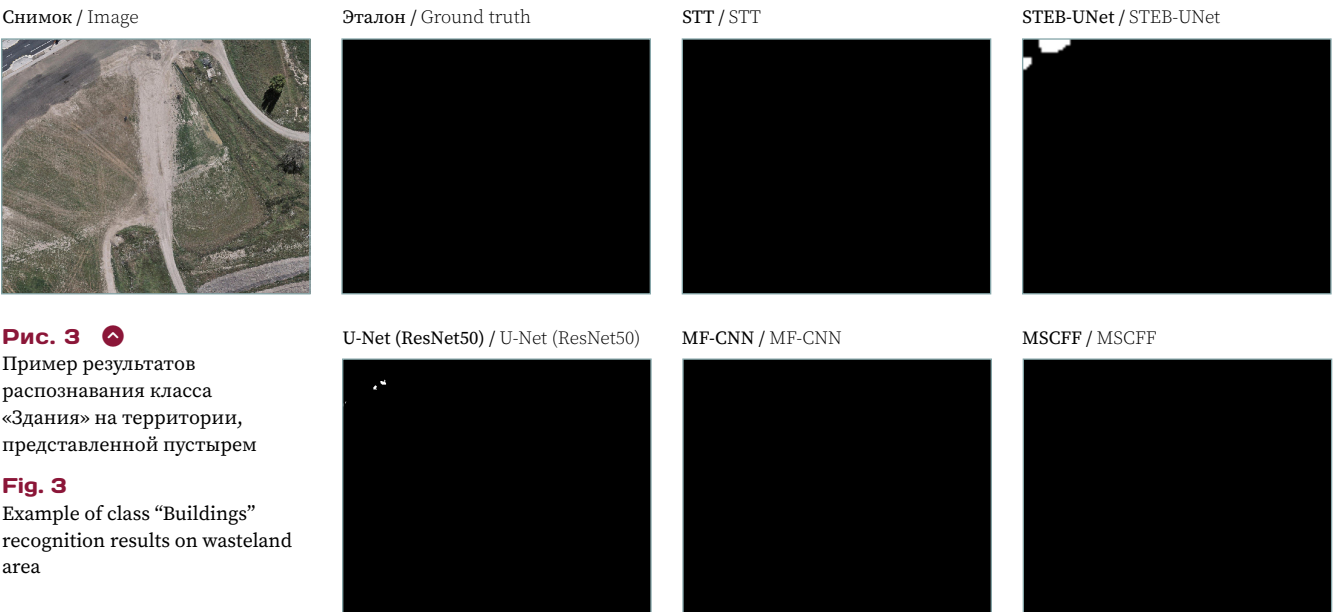
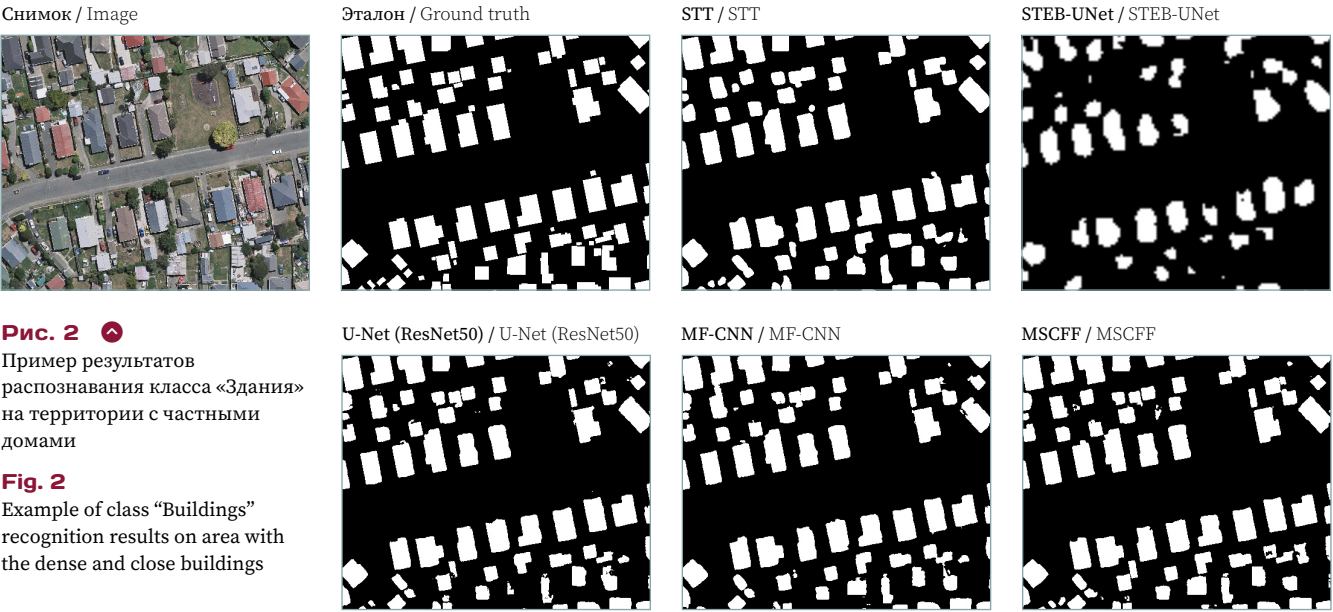


