

ГЛУБОКИЕ НЕЙРОСЕТИ ДЛЯ ВЫЧИСЛЕНИЯ ПАРАМЕТРОВ ЗДАНИЙ ПО ОДНОМОМЕНТНОМУ КОСМИЧЕСКОМУ ИЗОБРАЖЕНИЮ¹

© 2020 г. В. Ю. Игнатьев^{a,b,*}, А. Н. Трекин^{b,c,**}, П. Ю. Якубовский^{b,***}

^a ФИЦ ИУ РАН, Москва, Россия

^b Сколковский ин-т науки и технологий, Москва, Россия

^c НИИ “АЭРОКОСМОС”, Москва, Россия

*e-mail: vladimir.ignatiev.mipt@gmail.com

**e-mail: alexey.trekin@gmail.com

***e-mail: pavel.yakubovskiy2@skoltech.ru

Поступила в редакцию 06.04.2020 г.

После доработки 20.04.2020 г.

Принята к публикации 25.05.2020 г.

Высота здания – одна из основных характеристик, важных для геоаналитических сервисов, она может быть использована при оценке населения и функциональном зонировании внутри региона. Анализ высотной структуры городских территорий может быть полезен для понимания популяционной динамики. Предложен новый метод для оценки высоты зданий на основе однократного изображения дистанционного зондирования Земли, снятого не в надири. Оценка высоты производится симуляционным алгоритмом, который использует маски стены, тени и видимой части стены. Для сегментации изображения применяются глубокие сверточные нейронные сети, позволяющие извлекать маски крыш, теней и стен зданий. Модели сегментации интегрированы в полностью автоматическую систему для картирования зданий и оценки их высоты. Также представлен тестовый набор данных, который содержит размеченный набор разнообразных зданий, для оценки результата. Представленный метод исследован на собранном наборе данных и показывает среднюю абсолютную ошибку менее 4 м.

DOI: 10.31857/S0002338820050078

Введение. Задача определения физических параметров зданий по спутниковым изображениям является одной из основных задач, решаемых картографическими службами в процессе ручного дешифрирования. Данный процесс требует высокой квалификации картографов и значительного времени при обработке снимков на большие территории. При этом часть параметров, такие, как контуры объектов, площадь и высота, может быть извлечена с высокой точностью, что открывает возможность применения современных нейросетевых подходов для разработки алгоритмов распознавания объектов и автоматизации процедуры восстановления 3D-моделей объектов, таких, как здания. Высота здания является одной из важных характеристик, которые необходимо учитывать при картографировании и анализе городских территорий. По высоте и площади жилого дома можно оценить количество жителей. Кроме того, карта высоты здания может быть использована для полуавтоматического определения основания зданий на ненадирных снимках, для мониторинга хода строительства, а также для других приложений.

Первые автоматизированные алгоритмы определения высоты объекта по данным дистанционного зондирования применяли либо стереопары изображений, либо одно изображение [1]. В настоящее время эта проблема может быть решена в полностью автоматическом режиме. Пример алгоритма оценки высоты здания по данным дистанционного зондирования на основе фотogramметрии приведен в работе [2]. Решение показывает высокую точность, но требует дорогостоящих данных – стереопар спутниковых снимков субметрового разрешения. Другие подходы основываются на картировании по радарным данным или материалам воздушно-лазерного

¹ Работа выполнена при частичной финансовой поддержке Минобрнауки России (уникальный идентификатор проекта RFMEFI60719X0312).

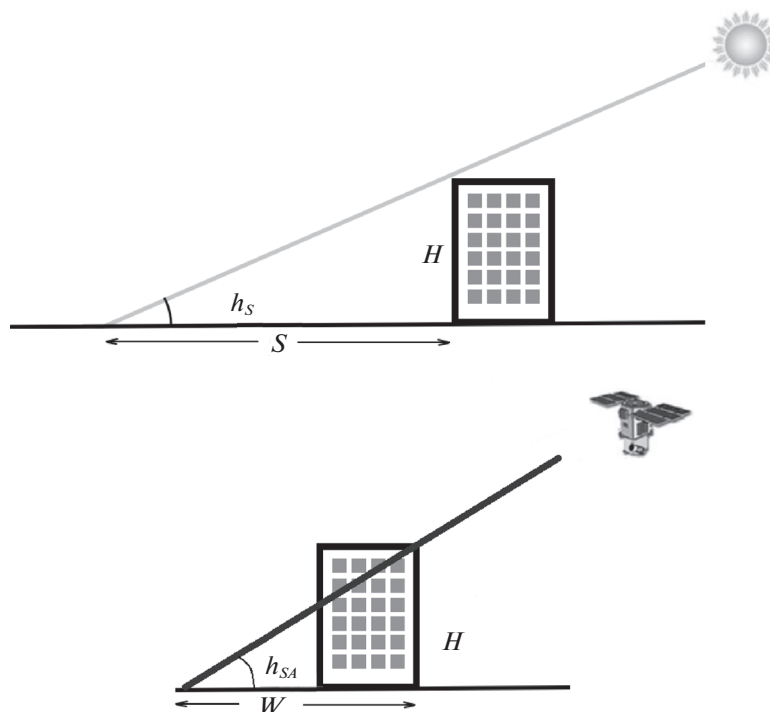


Рис. 1. Геометрия теней и стен для расчета высоты

сканирования [3]. В данной статье разрабатывается решение, которое, с одной стороны, должно показывать точность, достаточную для определения количества этажей в автоматическом режиме. С другой стороны, алгоритмы распознавания объектов на космоснимках должны быть устойчивы к изменяющейся среде при использовании относительно дешевых спутниковых данных. Для этого применяется идея классического картографического подхода, когда высота здания определяется геометрическими формами его стен, тени и крыши. Извлечение геометрии зданий и их теней осуществляется с помощью современных методов глубокого обучения для сегментации изображений.

