Санкт-Петербургский государственный университет

Кафедра системного программирования

Группа 21.Б07-мм

Фильтрация эффекта от облаков на спутниковых и аэроснимках

Урбанский Денис Андреевич

Отчёт по учебной практике в форме «Сравнение»

Научный руководитель:
старший преподаватель кафедры системного программирования, М. Н. Смирнов

Санкт-Петербург
2024

Оглавление

Введение **3**

1. Постановка задачи **5**

1.1. Цель 5

1.2. Задачи 5

2. Обзор **6**

2.1. TSMM 6

2.2. LANA 7

2.3. DeepMask 9

2.4. AFMUNet 11

2.5. Сравнение 12

3. Эксперимент **15**

3.1. Датасеты 15

3.2. Метрики 15

Заключение **16**

Список литературы **17**

Введение

В настоящее время спутниковая съемка применяется с самыми различными целями. Полученные со спутника фотографии используются для картографирования или гидрологических исследований (например, тенденции изменения русла реки, прогнозирования паводка или оценки ущерба от него). Спутниковую съемку используют вулканологи при изучении извержений, климатологи, океанологи и ряд других специалистов. Спутниковые снимки могут быть полезны в качестве опорных снимков для прокладки маршрутов летательных аппаратов.

Фотографии, полученные со спутников, могут быть использованы не только для решения глобальных задач. Региональные задачи, которые могут быть решены при помощи спутниковой съемки, не менее разнообразны. Также, постоянный мониторинг местности помогает выявить места незаконных свалок [17], вырубок леса или разработок полезных ископаемых. По динамике изменений снимков можно определить прогресс строительства, состояние дорожной транспортной сети, оценить загрязнение окружающей среды лесными и степными пожарами [18].

Вместе с тем имеются существенные ограничения на возможности использования полученных со спутника фотографий в качестве источника данных о планете из-за атмосферы: облака покрывают свыше половины поверхности Земли [10], что затрудняет спутниковое наблюдение за изменениями в ландшафте, загрязняя около двух третей снимков [3], особенно в регионах низких широт, отличающихся большей облачностью.

Оценка качества предлагаемых снимков позволила бы, с одной стороны, производить более качественный отбор для вторичной обработки (выявления свалок, вырубок, паводков и т. д.), а с другой - сократить расходы заказчика снимков на подрядчиков (при покупке снимков автоматически определять, в какой мере их качество удовлетворяет условиям контракта).

В связи с этим весьма актуальной представляется разработка приложения для обнаружения облаков и теней от них с возможностью рас-

ширения функциональности до удаления эффектов от облаков путем их замены на чистые участки с других снимков.

1. Постановка задачи

1.1. Цель

Целью является реализация инструмента для обнаружения эффектов от облаков на спутниковых снимках.

Для ее достижения были поставлены следующие задачи:

1.2. Задачи

1. Анализировать существующие методы обнаружения облаков и их теней.
2. Разработать алгоритм обнаружения облаков и их теней.
3. Реализовать инструмент для обнаружения эффектов от облаков
4. Оценить эффективность полученного метода.
5. Анализировать методы ускорения и увеличения точности фильтрации.

2. Обзор

2.1. TSMM

TSMM [14] (Time series maximum and minimum mask method — метод максимальной и минимальной маски временного ряда) работает с изображениями временных рядов и использует процессор оценки качества CloudScore+ [6].

Следует данному алгоритму:

1. Исходное изображение умножается на соответствующую ему маску облаков и теней, являющуюся результатом работы CloudScore+.

2. Вычисляется σ - соотношение первого и второго максимальных изображений синего диапазона. Если σ больше некоторого значения, которое авторы рекомендуют брать равным 1,2, то соответствующие пиксели на максимуме — облачные. Аналогично с двумя минимальными из около-инфракрасного диапазона: σ - отношение второго к первому минимальному изображению, и если σ больше 1,2, то соответствующие пиксели минимума — теневые.

3. Облачные и теневые пиксели максимума и минимума заменяются на пиксели с теми же координатами на втором наибольшем изображении синего диапазона и втором наименьшем изображении около-инфракрасного диапазона соответственно.

