



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ _____ «Информатика и системы управления»

КАФЕДРА _____ «Теоретическая информатика и компьютерные технологии»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ
НА ТЕМУ:
«Анализ гиперспектральных изображений с
помощью осцилляторных нейронных сетей»

Студент _____
(Группа)

(Подпись, дата)

(И.О. Фамилия)

Руководитель ВКР

(Подпись, дата)

(И.О. Фамилия)

Консультант

(Подпись, дата)

(И.О. Фамилия)

Консультант

(Подпись, дата)

(И.О. Фамилия)

Нормоконтролер

(Подпись, дата)

(И.О. Фамилия)

2024 г.

АННОТАЦИЯ

Записка состоит из 26 страниц, содержит 2 разделов, 1 таблиц и 1 рисунков. Окончания согласованы криво, потому что подсчёт страниц, разделов, таблиц, рисунков и листингов осуществляется автоматически с помощью пакетов L^AT_EX. Если захочется автоматически посчитать что-то ещё, нужно добавить в преамбулу документа строку `\DeclareTotalCounter{что-то ещё}` и затем воспользоваться командой `\totalчто-то ещёs{}` там, где потребуется.

СОДЕРЖАНИЕ

| | |
|---|----|
| ВВЕДЕНИЕ | 4 |
| 1 Гиперспектральные изображения в задачах анализа спутни- ковых снимков | 6 |
| 1.1 Типы спутниковых снимков | 6 |
| 1.2 Классификация задач анализа спутниковых снимков . . | 9 |
| 1.3 Анализ методов обработки спутниковых снимков | 12 |
| 1.4 Выводы по разделу 1 | 14 |
| 2 Осцилляторные нейронные сети | 16 |
| 2.1 Исследование входных данных при разработке осцил- ляторных нейронных сетей | 16 |
| 2.2 Архитектуры осцилляторных нейронных сетей | 17 |
| 2.3 Обучение осцилляторных нейронных сетей | 20 |
| 2.4 Выводы по разделу 2 | 24 |
| ЗАКЛЮЧЕНИЕ | 25 |
| СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ | 26 |
| ПРИЛОЖЕНИЕ А | 26 |

ВВЕДЕНИЕ

Количество спутников дистанционного зондирования Земли значительно увеличилось за последние десятилетия. Начиная с 2010 года, наблюдается особенно быстрый рост числа запусков, связанный с развитием технологий и увеличением потребности в данных дистанционного зондирования [1]. Также многие ведущие космические агентства из разных стран разрабатывают свои программы по развитию собственных группировок спутников дистанционного зондирования Земли [2, 3, 4]. Эти данные подтверждают тенденцию к увеличению количества спутников дистанционного зондирования Земли.

Одним из основных типов информации, получаемых с помощью спутников дистанционного зондирования Земли, являются гиперспектральные изображения. Гиперспектральные изображения представляют собой данные, собранные и обработанные из различных спектров света. Эти данные могут быть использованы для идентификации и анализа объектов, материалов и процессов в различных областях, включая экологию, сельское хозяйство, безопасность, геологию, промышленность и медицину. При этом рынок технологий дистанционного зондирования Земли по оценкам вырастет с 14.4 млрд. долл. США до 33.3 млрд. долл. США в 2030 [5]. Поэтому разработка методов для анализа гиперспектральных изображений является **актуальной задачей**.

Цель выпускной работы состоит в разработке и создании системы анализа гиперспектральных изображений с применением осцилляторных нейронных сетей.

Для достижения указанной цели в работе решаются следующие **основные задачи**:

- Изучение особенности работы с гиперспектральными изображениями;
- Обзор и исследование существующих методы анализа и тенденции их развития;
- Разработка и экспериментальное исследование осцилляторной нейронной сети для классификации гиперспектральных изображений;
- Исследование и выбор архитектуры для улучшения качества работы осцилляторной нейронной сети, в том числе с применением многоагентных систем и других видов нейронных сетей;

— Реализация выбранной архитектуры и исследование ее работоспособность в области дистанционного зондирования Земли.

Объектом исследования являются гиперспектральные изображения.

Предметом исследования являются системы анализа гиперспектральных изображений, использующие осцилляторные нейронные сети.