MF-CNN / MF-CNN



MSCFF / MSCFF





4 Обсуждение

На основании полученных в ходе обучения и валидации результатов можно заключить следующее:

- 1) MSCFF имеет наилучшую обучаемость среди всех отобранных методов, однако используемое в архитектуре данного метода мультимасштабное слияние признаков посредством свертки приводит к значительному увеличению временных затрат;
- 2) STEB-UNet, реализующий механизм внимания посредством трансформера со скользящим окном, имеет низкий предел обучаемости;
- 3) переобучение моделей в ходе проводимого тестирования отсутствует.

На основе результатов выполненного тестирования в рамках задачи выделения зданий как задачи семантической сегментации геополей сделаны следующие основные выводы о концептуальных решениях:

- 1) использование обработки глобальных признаков обеспечивает наилучшие результаты относительно точности модели, что подтверждается высокими значениями метрики IoU;
- 2) на основе анализа результатов оптическим способом выявлено, что:
 - мультимасштабное слияние признаков позволяет обеспечить отсутствие краевых аномалий при распознавании класса «Здания» на цифровых изображениях, покрывающих область интереса без застройки;
 - использование механизмов внимания обеспечивает решение проблемы распознавания отдельных зданий как одного объекта класса «Здания»;
 - шероховатость на сельскохозяйственных землях, вызванная наличием растительности, является наиболее сложным случаем для выделения зданий, что указывает на еще одну проблему, относящуюся к исходному множеству геополей, которая влияет на точность перехода к результирующему геополю;
- 3) наиболее быстрой с точки зрения выдачи результатов распознавания является MF-CNN, а наиболее медленной — MSCFF, что отражает наличие зависимости между реализацией обработки глобальных признаков в рамках архитектур нейронных сетей и временными затратами на выполнение выделения зданий;
- 4) механизмы внимания и мультимасштабное слияние признаков наиболее эффективны с точки зрения обеспечения точности и качества моделей концептуальными решениями в рамках архитектур отобранных методов нейронных сетей.

Исходя из вышеизложенного, авторы данной статьи предлагают:

- 1) обеспечивать работу с глобальными признаками в рамках архитектур нейронных сетей посредством мультимасштабного слияния признаков;
- 2) в зависимости от временных ограничений и области интереса рассматривать в качестве концептуальных решений, которые необходимо внедрять в архитектуры нейронных сетей, механизмы внимания и обработку глобальных признаков.

5 Выводы

Сформулирована задача выделения зданий в контексте семантической сегментации геополей, определена и проанализирована проблематика выделения зданий в данном контексте.

В качестве основных проблем при выделении зданий указаны ложноположительные и ложноотрицательные срабатывания при выделении зданий, выделение отдельных зданий как одного объекта и распознавание объектов класса «Здания» на снимках, имеющих разные яркостные характеристики.

На основе анализа указанных проблем определены их основные концептуальные решения в архитектурах нейронных сетей — механизмы внимания и обработка глобальных признаков.

Экспериментальные результаты указывают на наличие зависимости между временными затратами и конкретной реализацией концептуальных решений.

Кроме того, исходя из данных результатов определено наиболее эффективное с точки зрения обеспечиваемой точности концептуальное решение — обработка глобальных признаков посредством мультимасштабного слияния признаков. Это концептуальное решение, по мнению авторов, требует развития с точки зрения снижения временных затрат, требуемых для обучения модели и выделения зданий, при его внедрении в архитектуры нейронных сетей.

