

Экспериментальная оценка точности мультиклассовой сегментации объектов на спутниковых снимках на основе модифицированной свёрточной нейронной сети U-net

С. М. Багаев, Е. В. Медведева

Вятский государственный университет, Киров, 610000, Россия
E-mails: hedgehog_rus1@mail.ru, emedv@mail.ru

Представлена методология исследования точности сегментации целевых объектов на спутниковых снимках, принадлежащих разным классам, на основе модифицированной свёрточной нейронной сети U-net. Рассмотрены основные этапы сегментации: подготовки входных данных; модификации структуры свёрточной нейронной сети (СНС) с учётом параметров исследуемых изображений; обучения СНС на сформированной обучающей выборке; сегментации тестовых снимков. Для обучения и тестирования модифицированной СНС U-net использованы мультиспектральные снимки, полученные со спутника WorldView-3. Проведены экспериментальные исследования по повышению точности сегментации целевых объектов. Для реализации метода мультиклассовой сегментации предложены два алгоритма. Первый алгоритм реализуется на единой СНС. Второй — на основе отдельно обученных СНС для работы с каждым конкретным классом и с последующим объединением результатов их работы. Средняя точность сегментации первым алгоритмом по метрике Жаккара составила 76 %, по метрике *F*-мера — 0,66. Второй алгоритм повысил точность сегментации объектов похожих и редких классов на 23 % по метрике Жаккара и на 26 % по метрике *F*-мера при ограниченной тренировочной выборке изображений и вычислительных ресурсах.

Ключевые слова: мультиклассовая сегментация, спутниковые снимки, свёрточная нейронная сеть, оценка точности сегментации

Одобрена к печати: 09.11.2021

DOI: 10.21046/2070-7401-2021-18-6-35-45

Введение

Быстрое развитие спутниковых систем дистанционного зондирования Земли (ДЗЗ), всё большая доступность к спутниковым базам данных позволяют получать информацию о различных географических объектах с большим пространственным разрешением и с высокой частотой повторения, создавать новые методы и подходы для решения задач определения типов земной поверхности, оценки площадей областей землепользования, лесных массивов, природных пожаров, нефтяных загрязнений и т. п. (Шовенгердт, 2010; Persello et al., 2019; Yuan, Wang, 2014).

Базовой операцией для решения перечисленных задач выступает сегментация. Существуют различные подходы к сегментации спутниковых снимков, применение которых позволяет ускорить процесс выделения целевых объектов. Для сегментации объектов, внутри которых яркость или цвет пикселей изменяются незначительно (например, искусственные строения, дороги, реки), могут быть применены методы на основе выделения контуров с последующим закрашиванием выделенных и относящихся к объектам интереса сегментов (Гонсалес, Вудс, 2012; Шапиро, Стокман, 2006; Medvedeva, Kurbatova, 2015). Методы, использующие информацию о контурах, не требуют проведения обучения или заготовки специальных шаблонов и позволяют обнаруживать границы объектов любых размеров и форм. Однако из-за изменчивости яркости, нечётких границ контурные методы могут создавать неполные объекты и разбивать их на несколько сегментов.

В случае если целевые объекты имеют нечётко выраженные границы и характерную текстуру (растительность, поля, горные массивы), применяют методы на основе извлечения признаков текстуры. Основные методы извлечения текстур — структурные, статистические и спектральные. Структурные методы эффективны при анализе регулярных текстур (Li et al.,

2014), спектральные — при отыскании периодичности в изображении (Yuan, Wang, 2014). Более универсальными представляются статистические методы, в которых в качестве статистических характеристик используют матрицы смежности, длины серий по направлениям, автокорреляцию, разность уровней серого, локальную частоту флуктуаций яркости и т.п. (Медведева и др., 2017; Родионова, 2012; Borne, Viennois, 2017; Medvedeva, Evdokimova, 2019). Качество сегментации текстурными методами зависит от однородности текстурных областей, их размеров и формы.

В последнее время всё большее количество исследователей для решения задач сегментации и классификации спутниковых изображений стали использовать методы глубокого обучения, в частности свёрточные нейронные сети (СНС). СНС благодаря своей архитектуре хорошо учитывают структуру и локальные особенности цифровых изображений, что позволяет достичь высокой точности выделения целевых объектов на определённом наборе данных. Например, в работе (Захваткина и др., 2020) приведён алгоритм автоматизированной классификации морского льда Арктики по возрастным стадиям на основе метода нейронных сетей. Точность классификации для однолетнего льда составила 75 %, для старого льда — 90 %. В исследовании (Persello et al., 2019) для обнаружения границ сельскохозяйственных полей по снимкам спутника WorldView-3 используется СНС SegNet с последующей обработкой методом на основе водораздела и комбинаторной группировки. В работе (Fichtel et al., 2021) для оптимизации лесохозяйственной деятельности предлагается метод обнаружения и идентификации отдельных деревьев по снимкам, полученным с беспилотного летательного аппарата. Для реализации алгоритма выбрана СНС Faster R-CNN с архитектурой Inception v2, которая обеспечивает надёжные результаты распознавания, хорошие возможности обучения, приемлемое разрешение изображения и малое время вычислений. Точность обнаружения деревьев составляет более 75 %.