1 Гиперспектральные изображения в задачах анализа спутниковых снимков

1.1 Типы спутниковых снимков

Существует несколько типов изображений, которые получают со спутников. Они варьируются в зависимости от используемого спутника, его орбиты, а также специфических целей и задач, которые эти изображения должны выполнить.

Оптические изображения. Это наиболее распространенный тип изображений, которые мы получаем со спутников. Они представляют собой визуальные изображения Земли, снятые в различных спектрах света.

В настоящее время для съемок из космоса наиболее часто используются многоспектральные оптико-механические системы — сканеры, установленные на ИСЗ различного назначения. При помощи сканеров формируются изображения, состоящие из множества отдельных, последовательно получаемых элементов.

Важнейшей характеристикой сканера являются угол сканирования (обзора) и мгновенный угол зрения, от величины которого зависят ширина снимаемой полосы и разрешение. В зависимости от величины этих углов сканеры делят на точные и обзорные. У точных сканеров угол сканирования уменьшают до $\pm 5^\circ$, а у обзорных увеличивают до $\pm 50^\circ$. Величина разрешения при этом обратно пропорциональна ширине снимаемой полосы.

Космические системы, такие как Landsat-7, 8, Sentinel-1, 2 и MODIS / NAIP, предоставляют такие изображения.

Инфракрасные изображения. Космические снимки, полученные в инфракрасном спектральном диапазоне, содержат информацию о собственном тепловом излучении объектов земной поверхности. Природные объекты обладают различными тепловыми и излучательными свойствами и по-разному реагируют на поступление солнечного излучения, что находит отображение в яркости их изображения на тепловых снимках. По тепловому режиму можно выделить объекты земной поверхности с температурой, обусловленной внутренним теплом (вулканы, термальные воды и др.), и нагреваемые Солнцем (растительность, почвы и др.). Наиболее динамичный температурный режим у объектов второй группы, который зависит от их физических и химических свойств, а также интенсивности солнечного излучения. Поэтому максимальные температурные контрасты при-

родных объектов наблюдаются около полудня, они меньше вечером и значительно сглаживаются ночью, достигая минимума предутренние часы.

Инфракрасный диапазон спектра электромагнитных колебаний условно делится на три части (в мкм):

- ближний (0,74—1,35);
- ближний (0,74—1,35);
- ближний (0,74—1,35).

Солнечное (внешнее) и эндогенное (внутреннее) тепло нагревает геологические объекты по-разному в зависимости от литологических свойств пород, тепловой инерции, влажности, альбедо и многих других причин.

ИК-излучение, проходя через атмосферу, избирательно поглощается, в связи с чем тепловую съемку можно вести только в зоне расположения так называемых “окон прозрачности” — местах пропускания ИК-лучей. Опытным путем выделено четыре основных окна прозрачности (в мкм): 0,74—2,40; 3,40—4,20; 8,0—13,0; 30,0—80,0 [6].

Наиболее широко тепловые космические снимки низкого и очень низкого пространственного разрешения используются на глобальном уровне (вся планета, континенты, океаны) для изучения пространственно-временного распределения температур земной [7] и океанической [8, 9, 10] поверхности, состояния атмосферы [11, 12], исследования вулканической активности [13]. Менее распространены работы, посвященные анализу локальных географических объектов [14].

Тепловые космические снимки также часто находят применение в сельском и лесном хозяйстве. Анализируя тепловые изображения, ученые могут определять состояние сельскохозяйственных культур, выявлять области, пораженные вредителями или болезнями, а также отслеживать изменения растительности с течением времени.

Радарные изображения. Радарная космическая съемка выполняется в ультракоротковолновой области радиоволн, подразделяемой на X-, C- и L- диапазоны (Таблица 1). Существуют пять видов диапазонов:

- X-диапазон (от 2,4 до 3,75 см). Его используют для решения некоторых целей гражданской жизни, но в основном военных задач;
- C-диапазон (от 3,75 до 7,5 см). Используется не только для построения ЦММ и ЦМР, но и при мониторинге смещений литосферных плит;

- S-диапазон (от 7,5 до 15 см). Используется в приложениях для военных и гражданских жителей;
- L-диапазон (от 15 до 30 см). Используют для наблюдения за почвенным покровом, ледяным и снежным, а также за растительность и лесами до средней густоты;
- P-диапазон (от 30 до 100 см). Данный диапазон применяют для оценки биомассы и реализован в основном на авиа-носителях.