Исследование задачи выделения зданий в контексте семантической сегментации геополей закладывает новую основу для рассмотрения задач распознавания образов. В связи с этим представляется целесообразным изучение других задач компьютерного зрения, направленных на автоматизированную обработку данных дистанционного зондирования Земли с помощью методов нейронных сетей в этом контексте и развитие методов выделения зданий с точки зрения используемых функций потерь.

БЛАГОДАРНОСТИ

Авторы выражают благодарность рецензентам и редакторам за ценные замечания и оценку работы.

БИБЛИОГРАФИЯ

1. Гвоздев О.Г., Матерухин А.В., Майоров А.А. Сравнительный анализ методов определения сходства геополей // Вестник СГУГиТ. 2022. Т. 27. № 6. С. 120–130. DOI:10.33764/2411-1759-2022-27-6-120-130.
2. Копнов М.В. Модели, алгоритмы и программные средства для анализа геополей: дис. ... канд. техн. наук: 05.13.01 «Системный анализ, управление и обработка информации: промышленность». 2010. 160 с.
3. Каковкин П.А., Друки А.А. Применение нейросетевых алгоритмов для решения задачи восстановления двумерных геополей // Известия Томского политехнического университета. Инжиниринг георесурсов. 2016. Т. 327. № 2. С. 35–43.
4. Копнов М.В., Марков Н.Г. Геоинформационная система для восстановления пространственно-временных геополей // Проблемы информатики. 2011. № 5. С. 172–179.
5. Ostankovich V., Afanasyev I. Illegal buildings detection from satellite images using GoogLeNet and cadastral map // Proceedings of 9th IEEE International Conference on Intelligent Systems. 2018. P. 1–8. DOI:10.1109/IS.2018.8710565.
6. 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. 2585. DOI:10.3390/rs15102585.
7. Miao Y., Jiang S., Xu Y., Wang D. Feature residual analysis network for building extraction from remote sensing images // Applied Sciences. 2022. Vol. 12. Iss. 10. 5095. DOI:10.3390/app12105095.
8. You D., Wang S., Wang F., et al. EfficientUNet+: a building extraction method for emergency shelters based on deep learning // Remote Sensing. 2022. Vol. 14. Iss. 9. 2207. DOI:10.3390/rs14092207.
9. Xu G., Deng M., Sun G., et al. Improving building extraction by using knowledge distillation to reduce the impact of label noise // Remote Sensing. 2022. Vol. 14. Iss. 22. 5645. DOI:10.3390/rs14225645.
10. Dikmen M., Halici U. A learning-based resegmentation method for extraction of buildings in satellite // IEEE Geoscience and Remote Sensing Letters. 2014. Vol. 11. Iss. 12. P. 2150–2153. DOI:10.1109/LGRS.2014.2321658.
11. Mia M.R., Islam K.S., Islam M.D. Automatic building footprint extraction from high resolution stereo satellite image // Plan Plus. 2021. Vol. 11. Iss. 1. P. 17–28. DOI:10.54470/planplus.v11i1.2.