К общим недостаткам нейросетевых методов следует отнести наличие большого объёма выборок данных для обучения алгоритма, значительные вычислительные мощности и временные затраты на обучение и оптимизацию модели. Сложность решения задачи сегментации с помощью нейросетевых методов существенно возрастает с увеличением количества классов сегментируемых объектов. Это связано с усложнением настройки параметров сети и повышением требований к подготовке тренировочной базы (необходимостью учёта размерности и детальности изображений, а также равномерности распределения классов для корректной тренировки сети). В связи с этим настоящая работа посвящена разработке метода мультиклассовой сегментации целевых объектов на спутниковых снимках на основе СНС U-net и проведению исследований, направленных на повышение точности сегментации.

Методология

За последние годы появилось множество структур СНС, успешно зарекомендовавших себя в различных задачах. Для сегментации спутниковых снимков была выбрана СНС U-net (Gudzius et al., 2019; Overton, Tucker, 2020; Ronneberger et al., 2015), которая отличается высокой производительностью и обеспечивает хорошие результаты по сегментации объектов, используя небольшое количество данных.

Базовая архитектура СНС U-net состоит из сжимающего и расширяющего путей и выполняет следующую процедуру: пошагово уменьшает размерность изображения с некоторыми преобразованиями, а затем восстанавливает предсказанные маски из сжатого изображения. Для выполнения задачи мультиклассовой сегментации на первом этапе были введены следующие изменения в архитектуру сети U-net:

- для расширения формируемой карты признаков увеличена глубина СНС за счёт добавления новых слоёв сжатия и пулинга;
- изменена размерность входного окна для обработки изображений 112×112 , а также размерности выходного слоя и параметры остальных слоёв;
- для ускорения обучения СНС добавлены слои пакетной нормализации (*англ.* batch-normalization) после каждого слоя сжатия и Dropout (процедура прореживания нейронов);

- для повышения эффективности обучения и с целью избежать проблем, связанных с появлением “Dead ReLU”, в качестве функции активации использована экспоненциальная линейная единица — функция ELU. “Dead ReLU” приводит к тому, что в результате обновления весов сети нейрон никогда не активируется для любых входных значений, т.е. всегда выдаёт ноль. Предотвратить данную проблему можно с помощью функций, имеющих небольшой наклон в отрицательном диапазоне, например ELU или LReLU (Leak ReLU). В работе экспериментально была выбрана функция ELU, так как она позволила выполнить обучение в среднем на 3–8 % быстрее по сравнению с функцией LReLU.

Формула для вычисления функции ELU имеет вид (1):

$$\text{ELU}(x) = \begin{cases} x, & \text{если } x > 0, \\ \alpha(e^x - 1), & \text{если } x < 0, \end{cases} \quad (1)$$

где $\alpha = 0,1, \dots, 0,3$.

На рис. 1 показана архитектура модифицированной СНС U-net.