Радарные изображения обычно используют для фотографирования местности или объектов, которые покрыты облаками или испытывающие на себе другие факторы, не позволяющие провести оптическое наблюдение.

Таблица 1 — Диапазоны радиоволновой области электромагнитного спектра, которые применяются для спутниковой съемки.

| Диапазон | Частоты, ГГц | Длины волн, см. | Спутниковые системы |
|----------|--------------|----------------------------|--|
| X | 5.20 - 10.90 | 2.75 - 5.77 (2.4 - 3.8) | USGS SLAR, TerraSAR-X/TanDEM-X, Cosmo-SkyMed 1-4 |
| C | 3.9 - 6.2 | 3.8 - 7.6 | ERS-1,2; ENVISAT-ASAR RADARSAT-1,2 |
| L | 0.39 - 1.55 | 19.3 - 76.9 (15 - 30) | SIR-A,B, JERS, ALOS PALSAR |

Гиперспектральные изображения. Это изображения, которые получаются в результате сбора и комбинирования данных в сотнях узких спектральных диапазонов.

В результате гиперспектральной съемки получается большое количество гиперспектральных изображений. Их количество может варьироваться от нескольких десятков до нескольких сотен и зависит от технических возможностей аппаратуры. При этом данные получают с помощью специальных сенсоров [15].

Каждый снимок представляет собой узкополосный диапазон длин волн электромагнитного спектра (спектральный диапазон). Изображения объединяются для формирования трехмерного гиперспектрального куба данных для последующей обработки и анализа [16]. Этот куб содержит спектральные данные по одному измерению и пространственные данные по двум другим. При этом каждый пиксель изображения характеризуется собственным спектром [15].

Гиперспектральные снимки обычно включают сотни спектральных полос относительно узкой полосы пропускания (5—10 нм). Узкие спектральные каналы способны обнаруживать незначительные спектральные особенности, которые остаются незамеченными в более широких диапазонах.

Гиперспектральные снимки полезны своей высокой информативностью, так как вся доступная полезная информация заключена именно в спектре восходящего от объекта зондирования излучения [17]. Стоит отметить, что гиперспектральные снимки обладают высоким спектральным, но низким пространственным разрешением [18]. На Рисунке 1 приведена информация о доступных спутниковых системах с гиперспектральными камерами [19].