12. Farhadi H., Ebadi H., Kiani A. F2BFE: development of feature-based building footprint extraction by remote sensing data and GEE // *International Journal of Remote Sensing*. 2023. Vol. 44. Iss. 19. P. 5845–5875. DOI:10.1080/01431161.2023.2255351.
13. He H., Gao K., Tan W., et al. Impact of deep learning based super resolution on building footprint extraction // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022. Vol. XLIII-B1-2022. P. 31–37. DOI:10.5194/isprs-archives-XLIII-B1-2022-31-2022.
14. Liao C., Hu H., Yuan X., et al. BCE-Net: reliable building footprints change extraction based on historical map and up-to-date images using contrastive learning // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2023. Vol. 201. P. 138–152. DOI:10.1016/j.isprsjprs.2023.05.011.
15. Wang Z., Li S., Zhu Z. Rural building extraction based on joint U-Net and the generalized Chinese restaurant franchise from remote sensing images // *Sustainability*. 2023. Vol. 15. Iss. 5. 4685. DOI:10.3390/su15054685.
16. Quan Y., Yu A., Cao X., et al. Building extraction from remote sensing images with DoG as prior constraint // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022. Vol. 15. P. 6559–6570. DOI:10.1109/JSTARS.2022.3195808.
17. Jin H., Fu W., Nie C., et al. Extraction of building from remote sensing imagery base on multi-attention L-CAFSFM and MFFM // *Frontiers in Earth Science*. 2023. Vol. 11. 1268628. DOI:10.3389/feart.2023.1268628.
18. Sun Y., Chen J., Huang X., Zhang H. Multi-level perceptual network for urban building extraction from high-resolution remote sensing images // *Photogrammetric Engineering & Remote Sensing*. 2023. Vol. 89. No. 7. P. 427–434. DOI:10.14358/PERS.22-00103R1.
19. Qiu Y., Wu F., Yin J., et al. MSL-Net: An efficient network for building extraction from aerial imagery // *Remote Sensing*. 2022. Vol. 14. Iss. 16. 3914. DOI:10.3390/rs14163914.
20. Angelis G.-E., Domi A., Zamichos A., et al. On the exploration of vision transformers in remote sensing building extraction // *IEEE International Symposium on Multimedia (ISM)*. 2022. P. 208–215. DOI:10.1109/ISM55400.2022.00046.
21. Zhou J., Liu Y., Nie G., et al. Building extraction and floor area estimation at the village level in rural China via a comprehensive method integrating UAV photogrammetry and the novel EDSANet // *Remote Sensing*. 2022. Vol. 14. Iss. 20. 5175. DOI:10.3390/rs14205175.
22. Atas İ. Performance evaluation of Jaccard-Dice coefficient on building segmentation from high resolution satellite images // *Balkan Journal of Electrical & Computer Engineering*. 2023. Vol. 11. No. 1. P. 100–106. DOI:10.17694/bajece.1212563.
23. He W., Li J., Cao W., et al. Building extraction from remote sensing images via an uncertainty-aware network // *Preprint arXiv.org*, 2023. [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/2307.12309.pdf> (дата обращения: 27.11.2023).
24. Weng Q., Wang Q., Lin Y., Lin J. ARE-Net: an improved interactive model for accurate building extraction in high-resolution remote sensing imagery // *Remote Sensing*. 2023. Vol. 15. Iss. 18. 4457. DOI:10.3390/rs15184457.
25. Yuan W., Ran W., Shi X., Shibasaki R. Multiconstraint transformer-based automatic building extraction from high-resolution remote sensing images // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023. Vol. 16. P. 9164–9174. DOI:10.1109/JSTARS.2023.3319826.
26. Boonpook W., Tan Y., Torsri K., et al. PCL-PTD Net: parallel cross-learning-based pixel transferred deconvolutional network for building extraction in dense building areas with shadow // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022. Vol. 16. P. 773–786. DOI:10.1109/JSTARS.2022.3230149.
27. Wang Z., Xu N., Wang B., et al. Urban building extraction from high-resolution remote sensing imagery based on multi-scale recurrent conditional generative adversarial network // *GIScience & Remote Sensing*. 2022. Vol. 59. Iss. 1. P. 861–884. DOI:10.1080/15481603.2022.2076382.
28. Якушенков Ю.Г. Некоторые особенности компьютерного моделирования оптико-электронных систем визуализации // *Международная научно-техническая конференция, посвященная 225-летию МИИГАиК*. 2004. С. 116–122.
29. Shrestha S., Vanneschi L. Improved fully convolutional network with conditional random fields for building extraction // *Remote Sensing*. 2018. Vol. 10. Iss. 7. 1135. DOI:10.3390/rs10071135.