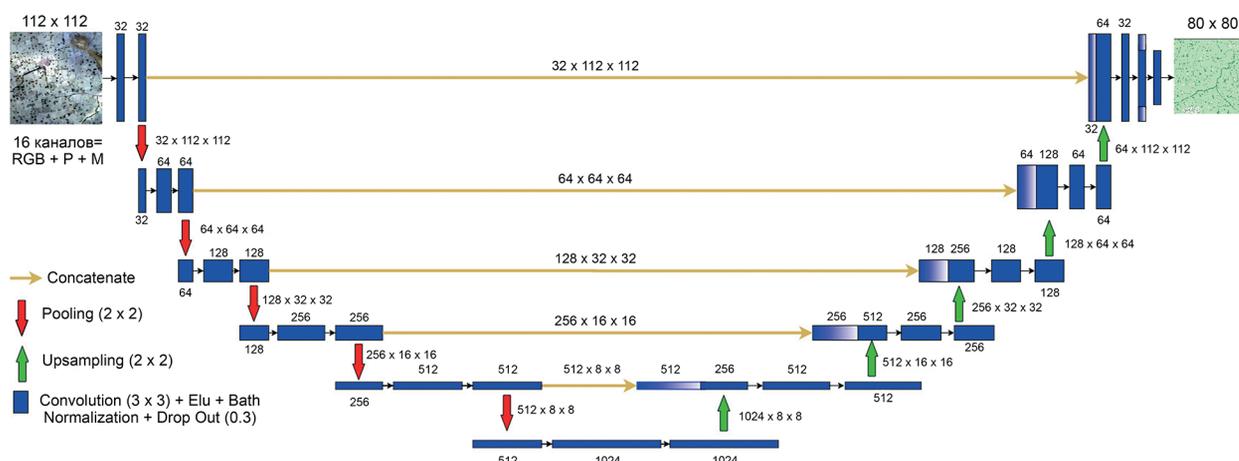


Рис. 1. Архитектура модифицированной СНС U-net

На следующем этапе выполнялась подготовка входных данных и предварительное обучение СНС. Для обучения входные снимки масштабировались к одному размеру (3472×3472 пикселей) дополнением нулями по краям изображения и с последующим заполнением их зеркальным отражением исходного изображения. Далее снимки разделялись на пересекающиеся фрагменты размером 112×112 пикселей. Для расширения тестовой выборки применялись процедуры зеркального отображения по вертикали и горизонтали, а также поворот фрагментов изображения на произвольный угол. Фрагменты изображения собирались в один пакет и подавались на вход СНС. Таким образом, обучение сети проводилось на 13 448 выборках.

Далее выполнялось тестирование модифицированной СНС и оценка результатов сегментации.

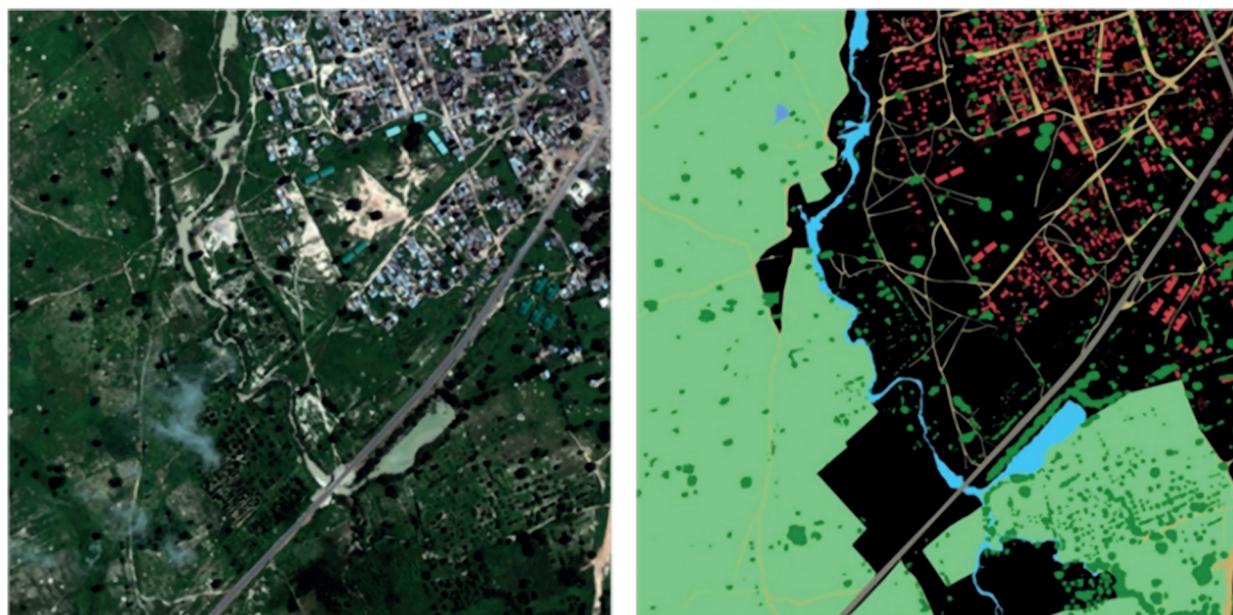
Данные

Для обучения и тестирования модифицированной СНС U-net применялись мультиспектральные снимки видимого диапазона и 16 инфракрасных спектральных диапазонов, сформированные датчиками спутника WorldView-3. Характеристики сканеров спутника WorldView-3 приведены в табл. 1.

Таблица 1. Характеристики сканеров спутника WorldView-3

Тип сканера	Спектральные каналы	Длина волн, нм	Разрешение, м	Динамический диапазон, бит/пиксель	Разрешение файлов, пиксель
Панхромный	–	450–800	0,31	11	~3396×3348
Мультиспектральный	RGB (Red)	630–690			
	RGB (Blue)	450–510			
	RGB (Green)	510–580			
	Coastal	400–450	1,24	~849×837	
	Blue	450–510			
	Green	510–580			
	Yellow	585–625			
	Red	630–690			
	Red Edge	705–745			
	Near-IR1	770–895			
	Near-IR2	860–1040			
Средний инфракрасный	SWIR-1	1195–1225	7,5	14	~136×134
	SWIR-2	1550–1590			
	SWIR-3	1640–1680			
	SWIR-4	1710–1750			
	SWIR-5	2145–2185			
	SWIR-6	2185–2225			
	SWIR-7	2235–2285			
	SWIR-8	2295–2365			

Примечание: Red — красный; Blue — синий; Green — зелёный; Coastal — фиолетовый; Red Edge — дальний инфракрасный; Near-IR — *англ.* near infrared, ближний инфракрасный; SWIR — *англ.* short wave infrared, коротковолновый инфракрасный.