4. Полученные маски максимума и минимума сравниваются с синей и около-инфракрасной составляющей интересующего изображения: пиксели синего диапазона, большие соответствующих пикселей маски максимума, являются облаками; пиксели около-инфракрасного диапазона, меньшие соответствующих пикселей маски минимума, являются тенями от облаков. В результате этого шага получаются предварительные маски облаков и теней.

5. Поскольку возможна ложная классификация пикселей, проводится свертка: вычисляется среднее значение μ ядра свертки — соседей ложно классифицированного пикселя на предварительной маске облаков или теней. Если μ больше или равно некоторому значению, которое авторы рекомендуют брать равным 0,3, то данный пиксель принадлежит облаку или тени. При наличии пикселя и на маске облака, и на маске тени, он остается на маске облака и удаляется с маски тени, что влечет за собой результат алгоритма.

Точность определения облаков и их теней увеличивается с увеличением размера временного ряда (60 и больше) и при больших значениях σ (2 и 1,4, соответственно). Однако точность определения незакрытых пикселей выше в обратном случае — при размере ряда, равном 10, и $\sigma = 1$. Чувствительность так же выше при малых значениях: при размере ряда, равном 10 и 15 для облаков и теней соответственно, и $\sigma = 1$. Для незакрытых пикселей — размер ряда, равный 60 и больше, и $\sigma = 2$

На датасете S2ccs (датасет производства авторов) TSMM по сравнению с CloudScore+ показал увеличение общей точности по определению эффекта от облаков на 3% и F-меры — на 0,09 (общая точность составила 93%, F-мера — 0,85), а на датасете CloudSEN12 [2] имел такую же общую точность, что и CloudScore+ (равную 87%), и F-меру, большую на 0,04 (и составляющую 0,66).

2.2. LANA

LANA [15] (Learning Attention Network Algorithm — алгоритм обучающейся сети внимания) — U-сеть[1], подразумевавшаяся авторами как более точный аналог Fmask [16], разработанного для Landsat 8 и Landsat 9. Работает с 8-полосными участками снимков размером 512 x 512 пикселей и обрабатывает 64 участка за итерацию.

Кодировщик состоит из четырех сверточных блоков, в каждом из которых происходит следующее:

[1]U-net — CNN (сверточная нейронная сеть) с U-образной структурой сети, имеющей кодировщик, узкое место и дешифратор.

1. Две свертки с ядрами 3 x 3 пикселя.

2. Нормализация серии — значение каждого нейрона нормализуется по среднеквадратичному отклонению.

После каждого блока происходит макс-пулинг с окном 2 x 2, в результате которого размер карт признаков уменьшается в 2 раза и их количество увеличивается в 2 раза.

Узкое место состоит из такого же сверточного блока, что и кодировщик.

Аналогично кодировщику, дешифратор состоит из четырех сверточных блоков, но перед каждым из которых находится транспонированный сверточный слой, добавляющий строку из нулей после каждой строки и столбец из нулей после каждого столбца и производящий свертку с ядром 2 x 2.

В случае получения отрицательных значений они приравниваются к нулю.

Для предотвращения потери информации после работы кодировщика между каждыми двумя блоками кодировщика и дешифратора, работающими с матрицами одинакового размера, проложены пропускные соединения: копия результата работы блока кодировщика передаётся соответствующему блоку дешифратора, но предварительно проходит через механизм внимания. Он, в свою очередь, трансформирует каждый признак во взвешенную комбинацию всех признаков.

В результате оптимизации для параметров были выбраны следующие значения: коэффициент скорости обучения — 0,0005, метода снижения скорости обучения — косинусный распад (cosine decay [12]; согласно данному методу, коэффициент скорости обучения сначала равен нулю и линейно увеличивается до указанного значения, после чего уменьшается до нуля по косинусоиде), алгоритм оптимизатора обучения — RMSProp [11]; к последнему сверточному блоку и всем блокам дешифратора перед транспонированным слоем применяется пространственное выпадение [9] (пространственное выпадение отбрасывает случайные карты признаков).

По сравнению с Fmask наблюдается увеличение общей точности в 2,93% (сама общая точность составляет 88,84%) и увеличение F-меры для облаков, их теней и чистых пикселей в 0,0261, 0,1211 и 0,0093 соответственно (сама F-мера составляет 0,9242, 0,5753 и 0,8902, соответственно).