1. Обзор предшествующих работ. В этом разделе представлен краткий обзор методов, использующих одно входное спутниковое изображение сверхвысокого пространственного разрешения для решения поставленной задачи. За скобками оставляем решения, которые применяют более сложные конфигурации входных данных, таких, как серии спутниковых изображений и стереосъемку.

Известны несколько работ, в которых информация о тенях используется для автоматической оценки высоты здания по монокулярным спутниковым изображениям. В основном задача решается путем сегментации теней и последующего расчета высоты на основе геометрии сцены.

Подход, применяющий сегментацию изображений на основе правил для оценки высот зданий, описан в [4]. Пространственное разрешение используемых данных (2.5 м на пиксель) относительно невелико, но это позволило авторам решить поставленную задачу. В результате все здания были сгруппированы по высоте в четыре категории. Недостатком этого метода является плохая устойчивость к изменяющимся условиям съемки.

В [5] предлагается подход для сегментации теней на основе “активных” контуров. Водные объекты фильтруются по порогу, поскольку характеризуются более низкой дисперсией яркости по сравнению с тенями. Мелкие объекты-артефакты удаляются с помощью морфологической обработки эрозии и дилатации. Заявленные результаты точности сегментации теней по метрике $F1 - score = 94.6$. Для оценки высоты этот метод использует среднюю длину тени в направлении солнечных лучей, как показано на рис. 1, и достигает среднеквадратичного отклонения 4.2%.

Однако есть и несколько недостатков. Алгоритм сегментации не удаляет затененные стены из полученной маски теней, что приводит к снижению точности оценки высоты.

В [6] применяется ближний инфракрасный (NIR) канал для вычисления индекса NDVI, который впоследствии используется для удаления растительности и отбрасываемых ей теней. Кроме того, для удаления мелких объектов применяется морфологическая фильтрация. Затем моделируются искусственные тени с помощью контура основания здания и данных о положении Солнца с различными высотами, перебором с заданным шагом. Разница между реальной и моделируемой теневой областью на каждом шаге увеличения высоты вычисляется с использованием меры Жаккара [7]. Расчетная высота здания соответствует моделируемой теневой области, которая доставляет максимум выражения для меры Жаккара. Тестирование алгоритма показывает среднюю абсолютную ошибку (MAE) в 0.65 м и среднеквадратичную ошибку (RMSE) в 1.22 м. Однако тестовый набор данных включает только участки одного изображения, содержащие 17 зданий, что не позволяет сделать выводы об устойчивости метода.

Полностью автоматический подход на основе глубоких сверточных нейросетей для прогнозирования карты высот непосредственно из изображений RGB предложен в [8]. Хотя результаты показывают, что сеть может отделять земную поверхность от крыш зданий, демонстрируя высокую точность построения модели рельефа по выбранной метрике, тем не менее она не позволяет найти высоты отдельных зданий.

Обзор литературы показывает, что большинство предлагаемых методов опираются на традиционные подходы к обработке изображений и используют небольшие наборы данных для тестирования качества работы алгоритмов [9–12]. В данной работе поставлена задача создать полностью автоматический метод, обеспечивающий стабильные результаты для высотных зданий, и изучить его результаты на репрезентативном наборе данных.

2. Набор данных. Были изучены различные открытые наборы данных, содержащие изображения дистанционного зондирования Земли (ДЗЗ) и разметку для семантической сегментации: спутниковые данные: SpaceNet [13], DSTL [14], аэрофотосъемка: ISPRS [15], Inria [16].

Однако все эти наборы данных оказались неприменимыми для поставленных целей, так как они содержали недостаточно многоэтажных строений, а также данных о высоте зданий и разметке стен и теней.

Подготовлен собственный набор данных, основанный на данных сверхвысокого разрешения WorldView-2/3, которые имеют пространственное разрешение 0.3 м на пиксель и три спектральных канала. Разметка зданий валидирована картографами, и основана на карте OpenStreetMap [17], высоты – из официальных данных [18].

На картографических источниках, как правило, присутствует информация о положении фундамента здания, и из нее требуется получить разметку крыш, стен и теней для конкретного изображения аналогично с конструированием *сгенерированной сцены* при оценке высоты (см. разд. 3.6).

Обучающий набор данных собран по территории 30 городов России и состоит из 51 изображения площадью около 10 кв. км каждое, что дает общее покрытие более 500 кв. км. Все изображения имеют пространственное разрешение 0.3 м и размечены на четыре класса: фундаменты, крыши, стены и тени.

Тестовый набор данных собран на основе изображения спутника WorldView-3 (г. Мытищи МО). Набор данных состоит из двух участков одного изображения общей площадью около 12 кв. км, с разметкой, которая соответствует классам в обучающем наборе данных, размеченных вручную картографами. Для разметки высот жилых зданий использовались панорамные снимки с поверхности Земли в предположении, что один этаж равен 3 м. Тестовый набор данных открыто опубликован для сравнения различных методов [19]. Угол возвышения Солнца составляет 34° , спутника – 41° .

3. Метод. Общий вид сцены представлен на рис. 1, 2, где H – высота здания, α_s и α_{SA} – азимуты на Солнце и спутник, измеренный в плоскости земли в координатах объекта, h_s и h_{SA} – углы возвышения Солнца и спутника соответственно, измеренные в вертикальных плоскостях, S – длина тени в направлении от Солнца к зданию, W – видимый размер стены, т.е. проекция стены на поверхность Земли в направлении обзора спутника.