| # | Hyperspectral Imager | Satellites / Platform | Orbit | Hyperspectral Imaging Approach | Launch Year | Demission Year | Spectral Range (μm) | Number of Spectral Bands | Spectral Sampling Interval (nm) | Ground Sampling Distance (m) | Swath Width (km) |
|----|----------------------|-----------------------|------------------------|--------------------------------|--------------------------|----------------|------------------------|--------------------------|---------------------------------|------------------------------|-----------------------|
| 1 | SPIMs 1-5 | MSX | LEO | Dispersive element | 1996 | 2008 | 0.11–0.9 | 272 | 0.5–4.3 | 770 | 15 |
| 2 | HSI | LEWIS | LEO | Dispersive element | 1997 | 1997 | 0.4–1.0 0.9–2.5 | 128 256 | 5.0 6.5 | 30 | 7.7 |
| 3 | Hyperion | EO-1 | LEO | Dispersive element | 2000 | 2017 | 0.4–2.5 | 220 | 10 | 30 | 7.7 |
| 4 | CHRIS | PROBA | LEO | Dispersive element | 2001 | Active | 0.4–1.0 | 19-62 | 1.25–11 | 25–50 | 13 |
| 5 | MERIS | ENVISAT | LEO | Dispersive element | 2002 | 2012 | 0.39–1.04 | 520 (transmit 15) | 1.25 | 300 | 1150 |
| 6 | VIRTIS | Rosetta | Comet 67P | Dispersive element | 2004 | 2016 | 0.28–1.10 1.05–5.13 | 432 432 | 1.89 9.47 | 0.014°× 0.014° (IFOV) | 3.67°× 3.67° (FOV) |
| 7 | VIRTIS | Venus Express | Venus | Dispersive element | 2005 | 2015 | 0.28–1.10 1.05–5.13 | 432 432 | 1.89 9.47 | 0.014°× 0.014° (IFOV) | 3.67°× 3.67° (FOV) |
| 8 | CRISM | MRO | Mars orbit (300 km) | Dispersive element | 2005 | Active | 0.37–3.92 | 455 | 6.55 | 18 | 10.8 |
| 9 | VIRTIS | NASA-Dawn | Vesta, Ceres | Dispersive element | 2007 | 2018 | 0.28–1.10 1.05–5.13 | 432 432 | 1.89 9.47 | 0.014°× 0.014° (IFOV) | 3.67°× 3.67° (FOV) |
| 10 | M3 | Chandrayaan-1 | Lunar orbit | Dispersive element | 2008 | 2009 | 0.43–3.0 | 260 | 10 | 70 | 40 |
| 11 | FTHSI | HJ-1A | LEO | Fourier Transform | 2008 | Active | 0.45–0.95 | 115 | 4 | 100 | 50 |
| 12 | HySI | IMS-1 | LEO | Spectral filter | 2008 | 2012 | 0.40–0.95 | 64 | 8 | 500 | 130 |
| 13 | ARTEMIS | TacSat-3 | LEO | Dispersive element | 2009 | 2012 | 0.40–2.50 | 400 | 5 | 4 | 4 |
| 14 | HICO | ISS | ISS orbit | Dispersive element | 2009 | 2014 | 0.35–1.08 | 128 | 5.7 | 90 | 51 |
| 15 | VNIS | Change'E | Lunar rover | Spectral filter | 2013 | 2015 | 0.45–0.95 0.9–2.4 | 100 | 5 | - | 8.5°×8.5° (FOV) |
| 16 | OLCI | Sentinel-3A | LEO | Dispersive element | 2016 | Active | 0.39–1.04 | 520 (transmit 21) | 1.25 | 300 | 1270 |
| 17 | MHRIS | GHGSat-D | LEO | Spectral filter | 2016 | Active | 1.6–1.7 | 512 | 0.2 | 50 | 15 |
| 18 | AaSI | Aalto-1 | LEO | Spectral filter | 2017 | Active | 0.50–0.90 | 6–20 | 7–10 | 192 | 97 |
| 19 | DESIS | ISS | ISS orbit | Dispersive element | 2018 | Active | 0.40–1.0 | 235 | 2.55 | 30 | 30 |
| 20 | HyperScout | GomX-4B | LEO | Spectral filter | 2018 | Active | 0.4–1.0 | 45 | 15 | 50 | 200 |
| 21 | AHSI | Gaofen-5 | LEO | Dispersive element | 2018 | Active | 0.40–2.50 | 330 | 5 (VNIR) 10 (SWIR) | 30 | 60 |
| 22 | HysIS | IMS-2 | LEO | Dispersive element | 2018 | Active | 0.40–2.4 | 256 | 10 | 30 | 30 |
| 23 | PRISMA | PRISMA | LEO | Dispersive element | 2019 | Active | 0.40–2.51 | 237 | 12 | 30 | 30 |
| 24 | HISUI | ISS | LEO | Dispersive element | 2019 | Active | 0.40–2.50 | 185 | 10 (VNIR) 12.5 (SWIR) | 30 | 20 |
| 25 | HIS/ GISAT-1 | GISAT | GEO | Dispersive element | June 2021 (scheduled) | N/A | 0.38–1.0 0.9–2.5 | 158 256 | 3.92 (VNIR) 6.25 (SWIR) | 318 191 | 160 190 |

Рисунок 1 — Список спутниковых систем с гиперспектральной съемкой, запущенных до 2021 года.

1.2 Классификация задач анализа спутниковых снимков

Задачи анализа спутниковых снимков могут быть классифицированы по различным критериям, включая типы данных, используемые методы анализа и цели

исследования. Методы анализа будут подробно рассмотрены далее. В данном подразделе проведен поверхностный обзор различных целей исследования, а также дана общая классификация задач.