а

б

Рис. 2. Пример изображения из обучающей выборки (а) и разметки (б)

Набор данных состоял из тренировочных и тестовых изображений. Тренировочный набор, используемый для подготовки сети, включал 80 различных снимков областей поверхности Земли с подготовленной разметкой. Данные снимки содержали объекты следующих классов: здания и другие искусственные сооружения, асфальтовые и грунтовые дороги, деревья, поля и сельскохозяйственные угодья, водоёмы и реки, автотранспорт. Перечисленные объекты представляют значительный интерес для различных сфер деятельности и часто становятся целевыми объектами в задачах обработки спутниковых снимков.

В качестве обучающей выборки использовались 20 изображений ($\sim 3300 \times 3300$ пикселей), размеченных на соответствующие классы. На *рис. 2* (см. с. 38) показан пример изображения из обучающей выборки (см. *рис. 2а*) и разметки (см. *рис. 2б*). Проверка работы сети осуществлялась на 60 изображениях, не используемых при тренировке.

Экспериментальные результаты

Для оценки точности сегментируемых областей использовались метрика Жаккара (Jaccard Index) и F -мера (F -score).

Метрика Жаккара оценивает сходство между сегментированным и размеченным изображениями:

$$J = \frac{TP}{TP + FP + FN}, \quad (2)$$

где TP — истинно положительное решение; FP — ложноположительное решение; FN — ложноотрицательное решение.

F -мера представляет собой гармоническое среднее между точностью (*англ.* precision) и полнотой (*англ.* recall):

$$F = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (3)$$

Обе метрики усреднялись отдельно по каждому классу объектов, а также по всему изображению.

Дополнительно для оценки результатов обучения СНС в качестве функции потерь использовались стандартная функция бинарной перекрёстной энтропии (*англ.* binary cross entropy) и стандартная метрика точности (*англ.* accuracy), позволяющая определить долю верно классифицированных изображений. Функция потерь и метрика точности вычисляются по формулам (4) и (5) соответственно:

$$H(y, p) = - \sum_{i=1}^n y_i \cdot \ln p_i, \quad (4)$$

где H — функция потерь; y — вектор меток (содержащий истинные результаты); p — вектор предсказаний нейронной сети; n — количество сегментированных классов;

$$\text{accuracy} = \frac{TP + TN}{TP + TP + FP + FN}, \quad (5)$$

где TN — истинно отрицательное решение.

Зависимости функции потерь и метрики точности от количества эпох для модифицированной сети представлены на *рис. 3* и *4* (см. с. 40).

Из графика на *рис. 3* видна сходимости нейросети к единице по точности, что свидетельствует о корректности архитектуры сети. После 50-й эпохи обучение сильно замедлилось. Это связано с ограниченным количеством тренировочных изображений. Наиболее часто встречаемые на изображениях классы хорошо определяются обучаемой сетью, что приводит к повышению общей точности сегментации и снижению корректировки весов нейронной сети в каждой итерации. При этом выделение более редких классов, таких как здания и грунтовые дороги, становится менее точным и практически не улучшается в дальнейшем.

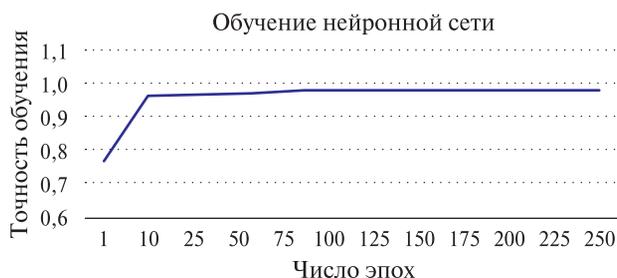


Рис. 3. Зависимость точности при обучении от количества эпох

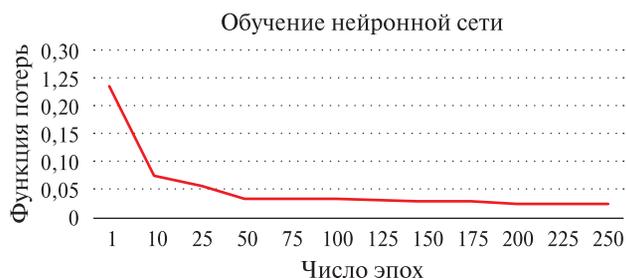


Рис. 4. График функции потерь

На рис. 5 приведён график валидации модифицированной СНС U-net, свидетельствующий о корректности работы алгоритма на тестовых изображениях.