Из данной схемы выводятся простые соотношения, позволяющие рассчитать высоту здания исходя из длины тени (3.1) или из видимого размера стены (3.2):

$$H = S \operatorname{tg} h_s, \quad (3.1)$$

$$H = W \operatorname{tg} h_{SA}. \quad (3.2)$$

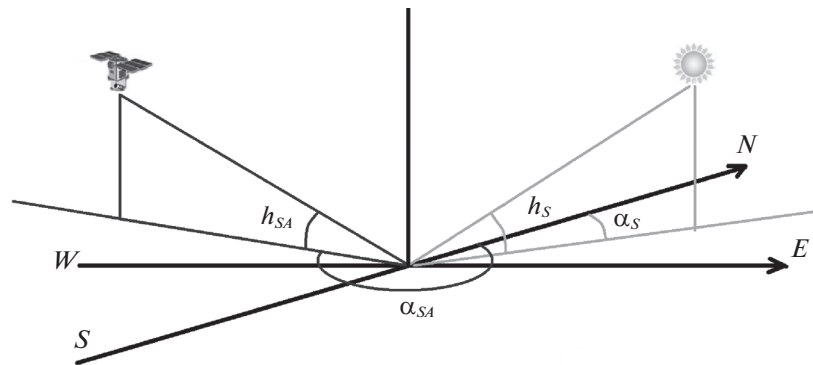


Рис. 2. Углы Солнца и спутника



Рис. 3. Блок-схема работы метода оценки высоты

Может быть использовано любое из выражений (3.1), (3.2) или усредненная величина в зависимости от ожидаемой ошибки. Если предположить, что ошибка в измерении S и W близка, то ошибки в расчете высоты H зависят от углов возвышения Солнца и спутника.

Совместное применение измерений тени и стены может увеличить точность по сравнению с использованием одного из этих параметров.

Введем несколько предположений о входных данных и изучаемых объектах. Поверхность Земли в месте расположения объекта плоская и горизонтальная. Здание имеет вертикальные стены, и его фундамент и крыша совпадают по форме и размерам. Космические изображения имеют три спектральных канала (RGB). Изображения не содержат облаков и их теней или они удалены в ходе предварительной обработки. Известны азимут и возвышение Солнца и спутника для каждого изображения. Данное ограничение можно обойти, решив обратные к (3.1) и (3.2) уравнения для зданий известной высоты, как в [20].

Общая схема предложенного метода изображена на рис. 3. Метод состоит из следующих шагов:

данные ДЗЗ подаются в нейронные сети для семантической сегментации стен и теней и сегментации крыш с разделением объектов;

растровые маски переводятся в векторную форму (полигоны) и подаются на вход модулю оценки высоты;

при оценке высоты строятся полигоны, соответствующие фундаментам зданий;

полученные векторные данные, содержащие контуры отдельных крыш и маски теней и стен, подаются в модуль оценки высоты, который описан в разд. 3.5, 3.6.

Введем два ключевых понятия разработанного метода, которые относятся к различным способам построения трехмерной модели здания – *предсказанная сцена* и *сгенерированная сцена*, соответствующие различным способам получения трехмерной модели здания. *Предсказанная сцена* состоит из крыши, стены и тени, найденные как результат сегментации изображения нейронными сетями. В *сгенерированной сцене* крыша также получена в результате сегментации изображения, но стена и тень выведены как результат генерации трехмерной модели здания при известной крыше, углах Солнца и спутника и при предположении об определенной высоте здания.

В процессе работы алгоритм оценки высоты создает сгенерированную сцену для различных предположений о высоте здания и ищет среди них наиболее подходящую к предсказанной сцене. Детальное описание всех этапов оценки высоты описано ниже.

3.1. Сверточные нейронные сети для сегментации изображений. Для семантической сегментации применены методы, основанные на глубоких нейронных сетях как наиболее эффективный метод выделения объектов на изображении. Используется широко распространенная схема “кодировщик-декодировщик” с соединениями по типу U-net [21]. Исследованы существующие архитектуры кодировщика, показавшие высокие результаты в аналогичных задачах: Densenet-169 [22], Inception-v3 [23], Inception-ResNet-v2 [24] и ResNet-50 [25], инициализированные весами, полученными при обучении на наборе данных ImageNet [26]. В [27] показано, что такая инициализация улучшает сходимость сети в случае обработки космических изображений.

Архитектура декодировщика общая для всех кодировщиков и состоит из последовательных сверточных слоев и операций transposed convolution [28] с ядром размером 4×4 чтобы избежать артефактов типа “шахматная доска”.

В качестве основной метрики для оценки результата сегментации выбран F1-score:

$$F1 - score = \frac{2pr}{p + r}, \quad (3.3)$$

где точность p и полнота r рассчитываются *попиксельно* для классов “тени зданий” и “стены зданий”, а для “крыши зданий” – *пообъектно*. При пообъектном расчете положительным результатом считается объект (крыша здания), для которого в эталонной разметке существует объект, дающий меру Жаккара $J > 0.5$.

3.2. Сегментация крыш. Как видно на рис. 1, для вычисления W и S необходимо положение фундамента здания.

Значительная часть исследований посвящена прямой сегментации фундамента здания, однако результаты четвертого раунда соревнования SpaceNet [29] показывают, что на изображениях, снятых не в нади́р, когда фундамент существенно не совпадает с крышей, точность сегментации существенно снижается. Проведенные эксперименты подтвердили эти результаты, поэтому применен метод, подобный тому, что применяют картографы при дешифрировании изображения: производится сегментация крыши здания, которая успешно выделяется независимо от угла съемки, и затем полученный контур параллельным переносом смещается в направлении основания видимой стены, как на рис. 4.