2.3. DeepMask

В качестве основы DeepMask [8] использует ResNet [7], сеть, которая содержит пропускные соединения и механизм остаточного обучения: вместо предсказания всей функции $H(x)$ сеть учится предсказывать $H(x) - x$, что позволяет ResNet иметь увеличенную глубину без проблемы затухающего и взрывающегося градиента, при которых необходимая для обучения функция градиента стремится к нулю и к слишком большому значению соответственно.

Используемая версия ResNet включает в себя следующее:

1. Свертка 3 x 3.
2. 3 остаточных блока с 16, 32 и 64 фильтрами 3 x 3 соответственно.
3. Средний пулинг.
4. Полносвязный слой.
5. Многопеременная логистическая функция.

Остаточные блоки состоят из равного количества меньших блоков со следующим содержимым:

1. Первая свертка 3 x 3.
2. Первая нормализация серии.
3. Первая функция активации, отбрасывающая отрицательные значения.
4. Выпадение, во время которого отбрасывается половина значений нейронов.

5. Вторая свертка 3 x 3.

6. Вторая нормализация серии.

7. Вторая функция активации, аналогичная первой. При этом значения до входа в первую свертку копируются в значения до второй функции активации.

DeepMask состоит из двух главных модулей: извлекателя местных регионов и классификатора центрального пикселя, основанного на ResNet. Извлекатель местных регионов принимает на вход изображение произвольного размера в нескольких спектральных диапазонах и извлекает из него местные регионы размеров 15 x 15 пикселей. Затем основанный на ResNet модуль возвращает вероятность, с которой центральный пиксель закрыт облаком или тенью.

Лучший результат показала ResNet, остаточные блоки которой состояли из 3 меньших блоков. Начальный коэффициент скорости обучения — 0,1, уменьшается в 10 раз, когда потеря подтверждения (сумма ошибок набора подтверждения[2]) начинает изменяться слишком медленно.

Было рассчитано время классификации для ПК с главным процессором Intel i7 и одним графическим процессором NVIDIA 1080 Ti: 7-8 минут для одного тестового изображения.

Точность для чистых и закрытых пикселей для DeepMask составляет 93,85% и 94,76%, что больше по сравнению с CFMask [5] на 2,83% и 6,65%, соответственно. Чувствительность — 94,1% и 92,8%, что на 17,21% больше и на 1,35% меньше результатов CFMask, соответственно. Авторы аргументируют уменьшение чувствительности для облаков и теней тем, что DeepMask сопоставляет им вероятности меньше 0,5 для некоторых пикселей, из-за чего эти пиксели автоматически определяются, как чистые. Наконец, F-мера составила 93,66% для чистых пикселей и 93,42% — для закрытых, что больше результатов CFMask

[2]Набор подтверждения — специальная часть датасета, выделенная для подтверждения производительности сети. Используется для корректировки параметров обучения без получения ответа на тестовый набор.

на 11,41% и 3,93%, соответственно.

2.4. AFMUNet

AFMUNet [1] (Attention Feature Fusion Network Based on a U-Shaped Structure — слияющая признаки сеть внимания, основанная на структуре U-сети) не изменяет кодировщик и узкое место типичной архитектуры U-сети, но дополняет дешифратор, использует особую функцию потери и имеет механизм внимания.

Кодировщик и дешифратор состоят из 4 слоев. На каждом слое кодировщика карты признаков уменьшаются в 2 раза и их количество возрастает в 2 раза. На слоях дешифратора — наоборот. Узкое место представляет собой 2 свертки 3 x 3.

После каждого обычного блока дешифратора следует канальный пространственный модуль внимания (Channel Spatial Attention Module). Он состоит из следующих друг за другом канального блока внимания (Channel Attention Block) и пространственного блока внимания (Spatial Attention Block). Результат работы каждого из них умножается на результат работы прошлого шага.

Канальный блок внимания призван улучшить внимание, выделяемое каналам, корректируя их коэффициенты. Принцип работы заключается в поэлементной сумме результатов работы двухслойной нейронной сети с распределением веса, примененной к результатам среднего и максимального пулинга входной карты признаков. Впоследствии применяется функция сигмоиды, что влечет за собой результат - признак внимания канала.

Пространственный блок внимания работает по похожему принципу: тоже используются результаты среднего и максимального пулинга, но вместо суммы к ним применяется конкатенация, свертка с размером фильтра 7 x 7 (по словам авторов, такой размер более производителен) и функция сигмоиды в качестве функции активации.