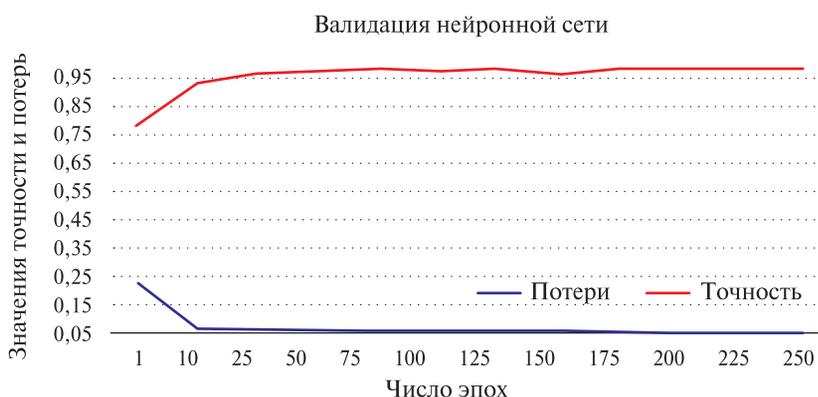


Рис. 5. Валидация сети

Точность сегментации и время обучения сети прямо пропорциональны количеству используемых для обучения эпох (проходов датасета через нейронную сеть). На эти параметры значительно влияет выбор алгоритма оптимизации и корректировки весов нейронной сети. Большинство алгоритмов обучения сети основаны на методе градиентного спуска. В настоящей работе были протестированы следующие алгоритмы: адаптивный градиентный алгоритм (англ. Adaptive Gradient Algorithm — AdaGrad), алгоритм естественного градиентного спуска (англ. Kalman-based stochastic gradient descent — kSGD), адаптивный метод оценки моментов (англ. Algorithm Development and Mining System — Adam), адаптивный метод оценки моментов с петлёй Нестерова (англ. Nesterov-accelerated Adaptive Moment Estimation — NAdam). На основании тестирования в качестве обучающего алгоритма был выбран метод NAdam, который обеспечил наибольшую точность сегментации за меньшее количество эпох (или время). Результаты тестирования методов по среднеклассовой метрике Жаккара представлены в табл. 2.

Таблица 2. Динамика значений среднеклассовой метрики Жаккара для различных тестируемых алгоритмов оптимизации

Число эпох	Метод обучения			
	AdaGrad	Adam	NAdam	kSGD
50	0,64	0,66	0,64	0,55
100	0,72	0,74	0,76	0,68
200	0,75	0,76	0,79	0,74

Важный параметр градиентных алгоритмов, влияющий на время обучения СНС, — коэффициент скорости обучения сети, который позволяет управлять величиной коррекции весов на каждой итерации. В *табл. 3* представлены результаты влияния скорости обучения сети на эффективность её работы. Низкий коэффициент скорости обучения снижает потери, но требует большего числа итераций для повышения точности и существенно замедляет процесс обучения. А высокое значение коэффициента не обеспечивает сходимость сети или приводит к её переобучению, что не даёт достигнуть высокой точности, но позволяет многократно быстрее обучать сеть. Наиболее оптимальным решением становится использование адаптивного коэффициента скорости обучения, снижающего значение в зависимости от числа пройденных эпох.

Таблица 3. Влияние коэффициента скорости обучения сети на эффективность обучения сети

Эпохи для обучения	Коэффициент скорости обучения	Потери обучения	Точность обучения	Коэффициент Жаккара
50	0,1	0,033	0,9805	0,66
	0,01	0,034	0,9711	0,62
	0,001	0,028	0,9653	0,54
100	0,1	0,032	0,9800	0,70
	0,01	0,033	0,9811	0,74
	0,001	0,027	0,9793	0,73

Модель СНС может быть расширена дополнительными слоями, что позволит получить больше карт функций, однако это технически сложно реализовывать из-за растущих требований к видеопамяти GPU. Также затруднительно увеличить размер пакета во время обучения или размеры входного слоя из-за нехватки оперативной памяти.

Точность выделения объектов и вероятность ложноположительных результатов определяются качеством обучения нейронной сети, а также схожестью набора тренировочных изображений с реальными снимками. Неравномерность распределения классов объектов в тренировочной выборке приводит к перенастройке сети для объектов часто встречаемых классов и к снижению точности выделения объектов более редких классов. Повысить точность выделения объектов редких классов при невозможности увеличения аппаратных ресурсов и нехватке в тренировочной выборке изображений с равномерным статистическим распределением классов возможно за счёт использования отдельно обученных сетей для работы с каждым конкретным классом и последующего объединения результатов их работы. Тренировка отдельной копии нейронной сети для каждого отдельного класса позволит достичь более высокой точности мультиклассовой сегментации. При таком подходе общее время обучения сети возрастает в n раз (n — число выделяемых классов), но снижается требование к равномерному распределению классов объектов на изображениях и к объёму необходимых тренировочных изображений.

На *рис. 6* (см. с. 42) приведён пример мультиклассовой сегментации фрагмента спутникового снимка: на *рис. 6а* показано исходное изображение, поданное на вход сети, на *рис. 6б* — результат сегментации изображения алгоритмом 1 и на *рис. 6в* — результат сегментации изображения алгоритмом 2. Алгоритмом 1 мультиклассовая сегментация выполнялась единой СНС, а алгоритмом 2 — с помощью отдельно обученных СНС для работы с каждым конкретным классом объектов и с последующим объединением результатов их работы.