30. Xu Y., Wu L., Xie Z., Chen Z. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters // *Remote Sensing*. 2018. Vol. 10. Iss. 1. 144. DOI:10.3390/rs10010144.
31. Wagner F.H., Dalagnol R., Tarabalka Y., et al. U-Net-Id, an instance segmentation model for building extraction from satellite images — case study in the Joanópolis city, Brazil // *Remote Sensing*. 2020. Vol. 12. Iss. 10. 1544. DOI:10.3390/rs12101544.
32. Schuegraf P., Bittner K. Automatic building footprint extraction from multi-resolution remote sensing images using hybrid FCN // *Information Journal of Geo-Information*. 2019. Vol. 8. Iss. 4. 191. DOI:10.3390/ijgi8040191.
33. Aslani M., Seipel S. A fast instance selection method for support vector machines in building extraction // *Applied Soft Computing Journal*. 2020. Vol. 97. Part B. 106716. DOI:10.1016/j.asoc.2020.106716.
34. Yu B., Chen F., Wang N., et al. MSFTrans: a multi-task frequency-spatial learning transformer for building extraction from high spatial resolution remote sensing images // *GIScience & Remote Sensing*. 2022. Vol. 59. Iss. 1. P. 1978–1996. DOI:10.1080/15481603.2022.2143678.
35. Song J., Zhu A.-X., Zhu Y. Transformer-based semantic segmentation for extraction of building footprints from very-high-resolution images // *Sensors*. 2023. Vol. 23. Iss. 11. 5166. DOI:10.3390/s23115166.
36. Han J., Wang Z., Wang Y., Hou W. Building extraction algorithm from remote sensing images based on improved DeepLabv3+ network // *Journal Of Physics: Conference Series*. 2022. Vol. 2303. 012010.
37. Larionov R., Khryashchev V., Pavlov V., Demidov P.G. Separation of closely located buildings on aerial images using U Net neural network // *Proceedings of 26th Conference of Open Innovations Association (FRUCT)*. 2020. P. 1–6. DOI:10.23919/FRUCT48808.2020.9087365.
38. Taha L.G.E., Ibrahim R.E. Assessment of approaches for the extraction of building footprints from Pléiades images // *Geomatics and Environmental Engineering*. 2021. Vol. 15. No. 4. P. 101–116. DOI:10.7494/geom.2021.15.4.101.
39. Luo M., Ji S., Wei S. A diverse large-scale building dataset and a novel plug-and-play domain generalization method for building extraction // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023. Vol. 16. P. 4122–4138. DOI:10.1109/JSTARS.2023.3268176.
40. Wang Z., Zhou Y., Wang F., et al. A multi-scale edge constraint network for the fine extraction of buildings from remote sensing images // *Remote Sensing*. 2023. Vol. 15. Iss. 4. 927. DOI:10.3390/rs15040927.
41. Alsabhan W., Alotaiby T. Automatic building extraction on satellite images using Unet and ResNet50 // *Computational Intelligence and Neuroscience*. 2022. Vol. 2022. 5008854. DOI:10.1155/2022/5008854.
42. Chen K., Zou Z., Shi Z. Building extraction from remote sensing images with sparse token transformers // *Remote Sensing*. 2021. Vol. 13. Iss. 21. 4441. DOI:10.3390/rs13214441.
43. Xiao X., Guo W., Chen R., et al. A swin transformer-based encoding booster integrated in U-shaped network for building extraction // *Remote Sensing*. 2022. Vol. 14. Iss. 11. 2611. DOI:10.3390/rs14112611.
44. Shao Z., Pan Y., Diao C., Cai J. Cloud detection in remote sensing images based on multiscale features-convolutional neural network // *IEEE Transactions on Geoscience and Remote Sensing*. 2019. Vol. 57. Iss. 6. P. 4062–4072. DOI:10.1109/TGRS.2018.2889677.
45. Li Z., Shen H., Cheng Q., et al. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2019. Vol. 150. P. 197–212. DOI:10.1016/j.isprsjprs.2019.02.017.

АВТОРЫ **Бирюков Никита Андреевич**


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

 0000-0001-9856-4696

Майоров Андрей Александрович

ФГБОУ ВО «Московский государственный университет геодезии и картографии»
(МИИГАиК), Москва, Россия
кафедра информационно-измерительных систем, факультет геоинформатики
и информационной безопасности
д-р техн. наук, профессор
 0000-0001-5539-3719

Лапчинская Маргарита Петровна

ФГБОУ ВО «Московский государственный университет геодезии и картографии»
(МИИГАиК), Москва, Россия
кафедра информационно-измерительных систем, факультет геоинформатики
и информационной безопасности
канд. техн. наук, доцент
 0000-0002-7354-8283

Поступила 01.12.2023. Принята к публикации 22.02.2024. Опубликовано 28.02.2024.



Geofields semantic segmentation by neural networks on example of building extraction on satellite and aerial images problematic

Nikita A. Biryukov¹✉, Andrei A. Maiorov¹,
Margarita P. Lapchinskaya¹

¹ Moscow State University of Geodesy and Cartography, Moscow, Russia
✉ 4096414@gmail.com

CITATION Biryukov NA, Maiorov AA, Lapchinskaya MP. 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. DOI:10.30533/GiA-2024-004.