3.3. Разделение соседних зданий. Результаты предлагаемого метода опираются на правильное разделение зданий, так как вычисление высоты ведется для здания целиком, и объединение двух соседних зданий различной высоты в одно приведет к неверным результатам. Используемая в данной работе нейронная сеть для сегментации крыш имеет два выходных канала, один из которых является маской семантической сегментации, т.е. определяет принадлежность пикселя к классу “крыша”, а другой выдает расстояние от данного пикселя до границы здания. Отдельные экземпляры зданий получают с помощью метода водораздела (watershed), где расстояние до границы используется в качестве потенциала, а сегментационная маска ограничивает общую площадь объектов.

3.4. Сегментация стен и теней. Для данной задачи разделение объектов не требуется, так как отдельные здания определяются крышей, и нет необходимости разделять другие их части. Были исследованы два подхода для этой задачи: сегментация стен и теней двумя

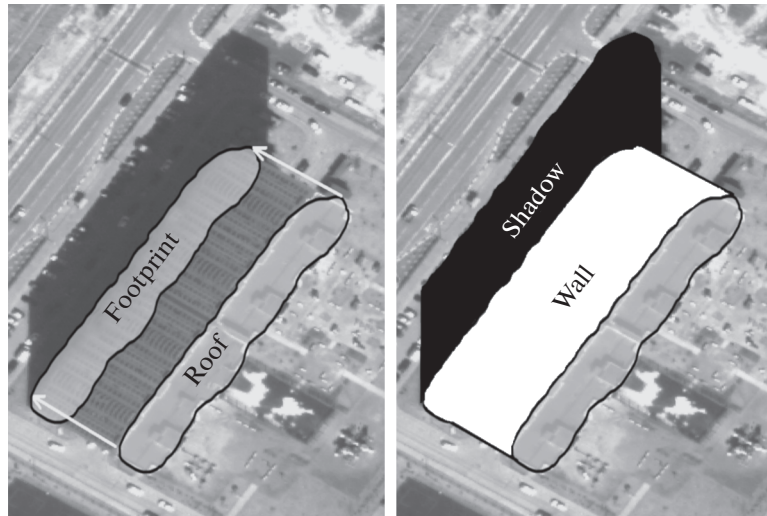


Рис. 4. Процесс генерации стены и тени

независимыми нейронными сетями и как два выхода одной нейронной сети. Второй подход показал результат, мало отличающийся от первого, при уменьшении времени вычислений вдвое (один проход нейросети вместо двух). Использована функция активации Softmax на последнем слое с множественными выходами (целевые классы плюс фон).

3.5. Оценка высоты. Для оценки высоты разработан метод, основанный на подходе, описанном в [6]. Модификация, предложенная в настоящей статье, учитывает съемку под углом и позволяет использовать не только тени, но и стены здания, а также оптимизирует поиск оптимальной высоты.

В качестве основной метрики результата рассчитывается средняя абсолютная ошибка (MAE). Поскольку конечная цель метода оценки высоты – вычисление количества этажей, то будем дополнительно рассчитывать значения $P1$ и $P2$, где $P1$ – процент зданий с абсолютной погрешностью определения количества этажей менее 1.5 м (правильное количество этажей), $P2$ – менее 4.5 м (ошибка не более чем на один этаж) соответственно.

3.6. Моделирование здания. Зная форму крыши здания и его высоту, можно построить трехмерную модель. Зная при этом углы возвышения и азимуты Солнца и спутника, можно восстановить вид здания, каким бы оно было при съемке, т.е. его проекцию на плоскость изображения. В дополнение к крыше можно смоделировать видимую часть стены и тень, падающую на поверхность Земли, чтобы получить *сгенерированную сцену* и сравнить ее с *предсказанной сценой*.

Чтобы получить позицию фундамента, необходимо сдвинуть крышу в направлении азимута спутника на $H_p \cdot \text{tg}(h_{SA})$, где H_p – предполагаемая высота здания. Видимая часть стены определяется как площадь, соединяющая крышу и фундамент.

Для построения изображения тени фундамент сдвигается в направлении от Солнца на $H_p \cdot \text{tg}(h_s)$. Тень образуется соединением сдвинутого и исходного фундаментов с последующим вычитанием из нее стены и крыши. Таким образом получается *сгенерированная сцена*, включающая в себя стену W_g и тень S_g для данной высоты H_p , результат и промежуточные стадии показаны на рис. 4.

Предсказанная сцена состоит из крыши, стены W_p , и тени S_p , являющихся результатом работы нейронных сетей. Для каждой *сгенерированной сцены* рассчитывается мера Жаккара для стен J_w и теней J_s и общая мера Жаккара J_t как их среднеарифметическое:

$$J_t = \frac{J_w + J_s}{2}, \quad (3.4)$$



Рис. 5. Модуль оценки высоты

$$J_w = \frac{W_p \cap W_g}{W_p \cup W_g}, \quad (3.5)$$

$$J_s = \frac{S_p \cap S_g}{S_p \cup S_g}. \quad (3.6)$$

В уравнении (3.7) оптимизируется общая мера Жаккара – J_t , а также J_s и J_w по высоте H_p методом градиентного спуска со следующими параметрами: начальная высота составляет 0 м, максимальная высота – 100 м, начальный шаг по высоте – 10 м, минимальный шаг по высоте равен 0.5 м:

$$H_p^{opt} = \arg \max_{H_p} J_t. \quad (3.7)$$

Архитектура модуля оценки высоты показана на рис. 5. Входными параметрами модуля являются: результаты сегментации крыш, стен, теней, а также углы возвышения и азимуты Солнца и спутника.