Для проведения исследований использовался персональный компьютер со следующей конфигурацией: видеокарта NVIDIA GTX 1080Ti 8 GB; процессор Ryzen 2600x (4.0 GHz); объём RAM — 16 Gb; жёсткий диск SSD Samsung EVO 860 — 256 Gb.

В *табл. 4* приведены результаты оценки точности сегментации по коэффициентам Жаккара и F -мере двумя алгоритмами. Результаты получены при следующих условиях: оптимизатор Nadam (шаг корректировки: 0,1 — до 50 эпох, 0,01 — до 100 эпох, далее —

0,001), Dropout (доля отключаемых нейронов слоя во время каждого обновления сети при обучении — 0,3).

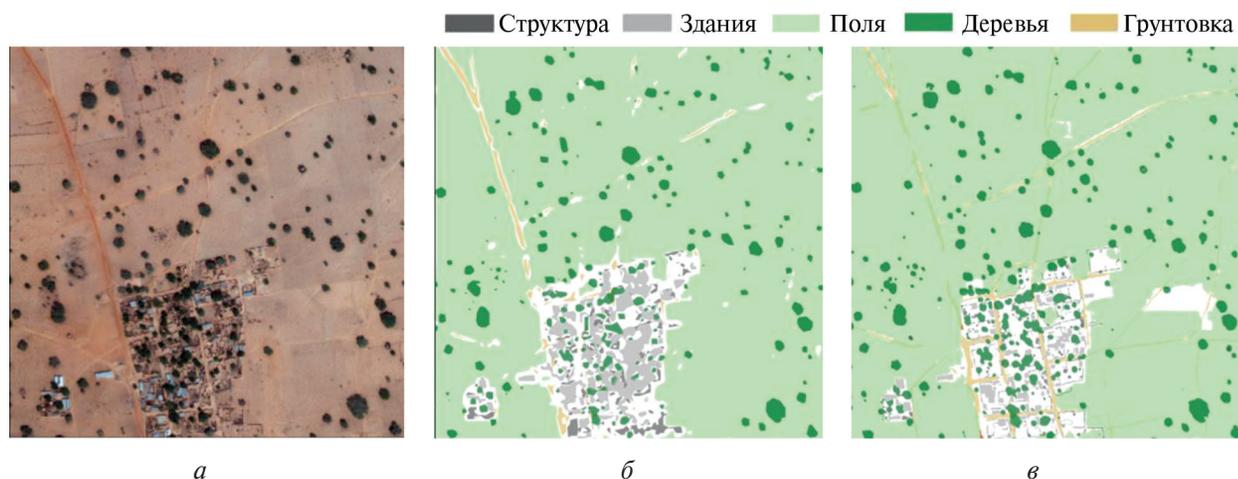


Рис. 6. Результаты сегментации спутникового снимка: а — исходное изображение; б — результат сегментации алгоритмом 1; в — результат сегментации алгоритмом 2

Таблица 4. Результаты оценки точности сегментации для различных классов

Класс	F-мера		Коэффициент Жаккара	
	Алгоритм 1	Алгоритм 2	Алгоритм 1	Алгоритм 2
Здания	0,72	0,86	0,72	0,84
Строения	0,69	0,81	0,65	0,82
Асфальтовые дороги	0,27	0,53	0,56	0,79
Грунтовые дороги	0,29	0,51	0,69	0,75
Деревья	0,89	0,94	0,97	0,95
Поля	0,92	0,93	0,92	0,96
Реки	0,79	0,82	0,76	0,89
Водоёмы	0,70	0,72	0,82	0,84
Оценка средней точности	0,66	0,77	0,76	0,86

Алгоритмом 1 хорошо определяются часто встречаемые на изображениях классы объектов (например, деревья, поля), что приводит к повышению общей точности сегментации и снижению корректировки весов СНС в каждой итерации. А выделение объектов более редких классов, таких как здания и дороги, становится менее точным и практически не улучшается в дальнейшем. Время обработки одного изображения размером до 3600×3600 пикселей не из обучающей выборки и тестового набора составило менее одной секунды. Оценка средней точности по коэффициенту Жаккара — 0,76, по коэффициенту F-мера — 0,66.

Алгоритм 2 увеличил общую точность сегментации за счёт повышения точности выделения объектов более редких классов. Наибольший выигрыш — у объектов схожих классов (грунтовые и асфальтовые дороги, здания и строения), который по коэффициенту Жаккара достигает 23 %, по F-мере — 26 %. Оценка средней точности коэффициентом Жаккара составила 0,86 %, по F-мере — 77 %. Время обработки одного изображения с учётом отдельного распознавания каждого класса и объединения результатов — около 100 с (для десяти классов объектов).