На пропускных соединениях между обычными модулями кодировщика и дешифратора расположены модули слияния признаков (Feature

Fusion Module). Он работает по следующему принципу:

1. Слияние двух карт признаков путем применения различных операций в гибридную.
2. Получение среднего значения всей полученной карты.
3. Свертка 1 x 1 и приравнивание отрицательных значений к нулю.
4. Свертка 1 x 1 и применение логистической функции.
5. Поэлементное умножение результата прошлого шага на гибридную карту признаков.
6. Поэлементное сложение результата прошлого шага с гибридной картой признаков.

В качестве функции потери используется линейная комбинация функции потерь перекрестной энтропии (наиболее часто используемая функция потерь для решения задачи сегментации) и функции потерь Дайса (помогающая нейронной сети разделять изображение на регионы без пересечений).

Авторы использовали иные метрики для оценки качества сегментации, включая пиксельную точность (pixel accuracy, отношение суммы истинно положительных пикселей и истинно отрицательных пикселей ко всем пикселям изображения) и среднее пересечение по объединению (mean intersection over union, отношение истинно положительных пикселей ко всем пикселям кроме истинно отрицательных). F-мера, как и прежде, вычисляется с использованием обычной точности и чувствительности.

На датасете GF1_WHU по сравнению с обычной UNet показал увеличение пиксельной точности на 1,07%, среднего пересечения по объединению — на 1,96% и F-меры — на 1,89%; результаты составили 97,4%, 93,28% и 95,1%, соответственно.

2.5. Сравнение

	Точность (%)	Чувствительность (%)	F-мера
TSMM	95	89	0,88
LANA	91,08	93,79	0,9242

Таблица 1: Результаты методов для облаков.

	Точность (%)	Чувствительность (%)	F-мера
TSMM	96	63	0,62
LANA	51,21	65,62	0,5753

Таблица 2: Результаты методов для теней.

	Точность (%)	Чувствительность (%)	F-мера
TSMM	93	86	0,85
DeepMask	94,76	92,8	0,9342

Таблица 3: Результаты методов для обоих эффектов.

	Точность (%)	Чувствительность (%)	F-мера
TSMM	93	90	0,9
LANA	92,15	86,09	0,8902
DeepMask	93,85	94,1	0,9366

Таблица 4: Результаты методов для чистых пикселей.

Учитывая, что наибольшее значение (0,951) общей F-меры достигает AFMUNet, разумно предположить, что в дальнейшем стоит использовать подобный метод; также нельзя не отметить высокие результаты DeepMask.

3. Эксперимент

3.1. Датасеты

1. CloudSEN12 [2]. Размер составляет примерно 1 терабайт. Состоит из 49400 блоков изображений. Снимался с 2018 по 2020 годы. Содержит 10 тысяч обозначенных изображений для чистых, покрытых тонкими облаками, покрытых плотными облаками и покрытых тенью пикселей. Используется алгоритмом TSMM.

2. Датасеты подтверждения CCA [4] (Cloud Cover Assesment Validation Datasets), созданные USGS. Основаны на снимках Landsat 7-9. Используются алгоритмом DeepMask и частично — сетью LANA.

3. GF-1_WHU [13]. Основан на снимках широкого поля зрения GaoFen-1. Облака и тени помечены вручную в Уханьском университете.

3.2. Метрики

- Обнаружение облаков: точность, чувствительность, F-мера.
- Обнаружение теней: точность, чувствительность, F-мера.
- Скорость работы.

Заключение

Итогом учебной практики за весенний семестр послужил выполненный анализ предметной области. Таким образом, на следующий год остаются эти задачи:

1. Проектирование алгоритма обнаружения облаков и их теней.
2. Реализация обнаружения облаков и их теней.
3. Оценка эффективности полученного метода.
4. Анализ методов увеличения точности фильтрации.

Список литературы

[1] AFMUNet: Attention Feature Fusion Network Based on a U-Shaped Structure for Cloud and Cloud Shadow Detection / Wenjie Du, Zhiyong Fan, Ying Yan et al. // Remote Sensing. — 2024. — 04. — Vol. 16. — P. 1574.