KEYWORDS geofields, semantic segmentation, building extraction, neural networks, pattern recognition, machine learning, computer vision, data preprocessing

ABSTRACT The Building extraction in a geofields semantic segmentation context is a task of transitioning from geofields set of its own and reflected by Earth's land, ocean and atmosphere elements electromagnetic radiation to a spatial distribution geofields set of its own and reflected by class "Buildings" objects electromagnetic radiation. Building extraction problematic as geofields semantic segmentation includes such main problems as: the false positives and false negatives, the extraction of separate buildings as one object and class "Buildings" objects recognition on images with different brightness. These problems in geofields semantic segmentation context considered as leading to the accuracy decrease of transition from original geofields set to a resulting geofields set and mainly applies to the original set. By mentioned problems analysis main conceptual solutions in the neural networks architectures identified those allowing reduce their effect these are attention mechanisms and global feature processing. As a result of the computational experiments with STT, STEB-UNet, U-Net, MF-CNN and MSCFF neural networks architectures those implement these conceptual solutions it was determined that there is a processing time and the specific conceptual solutions implementation and the most effective in accuracy terms conceptual solution that is the global features processing via multi-scale feature fusion implemented by MSCFF.

ACKNOWLEDGEMENTS The authors are grateful to the reviewers and the editors for their valuable comments and assessment of the work.


REFERENCES

1. Gvozdev OG, Materukhin AV, Maiorov AA. Comparative analysis of methods for the similarity assessment of geofields. *Vestnik SSUGT*. 2022;27(6): 120–130. (In Russian). DOI:10.33764/2411-1759-2022-27-6-120-130.
2. Kopnov MV. *Modeli, algoritmi i programmie sredstva dlia analiza geopolei* [Models, algorithms and software for geofields analysis]: PhD thesis. Tomsk: 2010; 160 p. (In Russian).
3. Kakovin PA, Druki AA. Primeneniye neirossetevykh algoritmov dlia resheniya zadachi vosstanovleniya dvumernykh geopolei [Application of artificial neural networks for solving the problem of two-dimensional geofields reconstruction]. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*. 2016;327(2): 35–43. (In Russian).
4. Kopnov MV, Markov NG. Geoinformatsionnaya sistema dlia vosstanovleniya prostranstvenno-vremennikh geopolei [Geographic information system for the spatiotemporal geofields reconstruction]. *Problemy informatiki*. 2011;5: 172–179. (In Russian).
5. Ostankovich V, Afanasyev I. Illegal buildings detection from satellite images using GoogLeNet and cadastral map. *Proceedings of 9th IEEE International Conference on Intelligent Systems*. 2018: 1–8. DOI:10.1109/IS.2018.8710565.
6. 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.
7. Miao Y, Jiang S, Xu Y, Wang D. Feature residual analysis network for building extraction from remote sensing images. *Applied Sciences*. 2022;12(10): 5095. DOI:10.3390/app12105095.
8. You D, Wang S, Wang F, et al. EfficientUNet+: a building extraction method for emergency shelters based on deep learning. *Remote Sensing*. 2022;14(9):2207. DOI:10.3390/rs14092207.
9. Xu G, Deng M, Sun G, et al. Improving building extraction by using knowledge distillation to reduce the impact of label noise. *Remote Sensing*. 2022;14(22): 5645. DOI:10.3390/rs14225645.
10. Dikmen M, Halici U. A learning-based resegmentation method for extraction of buildings in satellite. *IEEE Geoscience and Remote Sensing Letters*. 2014;11(12): 2150–2153. DOI:10.1109/LGRS.2014.2321658.
11. Mia MR, Islam KS, Islam MD. Automatic building footprint extraction from high resolution stereo satellite image. *Plan Plus*. 2021;11(1): 17–28. DOI:10.54470/planplus.v11i1.2.
12. Farhadi H, Ebadi H, Kiani A. F2BFE: development of feature-based building footprint extraction by remote sensing data and GEE. *International Journal of Remote Sensing*. 2023;44(19): 5845–5875. DOI:10.1080/01431161.2023.2255351.
13. He H, Gao K, Tan W, et al. Impact of deep learning based super resolution on building footprint extraction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022;XLIII-B1-2022: 31–37. DOI:10.5194/isprs-archives-XLIII-B1-2022-31-2022.
14. Liao C, Hu H, Yuan X, et al. BCE-Net: reliable building footprints change extraction based on historical map and up-to-date images using contrastive learning. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2023;201: 138–152. DOI:10.1016/j.isprsjprs.2023.05.011.
15. Wang Z, Li S, Zhu Z. Rural building extraction based on joint U-Net and the generalized Chinese restaurant franchise from remote sensing images. *Sustainability*. 2023;15(5): 4685. DOI:10.3390/su15054685.
16. Quan Y, Yu A, Cao X, et al. Building extraction from remote sensing images with DoG as prior constraint. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022;15: 6559–6570. DOI:10.1109/JSTARS.2022.3195808.
17. Jin H, Fu W, Nie C, et al. Extraction of building from remote sensing imagery base on multi-attention L-CAFSFM and MFFM. *Frontiers in Earth Science*. 2023;11: 1268628. DOI:10.3389/feart.2023.1268628.
18. Sun Y, Chen J, Huang X, Zhang H. Multi-level perceptual network for urban building extraction from high-resolution remote sensing images. *Photogrammetric Engineering & Remote Sensing*. 2023;89(7): 427–434. DOI:10.14358/PERS.22-00103R1.
19. Qiu Y, Wu F, Yin J, et al. MSL-Net: An efficient network for building extraction from aerial imagery. *Remote Sensing*. 2022;14(16): 3914. DOI:10.3390/rs14163914.
20. Angelis G-E, Domi A, Zamichos A, et al. On the exploration of vision transformers in remote sensing building extraction. *IEEE International Symposium on Multimedia (ISM)*. 2022: 208–215. DOI:10.1109/ISM55400.2022.00046.