4. Эксперименты. В рамках данной работы были проведены следующие эксперименты: выбор архитектуры нейронных сетей для сегментации крыш, стен и теней; оценка возможностей метода определения высоты с использованием эталонной (экспертной) разметки изображений; определение возможности непосредственной сегментации основания здания; сравнение результатов при применении методики разделения соседних зданий и без; выбор метода оценки высоты: с учетом стен, теней или стен и теней совместно.

4.1. В ы ч и с л и т е л ь н о е о б о р у д о в а н и е. Все вычислительные эксперименты проведены на ПК со следующими характеристиками:

графический ускоритель: 2 × Nvidia GeForce GTX 1080Ti, память 12 Гб;

центральный процессор: Intel Core i5-8600K;

оперативная память: 64 Гб.

Таблица 1. Сравнение точности сегментации крыш для различных архитектур нейронных сетей

Архитектура	F1 “крыши”
U-Net (baseline)	0.7476
DenseNet-169	0.8090
Inceptionv3	0.8094
Inception-ResNetv2	0.8110
Inception-ResNetv2 (no pretrain)	0.8044

Таблица 2. Сравнение производительности различных архитектур нейросетей

Архитектура	F1 “тени”	F1 “стены”	Параметры, млн
DenseNet-169	0.8113	0.8653	15.5
Inceptionv3	0.8101	0.8545	25.2
Inception-ResNetv2	0.8149	0.8537	57.3
DenseNet-169	0.8073	0.8606	15.5

4.2. Особенности процесса обучения нейронных сетей. Нейронная сеть для классов “стены” и “тени” исследована в двух вариантах: в виде отдельных моделей для каждого класса и как общая модель с двухканальным выходом.

Обучение нейронных сетей проводилось с помощью библиотеки Keras [30], при этом использованы следующие параметры:

набор данных разбивался случайным образом на изображения 512×512 пикселей;

на каждом шаге обучения в нейросеть подавалось шесть изображений;

каждая эпоха состояла из 500 шагов обучения и 300 шагов валидации, общее количество эпох – 100;

оптимизатор Adam [31], параметр скорости обучения (learning rate), был установлен как 10^{-4} в течение первых пяти эпох, 10^{-5} – в течение следующих 35 эпох, 10^{-6} – далее;

первые пять эпох веса кодировщика были зафиксированы;

аугментации набора данных применялась в случайной комбинации: сдвиг, поворот в пределах 90° градусов, отражение по вертикали и горизонтали, увеличение и уменьшение в пределах 20%, растяжение; метрика: F1 – score.

Точность сегментации, указанная ниже, определялась на выделенном валидационном поднаборе данных.

4.3. Сегментация крыш. Эксперименты по сравнению точности сегментации класса “крыши” для различных архитектур нейронных сетей представлены в табл. 1. Видно, что наилучшую точность показывает модель с архитектурой Inception-ResNetv2. Кроме того, показано, что предварительное обучение на базе ImageNet [26] позволяет достичь более высокой точности по сравнению со случайной инициализацией весов.

4.4. Сегментация стен и теней. Численные результаты по точности сегментации для различных архитектур моделей приведены в табл. 2. Из этого эксперимента видно, что хотя наилучшие показатели при сегментации стен и теней зданий демонстрируются различными моделями, разница (менее 0.01) при этом оказывается незначительной для практических целей. Таким образом, наиболее перспективна, с точки зрения соотношения точности и производительности, двухканальная модель DenseNet-169, выходом которой одновременно являются маски стен и теней.

4.5. Ограничения алгоритма определения высоты зданий. Точность разработанного алгоритма определения высоты зданий оценивается при сравнении с эталонной разметкой изображений (см. разд. 2), чтобы определить ограничения метода, оставляя за скобками автоматическую сегментацию по классам “тени зданий”, “стены зданий” и “крыши зданий”. Для проверки концептуальной применимости разработанного алгоритма оценки высот зданий организуем эксперимент следующим образом. Запустим вычисление высот зданий на данных эталонной разметки геометрии зданий, полученной при ручном дешифрировании

Таблица 3. Результаты эксперимента с разделением отдельных зданий и семантической сегментацией (N – количество зданий, подходящих для оценки высоты)

Метод	MAE	P1	F1	N
Семантическая сегментация	3.47	0.34	0.75	647
Отдельные здания	3.45	0.34	0.69	696

Таблица 4. Результаты экспериментов по расчету высоты с использованием стен и теней отдельно

Класс объектов	MAE	P1	P2	N
Стены и тени	3.81	0.33	0.72	512
Только стены	5.52	0.19	0.55	273
Только тени	3.45	0.32	0.69	696

картографом стен, теней и крыш зданий на снимке. При этом сравним оптимальные значения высоты H_p , найденные переборным методом и градиентным спуском для (3.7). Для полученных оценок значений высоты вычисляются метрики MAE, P1, P2 на тестовом наборе без учета зданий с нулевой высотой.

4.6. **Непосредственная сегментация фундамента.** Архитектура нейросети, показывающая лучший результат для сегментации крыш, была обучена непосредственно для сегментации фундаментов зданий.

Полученные результаты оказались значительно хуже: средняя мера Жаккара уменьшилась с 77.64 при сегментации крыш до 68.51, причем чем выше здание, тем больше ошибка из-за большего смещения фундамента. Таким образом, применение сегментации крыши и сдвига получается оправдано для оценки высоты.