По целям исследования задачи можно разделить на следующие категории:

1. Мониторинг различных объектов и процессов, таких как вырубка лесов, рост и созревание урожая, передвижение ледовых покровов, изменения в городской среде и других динамически изменяющихся объектов.
2. Прогнозирование. Объектом исследования может являться урожайность, количество населения, а также различные стихийные бедствия.
3. Исследование территорий. Поиск новых месторождений полезных ископаемых, подходящих мест для прокладки дорог или строительства сооружений являются одними из наиболее частых примеров.

Стоит отметить, что хоть задачи и различаются достаточно сильно по целям исследования, при их декомпозиции они распадаются на 3 основные задачи:

- Детекция (обнаружение) объектов на изображении;
- Классификация изображений;
- Семантическая сегментация.

Для каждой из этих задач разработаны и применяются отдельные методы решения, поэтому важно понимать специфику этих задач.

Детекция объектов на изображении – это процесс идентификации и локализации объектов на изображении. Эта задача является одной из ключевых в области компьютерного зрения и машинного обучения. Основные этапы детекции объектов:

1. Сканирование изображения. На первом этапе алгоритм сканирует изображение, используя так называемое "скользящее окно"(sliding window). Это окно перемещается по изображению, и в каждом положении окна алгоритм пытается определить, есть ли на этом участке изображения объект, который нужно обнаружить.
2. Выделение признаков. На этом этапе алгоритм анализирует каждый участок изображения, на котором может находиться объект, и выделяет признаки, которые помогут определить, есть ли на этом участке объект. Признаки могут включать в себя цвет, текстуру, форму и другие характеристики объекта.

3. Классификация. После выделения признаков алгоритм классифицирует каждый участок изображения, определяя, есть ли на нем объект, и если да, то какого типа этот объект.
4. Локализация. После того, как объекты были обнаружены и классифицированы, алгоритм определяет их точное положение на изображении. Это обычно делается путем определения ограничивающего прямоугольника (bounding box) вокруг каждого объекта.

Задача классификации изображений заключается в отнесении входного изображения к одному или нескольким классам на основе его визуального содержания. Основными этапами классификации изображений являются:

1. Предварительная обработка изображения: Перед обучением модели все изображения приводятся к одному и тому же размеру. Также могут быть применены различные техники аугментации данных, такие как повороты, сдвиги, изменение масштаба и т.д., чтобы увеличить объем тренировочных данных и сделать модель более устойчивой к изменениям в входных данных.
2. Извлечение признаков: На этом этапе из изображения извлекаются ключевые признаки, которые могут помочь в классификации. Это может быть сделано с помощью традиционных методов, таких как фильтры Габора или HOG (Histogram of Oriented Gradients), или с помощью сверточных нейронных сетей (Convolutional Neural Networks, CNN), которые могут автоматически извлекать и учить признаки из изображений.
3. Классификация: После извлечения признаков эти данные подаются на вход классификатору, который может быть основан на различных методах машинного обучения, включая SVM (Support Vector Machines), случайные леса (Random Forests) или глубокие нейронные сети. Классификатор обучается на основе этих признаков, чтобы отнести каждое изображение к одному из предопределенных классов.

Одной из наиболее перспективных и важных подзадач классификации является задача распознавания образов [20].

Задача семантической сегментации является наиболее сложной из представленных. Семантическая сегментация изображений - это процесс разделения изображения на различные сегменты, где каждый сегмент представляет собой

множество пикселей, которые принадлежат одному и тому же классу объектов. В отличие от задачи классификации изображений, где целью является определение класса всего изображения, или задачи детекции объектов, где целью является определение класса и положения отдельных объектов на изображении, семантическая сегментация стремится к классификации каждого пикселя изображения.

Основные этапы семантической сегментации схожи с этапами классификации за исключением заключительного шага. Он заменен на попиксельную классификацию, которая выполняется с помощью сверточных нейронных сетей и др.

1.3 Анализ методов обработки спутниковых снимков

В последнее время с развитием области глубокого обучения появилось множество новых методов для обработки спутниковых снимков в дополнение к классическим методам. Классическими методами обработки спутниковых снимков в данной работе будут называться те, которые не содержат в себе продвинутых нейросетевых алгоритмов, таких, например, как глубокие сверточные нейронные сети (от англ. *deep convolutional neural networks*, CNN).