21. Zhou J, Liu Y, Nie G, et al. Building extraction and floor area estimation at the village level in rural China via a comprehensive method integrating UAV photogrammetry and the novel EDSANet. *Remote Sensing*. 2022;14(20): 5175. DOI:10.3390/rs14205175.
22. Atas İ. Performance evaluation of Jaccard-Dice coefficient on building segmentation from high resolution satellite images. *Balkan Journal of Electrical & Computer Engineering*. 2023;11(1): 100–106. DOI:10.17694/bajece.1212563.
23. He W, Li J, Cao W, et al. Building extraction from remote sensing images via an uncertainty-aware network. *Preprint arXiv.org*, 2023. Available from: <https://arxiv.org/pdf/2307.12309.pdf> (Accessed: 27 Nov 2023).
24. Weng Q, Wang Q, Lin Y, Lin J. ARE-Net: an improved interactive model for accurate building extraction in high-resolution remote sensing imagery. *Remote Sensing*. 2023;15(18): 4457. DOI:10.3390/rs15184457.
25. Yuan W, Ran W, Shi X, Shibasaki R. Multiconstraint transformer-based automatic building extraction from high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023;16: 9164–9174. DOI:10.1109/JSTARS.2023.3319826.
26. Boonpook W, Tan Y, Torsri K, et al. PCL-PTD Net: parallel cross-learning-based pixel transferred deconvolutional network for building extraction in dense building areas with shadow. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022;16: 773–786. DOI:10.1109/JSTARS.2022.3230149.
27. Wang Z, Xu N, Wang B, et al. Urban building extraction from high-resolution remote sensing imagery based on multi-scale recurrent conditional generative adversarial network. *GIScience & Remote Sensing*. 2022;59(1): 861–884. DOI:10.1080/15481603.2022.2076382.
28. Yakushenkov YG. *Nekotore osobennosti komp'yuternogo modelirovaniia optiko-elektronnykh sistem vizualizatsii* [Some features of optical-electronic visualization systems computer modeling]. Mezhdunarodnaia nauchno-tehnicheskaiia konferentsiia, posviashennaya 225-letiyu MIIGAiK. 2004: 116–122.
29. Shrestha S, Vanneschi L. Improved fully convolutional network with conditional random fields for building extraction. *Remote Sensing*. 2018;10(7): 1135. DOI:10.3390/rs10071135.
30. Xu Y, Wu L, Xie Z, Chen Z. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sensing*. 2018;10(1): 144. DOI:10.3390/rs10010144.
31. Wagner FH, Dalagnol R, Tarabalka Y, et al. U-Net-Id, an instance segmentation model for building extraction from satellite images — case study in the Joanópolis city, Brazil. *Remote Sensing*. 2020;12(10): 1544. DOI:10.3390/rs12101544.
32. Schuegraf P, Bittner K. Automatic building footprint extraction from multi-resolution remote sensing images using hybrid FCN. *Information Journal of Geo-Information*. 2019;8(4):191. DOI:10.3390/ijgi8040191.
33. Aslani M, Seipel S. A fast instance selection method for support vector machines in building extraction. *Applied Soft Computing Journal*. 2020;97(B):106716. DOI:10.1016/j.asoc.2020.106716.
34. Yu B, Chen F, Wang N, et al. MSFTrans: a multi-task frequency-spatial learning transformer for building extraction from high spatial resolution remote sensing images. *GIScience & Remote Sensing*. 2022;59(1):1978–1996. DOI:10.1080/15481603.2022.2143678.
35. Song J, Zhu A-X, Zhu Y. Transformer-based semantic segmentation for extraction of building footprints from very-high-resolution images. *Sensors*. 2023;23(11): 5166. DOI:10.3390/s23115166.
36. Han J, Wang Z, Wang Y, Hou W. Building extraction algorithm from remote sensing images based on improved DeepLabv3+ network. *Journal Of Physics: Conference Series*. 2022;2303: 012010.
37. Larionov R, Khryashchev V, Pavlov V, Demidov PG. Separation of closely located buildings on aerial images using U Net neural network. *Proceedings of 26th Conference of Open Innovations Association (FRUCT)*. 2020: 1–6. DOI:10.23919/FRUCT48808.2020.9087365.
38. Taha LGE, Ibrahim RE. Assessment of approaches for the extraction of building footprints from Pléiades images. *Geomatics and Environmental Engineering*. 2021;15(4):101–116. DOI:10.7494/geom.2021.15.4.101.
39. Luo M, Ji S, Wei S. A diverse large-scale building dataset and a novel plug-and-play domain generalization method for building extraction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2023;16: 4122–4138. DOI:10.1109/JSTARS.2023.3268176.