4.7. **Разделение отдельных зданий.** Использование техники разделения отдельных зданий по сравнению с применением только семантической сегментации дает вклад в первую очередь в количестве зданий, которые корректно сегментированы; однако есть влияние и на среднюю точность определения высоты (см. табл. 3).

4.8. **Использование стен и теней при оценке высоты.** В данном эксперименте сравниваются результаты определения высот при оптимизации J_w (только стены), J_s (только тени) и J_t (тени и стены).

Результаты эксперимента в табл. 4 показывают, что применение только теней позволяет получать лучшие результаты на тестовых данных.

В данном эксперименте N – число зданий, для которых метод градиентного спуска сошелся к ненулевой высоте, из тех зданий, которые верно сегментированы.

5. Результаты и обсуждение. Полученные в данном исследовании результаты показывают, что точность сегментации теней зданий оказывается ниже, чем в некоторых работах, в частности [5]. Предполагается, что возможно достижение более высокой точности при определении высоты зданий по разработанному методу, если полученные результаты по сегментации крыши будут объединены с другими методами обнаружения теней, хотя при этом может снизиться степень автоматизации подхода.

5.1. **Ограничения алгоритма определения высоты зданий.** Результаты, представленные в табл. 5, показывают неожиданно высокий процент ошибок, когда на вход алгоритму определения высоты подаются геометрии здания, полученные из ручной разметки картографом. Детальное исследование причин возникновения ошибок показывает, что большая их часть вызвана естественными ограничениями предложенного алгоритма.

Например, метод не работает, если стены и тени полностью скрыты деревьями (рис. 6, а) или если дом является частью здания, состоящего из нескольких секций различной высоты, и стены не видны (рис. 6, в); также высота здания с более широким основанием может быть недооценена, если основание ошибочно принимается за уровень Земли (рис. 6, б).

Оказывается, что с помощью созданного тестового набора данных и предложенного метода, использующего видимые области стены и тени здания на изображении, достигается MAE около 4 м и определяются около 20% зданий, для которых метод не дает соответствующих результатов.

Таблица 5. Тестирование алгоритма на данных эталонной разметки (N – количество зданий, для которых оценивается высота)

Метод	MAE	P1	P2	N
Градиентный спуск	6.65	0.40	0.67	1207
Перебор	4.2	0.45	0.76	530
Градиентный спуск (без зданий с “нулевой” высотой)	3.45	0.32	0.69	696
Перебор (без зданий с “нулевой” высотой)	3.96	0.21	0.52	495

Эти значения необходимо учитывать при итоговой оценке качества работы метода определения высоты.

5.2. Применение стен и теней. При разработке метода предполагалось, что совместное использование теней и стен позволит получить более высокую точность определения высоты, чем применение одной из характеристик отдельно, однако в проведенном эксперименте использование только теней показало более высокие результаты. Однако из (3.1), (3.2) следует, что ошибка в определении высоты, связанная с ошибками сегментации, пропорциональна тангенсу угла, поэтому если угол возвышения спутника h_{SA} существенно меньше, чем угол Солнца h_S , то ошибка при определении высоты стены будет существенно выше, чем тени, и наоборот. В данном случае разница в точности может объясняться разницей углов возвышения Солнца и спутника, вследствие чего стены дают больший вклад в ошибку.

5.3. Сравнение результатов с другими работами. Оценка высоты здания в соответствии с процедурой, описанной в разд. 3.6, и оценка точности метода проводятся по всему тестовому набору данных. Сравнение разработанного алгоритма с лучшими алгоритмами, часть из которых реализована в коммерческом программном обеспечении (ПО), и результатами, опубликованными в других работах, дано в табл. 6.

Полуавтоматический метод, реализованный в ПО Photomod, требует задания определенного набора входных метаметров обработки, например параметры гауссовского фильтра. Кроме того, увеличение стоимости данного ПО в сравнении с предлагаемым методом обусловлено необходимостью получения стереопар высокого разрешения для обработки. Все методы, основанные на ручном экспертном дешифрировании аэрокосмических изображений, имеют более высокую точность. Однако их недостатками являются высокая трудоемкость, низкая скорость работы, что неизбежно приводит к появлению ошибок при обработке очень больших площадей. Разработанный метод имеет среднюю погрешность около 4 м, что сопоставимо с полуавтоматическим режимом, предложенным в ПО english Photomod, но предлагаемый подход отличается более высокой скоростью обработки по сравнению с другими методами. Применяя полученный метод, территория площадью 80 кв. км, содержащая около 7500 зданий, обрабатывается примерно за 25 мин с помощью персонального компьютера (ПК) с характеристиками, указанными в разд. 4.1.

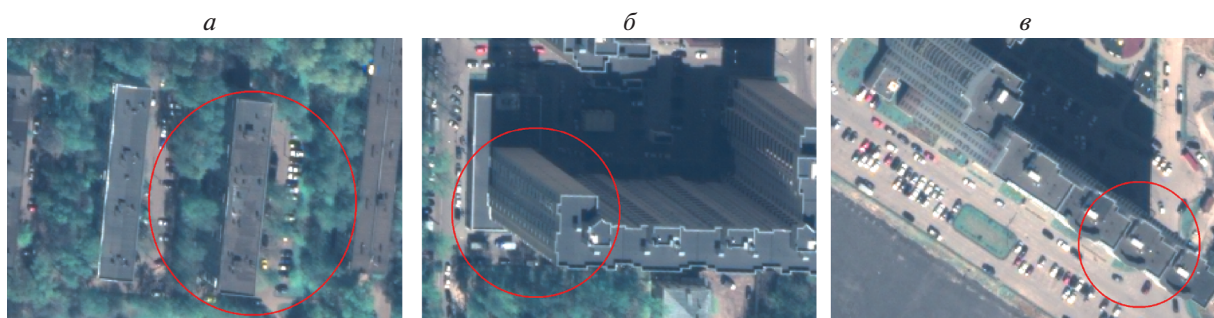


Рис. 6. Типичные источники ошибок, где алгоритм не работает по построению: *a* – стены и тени закрыты деревьями; *b* – несколько первых этажей здания имеют большую площадь, чем остальные; *в* – секции здания имеют разную высоту, стена одной из секций не видна.