Заключение

В работе предложен метод мультиклассовой сегментации спутниковых снимков на основе модифицированной СНС U-net. Проведены исследования, направленные на повышение точности сегментации спутниковых снимков путём разработки алгоритмов на основе единой СНС и отдельно обученных СНС для работы с каждым конкретным классом. Средняя оценка точности сегментированных областей алгоритмом 1 по метрике Жаккара составила 0,76, по метрике F -мера — 66 %. Алгоритм 2 повысил точность сегментации объектов похожих и редких классов на 23 % по метрике Жаккара и на 26 % — по метрике F -мера при ограниченной тренировочной выборке изображений и вычислительных ресурсах.

Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений: пер. с англ. 3-е изд., испр. и дополн. / пер. Л. И. Рубанова, П. А. Чочиа; ред. П. А. Чочиа. М.: Техносфера, 2012. 1104 с.
2. Захваткина Н. Ю., Бычкова И. А., Смирнов В. Г. Цифровая обработка данных Sentinel-1 для автоматизированного выделения кромки старых льдов // Современные проблемы дистанционного зондирования Земли из космоса. 2020. Т. 17. № 5. С. 23–34. DOI: 10.21046/2070-7401-2020-17-5-23-34.
3. Медведева Е. В., Курбатова Е. Е., Окулова А. А. Текстульная сегментация зашумленных изображений поверхности Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2017. Т. 14. № 7. С. 20–28. DOI: 10.21046/2070-7401-2017-14-7-20-28.
4. Родионова Н. В. Текстульная сегментация одноканальных изображений: примеры применения // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 3. С. 65–69.
5. Шапиро Л., Стокман Дж. Компьютерное зрение: пер. с англ. / пер. А. А. Богуславского; ред. С. М. Соколов. М.: Бином. Лаборатория знаний, 2006. 752 с.
6. Шовенгердт Р. А. Дистанционное зондирование. Модели и методы обработки изображений: пер. с англ. / пер. А. В. Кирюшина, А. И. Демьяникова. М.: Техносфера, 2010. 594 с.
7. Borne F., Viennois G. Texture-based classification for characterizing regions on remote sensing images // J. Applied Remote Sensing. 2017. V. 11(3). DOI: 10.1117/1.JRS.11.036028.
8. Fichtel L., Fruhwald A. M., Hosch L., Schreibmann V., Bachmeir C., Bohlander F. Tree Localization and Monitoring on Autonomous Drones employing Deep Learning // Proc. 29th Conf. Open Innovations Association (FRUCT). 2021. P. 132–140. DOI: 10.23919/FRUCT52173.2021.9435549.
9. Gudzius P., Kurasova O., Filatovas E. Optimal U-Net Architecture for Object Recognition Problems in Multispectral Satellite Imagery // Proc. 2019 IEEE/ACS 16th Intern. Conf. Computer Systems and Applications (AICCSA). 2019. Art. No. 19454678. 2 p. DOI: 10.1109/AICCSA47632.2019.9035305.
10. Li M., Zang S., Zhang B., Li S., Wu C. A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information // European J. Remote Sensing. 2014. V. 47. P. 389–411. DOI: 10.5721/EuJRS20144723.
11. Medvedeva E., Evdokimova A. Detection of Texture Objects on Multichannel Images // Proc. 24th Conf. Open Innovations Association (FRUCT). 2019. P. 249–254. DOI: 10.23919/FRUCT.2019.8711986.
12. Medvedeva E. V., Kurbatova E. E. Image Segmentation Based on Two-Dimensional Markov Chains // Computer Vision in Control Systems-2. Innovations in Practice. Switzerland: Springer Intern. Publishing, 2015. P. 277–295. DOI: 10.1007/978-3-319-11430-9_11.
13. Overton T., Tucker A. DO-U-Net for Segmentation and Counting: Applications to Satellite and Medical Images // Advances in Intelligent Data Analysis XVIII. 2020. P. 391–403. DOI: 10.1007/978-3-030-44584-3_31.
14. Persello C., Tolpekin V. A., Bergado J. R., de By R. A. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping // Remote Sensing of Environment. 2019. V. 231. Art. No. 111253. DOI: 10.1016/j.rse.2019.111253.
15. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Proc. Intern. Conf. Medical Image Computing and Computer-Assisted Intervention. Springer Intern. Publishing, 2015. P. 234–241. DOI: 10.1007/978-3-319-24574-4_28.
16. Yuan J., Wang D. L. Remote sensing image segmentation by combining spectral and texture features // IEEE Trans. Geoscience and Remote Sensing. 2014. V. 52(1). P. 16–24. DOI: 10.1109/TGRS.2012.2234755.