40. Wang Z, Zhou Y, Wang F, et al. A multi-scale edge constraint network for the fine extraction of buildings from remote sensing images. *Remote Sensing*. 2023;15(4): 927. DOI:10.3390/rs15040927.
41. Alsabhan W, Alotaiby T. Automatic building extraction on satellite images using Unet and ResNet50. *Computational Intelligence and Neuroscience*. 2022;2022: 5008854. DOI:10.1155/2022/5008854.
42. Chen K, Zou Z, Shi Z. Building extraction from remote sensing images with sparse token transformers. *Remote Sensing*. 2021;13(21): 4441. DOI:10.3390/rs13214441.
43. Xiao X, Guo W, Chen R, et al. A swin transformer-based encoding booster integrated in U-shaped network for building extraction. *Remote Sensing*. 2022;14(11): 2611. DOI:10.3390/rs14112611.
44. Shao Z, Pan Y, Diao C, Cai J. Cloud detection in remote sensing images based on multiscale features-convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*. 2019;57(6): 4062–4072. DOI:10.1109/TGRS.2018.2889677.
45. Li Z, Shen H, Cheng Q, Liu Y, et al. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2019;150: 197–212. DOI:10.1016/j.isprsjprs.2019.02.017.


AUTHORS**Nikita A. Biryukov**

Moscow State University of Geodesy and Cartography, Moscow, Russia
 Department of Information-Measuring Systems,
 Faculty of Geoinformatics and Information Security
 Postgraduate student
 0000-0001-9856-4696

Andrei A. Maiorov

Moscow State University of Geodesy and Cartography, Moscow, Russia
 Department of Information-Measuring Systems,
 Faculty of Geoinformatics and Information Security
 Dr. Sci. (Engineering), Professor
 0000-0001-5539-3719

Margarita P. Lapchinskaya

Moscow State University of Geodesy and Cartography, Moscow, Russia
 Department of Information-Measuring Systems,
 Faculty of Geoinformatics and Information Security
 PhD in Engineering, Assistant Professor
 0000-0002-7354-8283

Submitted: December 01, 2023. Accepted: February 22, 2024. Published: February 28, 2024.