Таблица 6. Сравнение разработанного метода определения высоты по спутниковым изображениям с аналогами

Статья	MAE	P1	P2	Автоматизация
[32]	6.03	—	—	—
[20]	0.82	—	—	Отсутствует
[6]	0.65	—	—	Частичная*
[5]	4.13%**	—	—	Частичная***
Предложенный метод	3.45	0.32	0.69	Полная
Предложенный метод и ручная верификация картографом	3.63	0.27	0.71	Отсутствует

*Требуется участие оператора для работы алгоритма сегментации основания зданий.

**Отсутствуют данные о средней ошибке в метрах и средней высоте зданий, рассматриваемых в работе.

***Определение положения и геометрии зданий не автоматизировано.

Достижимая точность определения высоты зданий в 4 м (1 этаж) оказывается достаточной для многих приложений: оценка плотности населения (ошибка в один этаж по высоте дает около 3–10% ошибки по плотности населения), автоматическое картирование зданий для выбранного региона. Тем не менее, представленный алгоритм имеет ряд недостатков, которые должны быть оптимизированы или приняты во внимание при использовании в приложениях:

точность ограничена условиями съемки и может быть снижена из-за пересечением теней от нескольких зданий и наложением объектов друг на друга на изображении;

алгоритм не может оценить высоту зданий в 1–2 этажа. Однако это не важно для поставленной задачи, так как в данном исследовании любое низкое и маленькое здание рассматривается как частный дом;

алгоритм не работает должным образом в районах с плотными зданиями, где перекрытие объектов одного или разных классов составляет более 50%.

Заключение. Разработан автоматический метод оценки высоты зданий по данным ДЗЗ. Метод использует современные нейросетевые модели для сегментации изображений и оценивает высоту каждого обнаруженного здания с использованием его геометрических параметров и трехмерного моделирования сцены.

В данной работе существующие методы расширены следующим вкладом:

- 1) применение современных методов сегментации изображений, которые показали достаточно высокий результат в задаче сегментации отдельных зданий;
- 2) сбор большого набора данных для обучения и тестирования моделей по изображениям территории России;
- 3) поиск высоты методом градиентного спуска вместо полного перебора, что значительно сокращает затраты времени;
- 4) реализация полностью автоматического метода, который принимает на вход изображения и их метаданные и генерирует карту зданий с рассчитанными высотами зданий.

Эксперименты показали, что результат существенно зависит от выбора тестовой области и собранный набор данных представляет собой сложный объект для методов определения высот зданий.

Результирующий алгоритм позволяет в полностью автоматическом режиме получать оценку высоты со средней ошибкой порядка одного этажа здания. Наилучшие результаты получены в областях редкой высокой застройки, которые типичны для территории России. В дальнейшей работе необходимо расширить возможности метода по работе в областях плотной исторической застройки и частных домов в сельской местности.

СПИСОК ЛИТЕРАТУРЫ

1. *Irvin R.B., McKeown D.M.* Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery // IEEE Trans. on Systems, Man, and Cybernetics: IEEE. 1989. V. 19. № 6. P. 1564–1575.