Experimental assessment of the accuracy of multiclass segmentation of objects from satellite images based on a modified convolutional neural network U-net

S. M. Bagaev, E. V. Medvedeva

Vyatka State University, Kirov 610000, Russia
E-mails: hedgehog_rus1@mail.ru, emedv@mail.ru

A methodology for studying the accuracy of segmentation of target objects in satellite images belonging to different classes based on a modified convolutional neural network U-net is presented. The main stages of segmentation are considered: preparation of input data; modification of the structure of the convolutional neural network (CNN) taking into account the parameters of the studied images; training of the CNN on the formed training sample; segmentation of test images. Multispectral images obtained from the WorldView-3 satellite were used to train and test the modified U-net network. Experimental studies were conducted to improve the accuracy of target object segmentation. Two algorithms are used to implement the multi-class segmentation method. The first algorithm is implemented on a single multiclass CNN. The second algorithm is based on separately trained CNN to work with each specific class and then combine the results of their work. The average segmentation accuracy by the first algorithm according to the Jaccard metric was 76 %, according to the *F*-measure — 0.66. The second algorithm increased the segmentation accuracy of objects of similar and rare classes by 23 % according to the Jaccard metric and by 26 % according to the *F*-measure metric with a limited training sample of images and computational resources.

Keywords: multiclass segmentation, satellite imagery, convolutional neural network, segmentation accuracy assessment

Accepted: 09.11.2021

DOI: 10.21046/2070-7401-2021-18-6-35-45

References

1. Gonzalez R., Woods R., *Digital Image Processing: Intern. Version 3rd Edition*, Prentice Hall, 2008, 976 p.
2. Zakhvatkina N. Yu., Bychkova I. A., Smirnov V. G., Digital processing of Sentinel-1 data for automated detection of old ice edge, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2020, Vol. 17, No. 5, pp. 23–34 (in Russian), DOI: 10.21046/2070-7401-2020-17-5-23-34.
3. Medvedeva E. V., Kurbatova E. E., Okulova A. A., Texture segmentation of Earth's surface noisy images, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2017, Vol. 14, No. 7, pp. 20–38 (in Russian), DOI: 10.21046/2070-7401-2017-14-7-20-28.
4. Rodionova N. V., One channel texture based segmentation: application examples, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 9, No. 3, pp. 65–69 (in Russian).
5. Shapiro L. G., Stockman G. C., *Computer Vision*, Prentice Hall, 2001, 580 p.
6. Schowengerdt R. A., *Remote Sensing, Models and Methods for Image Processing, 3rd Edition*, Amsterdam: Elsevier; Burlington: Academic Press, 2007, 515 p.
7. Borne F., Viennois G., Texture-based classification for characterizing regions on remote sensing images, *J. Applied Remote Sensing*, 2017, Vol. 11(3), DOI: 10.1117/1.JRS.11.036028.
8. Fichtel L., Fruhwald A. M., Hosch L., Schreibmann V., Bachmeir C., Bohlander F., Tree Localization and Monitoring on Autonomous Drones employing Deep Learning, *Proc. 29th Conf. Open Innovations Association (FRUCT)*, 2021, pp. 132–140, DOI: 10.23919/FRUCT52173.2021.9435549.
9. Gudzius P., Kurasova O., Filatovas E., Optimal U-Net Architecture for Object Recognition Problems in Multispectral Satellite Imagery, *Proc. 2019 IEEE/ACS 16th Intern. Conf. Computer Systems and Applications (AICCSA)*, 2019, Art. No. 19454678, 2 p., DOI: 10.1109/AICCSA47632.2019.9035305.
10. Li M., Zang S., Zhang B., Li S., Wu C., A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information, *European J. Remote Sensing*, 2014, Vol. 47, pp. 389–411, DOI: 10.5721/EuJRS20144723.
11. Medvedeva E., Evdokimova A., Detection of Texture Objects on Multichannel Images, *Proc. 24th Conf. Open Innovations Association (FRUCT)*, 2019, pp. 249–254, DOI: 10.23919/FRUCT.2019.8711986.

12. Medvedeva E. V., Kurbatova E. E., Image Segmentation Based on Two-Dimensional Markov Chains, In: *Computer Vision in Control Systems-2, Innovations in Practice*, Switzerland: Springer Intern. Publishing, 2015, pp. 277–295, DOI: 10.1007/978-3-319-11430-9_11.
13. Overton T., Tucker A., DO-U-Net for Segmentation and Counting: Applications to Satellite and Medical Images, In: *Advances in Intelligent Data Analysis XVIII*, 2020, pp. 391–403, DOI: 10.1007/978-3-030-44584-3_31.
14. Persello C., Tolpekin V. A., Bergado J. R., de By R. A., Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping, *Remote Sensing of Environment*, 2019, Vol. 231, Art. No. 111253, DOI: 10.1016/j.rse.2019.111253.
15. Ronneberger O., Fischer P., Brox T., U-net: Convolutional networks for biomedical image segmentation, *Proc. Intern. Conf. Medical Image Computing and Computer-Assisted Intervention*, Springer Intern. Publishing, 2015, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28.
16. Yuan J., Wang D. L., Remote sensing image segmentation by combining spectral and texture features, *IEEE Trans. Geoscience and Remote Sensing*, 2014, Vol. 52(1), pp. 16–24, DOI: 10.1109/TGRS.2012.2234755.