[2] Aybar Cesar, Ysuhuaylas Luis, Loja Jhomira. CloudSEN12, a global dataset for semantic understanding of cloud and cloud shadow in Sentinel-2 // Sci Data. — 2022. — Vol. 9. — P. 782. — URL: https://doi.org/10.1038/s41597-022-01878-2.

[3] Calculation of Radiative Fluxes From the Surface to top of Atmosphere Based on ISCCP and Other Global Data Sets: Refinements of the Radiative Transfer Model and the Input Data / Yuanchong Zhang, William Rossow, Andrew Lacis, Valdar Oinas // Journal of Geophysical Research. — 2004. — 10. — Vol. 109.

[4] Cloud Cover Assessment Validation Datasets. — 2021. — accessed on 28 November 2024. URL: https://www.usgs.gov/landsat-missions/cloud-cover-assessment-validation-datasets.

[5] Cloud detection algorithm comparison and validation for operational Landsat data products / Steve Foga, Pat L. Scaramuzza, Song Guo et al. // Remote Sensing of Environment. — 2017. — Vol. 194. — P. 379–390. — URL: https://www.sciencedirect.com/science/article/pii/S0034425717301293.

[6] Comprehensive quality assessment of optical satellite imagery using weakly supervised video learning / Valerie Pasquarella, Christopher Brown, Wanda Czerwinski, William Rucklidge. — 2023. — 06. — P. 2125–2135.

[7] Deep Residual Learning for Image Recognition / Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun // CoRR. — 2015. — Vol. abs/1512.03385. — arXiv : 1512.03385.

[8] DeepMask: an algorithm for cloud and cloud shadow detection in optical satellite remote sensing images using deep residual network / Ke Xu, Kaiyu Guan, Jian Peng et al. // CoRR. — 2019. — Vol. abs/1911.03607. — arXiv : 1911.03607.

[9] Efficient object localization using Convolutional Networks / Jonathan Tompson, Ross Goroshin, Arjun Jain et al. // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2015. — P. 648–656.

[10] Han Shuning, Wang Jianmei, Zhang Shaoming. Former-CR: A Transformer-Based Thick Cloud Removal Method with Optical and SAR Imagery // Remote Sensing. — 2023. — 02. — Vol. 15. — P. 1196.

[11] Improving neural networks by preventing co-adaptation of feature detectors / Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky et al. // CoRR. — 2012. — Vol. abs/1207.0580. — arXiv : 1207.0580.

[12] Loshchilov Ilya, Hutter Frank. SGDR: Stochastic Gradient Descent with Warm Restarts. — 2017. — 1608.03983.

[13] Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery / Zhiwei Li, Huanfeng Shen, Huifang Li et al. // Remote Sensing of Environment. — 2017. — Vol. 191. — P. 342–358. — URL: https://www.sciencedirect.com/science/article/pii/S003442571730038X.

[14] A Novel Method for Cloud and Cloud Shadow Detection Based on the Maximum and Minimum Values of Sentinel-2 Time Series Images / Kewen Liang, Gang Yang, Yangyan Zuo et al. // Remote Sensing. — 2024. — 04. — Vol. 16. — P. 1392.

[15] Zhang Hankui, Luo Dong, Roy David. Improved Landsat Operational Land Imager (OLI) Cloud and Shadow Detection with the Learning Attention Network Algorithm (LANA) // Remote Sensing. — 2024. — 04. — Vol. 16. — P. 1321.

[16] Zhu Zhe, Woodcock Curtis. Object-based cloud and cloud shadow detection in Landsat imagery // Remote Sensing of Environment. — 2012. — 03. — Vol. 118. — P. 83–94.

[17] Безуглый, Т. А. Использование спутниковых снимков для выявления проблемы несанкционированных свалок в Челябинской области / Т. А. Безуглый, А. Р. Сибиркина // Использование и охрана природных ресурсов в России. – 2023. – № 2(174). – С. 58-62. – EDN OIIRHU. — URL: https://www.elibrary.ru/item.asp?id=54356689.

[18] Трофимова Н. В., Сухинин А. И. Оценка загрязнения окружающей среды дымовыми выбросами от пожаров по спутниковым данным // Технологии гражданской безопасности. 2004. №4. — URL: https://cyberleninka.ru/article/n/otsenka-zagryazneniya-okruzhayuschey-sredy-dymovymi-vybrosami-ot