2. *Ahmad A., Rabiul L.* Generation of Three Dimensional Model of Building Using Photogrammetric Technique // Proc. 2011 IEEE 7th Intern. Colloquium on Signal Processing and its Applications. Penang, Malaysia: IEEE, 2011. P. 225–231.
<https://doi.org/10.1109/cspa.2011.5759877>.
3. *Baltsavias E.P.* A Comparison Between Photogrammetry and Laser Scanning // ISPRS J. Photogrammetry and Remote Sensing: Elsevier, 1999. V. 54. № 2–3. P. 83–94.
[https://doi.org/10.1016/s0924-2716\(99\)00014-3](https://doi.org/10.1016/s0924-2716(99)00014-3)
4. *Comber A., Umezaki M., Zhou R. et al.* Using Shadows in High-resolution Imagery to Determine Building Height // Remote Sensing Letters: Taylor&Francis. 2012. V. 3. № 7. P. 551–556.
<https://doi.org/10.1080/01431161.2011.635161>
5. *Liasis G., Stavrou S.* Satellite Images Analysis for Shadow Detection and Building Height Estimation // ISPRS J. Photogrammetry and Remote Sensing: Elsevier. 2016. V. 119. № 2–3. P. 437–450.
<https://doi.org/10.1016/j.isprsjprs.2016.07.006>
6. *Kadhim N., Mourshed M.* A Shadow-overlapping Algorithm for Estimating Building Heights from VHR Satellite Images // IEEE Geoscience and Remote Sensing Letters: IEEE. 2017. V. 15. № 1. P. 8–12.
<https://doi.org/10.1109/lgrs.2017.2762424>
7. *Jaccard P.* Etude Comparative de la Distribution Florale Dans Une Portion des Alpes et du Jura // Bulletin de la Société Vaudoise des Sciences Naturelles. 1901. V. 37. Iss. 142. P. 547–579.
<https://doi.org/10.5169/seals-266450>
8. *Mou L., Zhu X.X.* IM2HEIGHT: Height Estimation from Single Monocular Imagery via Fully Residual Convolutional-deconvolutional Network // URL: <https://arxiv.org/pdf/1802.10249.pdf>.
9. *Себряков Г.Г., Сошников В.Н., Кикин И.С., Ишутин А.А.* Алгоритм распознавания объектов на оптико-электронных изображениях наземных сцен, основанный на оценке ковариационных матриц градиентных функций поля яркости // Вестн. компьютерных и информационных технологий. 2013. Т. 109. № 7. С. 14–19.
10. *Визильтер Ю.В., Желтов С.Ю.* Использование проективных морфологий в задачах обнаружения и идентификации объектов на изображениях // Изв. РАН. ТиСУ. 2009. № 2. С. 125–138.
11. *Novikov G., Trekin A., Potapov G. et al.* Satellite Imagery Analysis for Operational Damage Assessment in Emergency Situations // Lecture Notes in Business Information Processing. 2018. V. 320. P. 347–358.
https://doi.org/10.1007/978-3-319-93931-5_25
12. *Ignatiev V., Trekin A., Lobachev V. et al.* Targeted Change Detection in Remote Sensing Images // Proc. SPIE 11041, Eleventh Int. Conf. on Machine Vision (ICMV 2018). Munich, Germany. 2018. V. 11041. doi.org/
<https://doi.org/10.1117/12.2523141>.
13. *Van E.A., Lindenbaum D., Bacastow T.M.* Spacenet: A Remote Sensing Dataset and Challenge Series // URL: <https://arxiv.org/pdf/1807.01232.pdf>.
14. *Kaggle Inc.* DSTL Satellite Imagery Feature Detection Challenge // URL: <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection/>.
15. *Rottensteiner F., Sohn G., Gerke M. et al.* ISPRS Semantic Labeling Contest // URL: <http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html>.
16. *Maggiori E., Tarabalka Y., Charpiat G. et al.* Can Semantic Labeling Methods Generalize to any city? The Inria Aerial Image Labeling Benchmark // Proc. 2017 IEEE Intern. Geoscience and Remote Sensing Sympos. (IGARSS). Fort Worth, Texas, USA: IEEE. 2017. P. 3226–3229.
<https://doi.org/10.1007/978-3-319-24574-4>
17. *OpenStreetMap Contributors.* Planet Dump Retrieved from <https://planet.osm.org> // URL: <https://planet.openstreetmap.org/>.
18. *Государственная корпорация – Фонд содействия реформированию жилищно-коммунального хозяйства* Открытые данные. Реформа ЖКХ // URL: <https://www.reformagkh.ru/opendata>.
19. *Aeronetlab, Skoltech* Open Dataset for Building Height Estimation // URL: <https://github.com/aeronetlab/open-datasets>.
20. *Qi F., Zhai J.Z., Dang G.* Building Height Estimation Using Google Earth // Energy and Buildings: Elsevier. 2016. V. 118. 1. P. 123–132.
<https://doi.org/10.1016/j.enbuild.2016.02.044>
21. *Ronneberger O., Fischer Ph., Brox Th.* U-net: Convolutional Networks for Biomedical Image Segmentation // Proc. 2015 Int. Conf. on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015. P. 234–241.
https://doi.org/10.1007/978-3-319-24574-4_28.
22. *Huang G., Liu Z., Van Der M.* Densely Connected Convolutional Networks // Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017. P. 4700–4708.
<https://doi.org/10.1109/cvpr.2017.243>.

23. Szegedy C., Vanhoucke V., Ioffe S. Rethinking the Inception Architecture for Computer Vision // Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. P. 2818–2826. <https://doi.org/10.1109/cvpr.2016.308>.
24. Szegedy C., Ioffe S., Vanhoucke V. Inception-v4, Inception-resnet and the Impact of Residual Connections on Learning // Proc. 2017 AAAI Conf. on Artificial Intelligence. San Francisco, California, USA: AAAI Press., 2017. P. 2818–2826. <https://doi.org/10.1109/cvpr.2016.308>.
25. He K., Zhang X., Ren Sh. et al. Deep Residual Learning for Image Recognition // Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016. P. 770–778. <https://doi.org/10.1109/cvpr.2016.90>.
26. Deng J., Dong W., Socher R. et al. Imagenet: A Large-scale Hierarchical Image Database // Proc. 2009 IEEE Conf. Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009. P. 248–255. <https://doi.org/10.1109/cvpr.2009.5206848>.
27. Iglovikov V., Shvets A. Terausnet: U-net with VGG11 Encoder Pre-trained on Imagenet for Image Segmentation // URL: <https://arxiv.org/pdf/1801.05746.pdf>.
28. Odena A., Dumoulin V., Olah C. Deconvolution and Checkerboard Artifacts // Distill. 2016. V. 1. Iss. 10. <https://arxiv.org/pdf/1802.10249.pdf>. <https://doi.org/10.23915/distill.00003>.
29. Weir N., Lindenbaum D., Bastidas A. SpaceNet MVOI: a Multi-View Overhead Imagery Dataset // URL: <https://arxiv.org/pdf/1903.12239.pdf>.
30. Chollet F. Keras // URL: <https://keras.io>.
31. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization // URL: <https://arxiv.org/pdf/1412.6980.pdf>.
32. Licciardi G.A., Villa A., Dalla M. et al. Retrieval of the Height of Buildings from WorldView-2 Multi-angular Imagery Using Attribute Filters and Geometric Invariant Moments // IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing: IEEE. 2012. V. 5. 1. P. 71–79. <https://doi.org/10.1109/jstars.2012.2184269>