Классические методы можно разделить на 2 группы:

- Методы, не связанные с обучением;
- Методы, основанные на обучении.

Методы, не связанные с обучением, в основном разрабатывались для решения задач детекции объектов. В них использовались различные подходы, например:

- Вейвлет-анализ с несколькими разрешениями;
- Масштабно-инвариантное преобразование объектов;
- Модель набора слов для пространственного разреженного кодирования (от англ. *spatial sparse coding bag-of-words*) [21].

С развитием методов машинного обучения задачу детекции объектов стали рассматривать как задачу классификации и стали применять для ее решения различные подходы с использованием обучения:

- Метод опорных векторов (от англ. *Support Vector Machines*);
- Модели гауссовой смеси (от англ. *Gaussian Mixture Model*);
- Квадратичный дискриминантный анализ;
- Решающие деревья, случайный лес (от англ. *Random Forest*).

Несмотря на развитие алгоритмов анализа спутниковых изображений, качество классификации для части изображений оставалось неадекватным. Это обуславливалось большим разнообразием пространственных и спектральных признаков изображений. Чтобы справиться с этой особенностью, были разработаны подходы, которые учитывали как пространственные, так и временные признаки изображения. Эти подходы основывались на таких методах, как:

- Марковские случайные поля (от англ. Markov Random Field) [22];
- Условные случайные поля (от англ. Conditional Random Field) [23];
- Методы составного ядра (от англ. Composite Kernel Methods) [24].

Описанные выше методы в настоящее время редко используются в практических задачах, поэтому не будем давать их подробное описание. На данный момент наиболее перспективными считаются методы, основанные на нейросетевых технологиях. Они тоже поддаются дополнительной классификации в зависимости от используемых подходов:

- Сверточные нейронные сети. В настоящее время применяются различные модификации классической сверточной нейронной сети, такие как глубокие сверточные сети, сети с остаточными блоками (от англ. residual), 3-D сверточные сети. Последние широко применяются в анализе гиперспектральных изображений, так как позволяют учитывать сразу спектральные и пространственные признаки при применении операции свертки;
- Сети глубокого доверия. Состоят из композиции обучающих модулей, каждая из которых является ограниченной машиной Больцмана. Преимущество данной сети состоит в возможности “послойного” обучения ограниченных машин Больцмана. Данные сети применяются для выделения скрытых признаков из распределения данных, оценки семантического расстояния между различными классами;
- Осцилляторные нейронные сети. Основная идея осцилляторной нейронной сети заключается в том, что нейроны могут быть представлены как осцилляторы, которые генерируют периодические сигналы. Эти сигналы могут быть синхронизированы с сигналами от других нейронов, что позволяет сети обрабатывать информацию и выполнять задачи по распознаванию образов. В осцилляторной нейронной сети, в общем случае, все

- нейроны полностью связаны (fully connected). Это означает, что каждый нейрон может получать сигналы от всех остальных нейронов в сети;
- Многослойные автоэнкодеры. Они состоят из двух основных частей: кодировщика (encoder), который преобразует входные данные в некоторое скрытое представление, и декодера (decoder), который восстанавливает данные из этого скрытого представления. Для задач обработки гиперспектральных изображений интерес представляет распределение признаков в скрытом представлении. Его еще называют латентным пространством (от англ. latent space). Выбор признаков из данного пространства помогает уменьшить размерность данных, при этом сохранив ключевую информацию, заключенную в изображении;
 - Визуальные трансформеры (от англ. visual transformers). Это модели машинного обучения, которые применяют архитектуру трансформера, изначально разработанную для обработки естественного языка, к задачам компьютерного зрения. Визуальные трансформеры начинают с разделения входного изображения на решетку небольших патчей или "визуальных токенов". Эти патчи затем линейризуются и обрабатываются с использованием трансформера. Трансформер состоит из нескольких слоев, каждый из которых содержит два основных компонента: механизм внимания и полносвязные сети (от англ. feed-forward networks, FFN). Механизм самовнимания позволяет модели учесть контекст каждого визуального токена, учитывая все остальные токены, а FFN обеспечивают дополнительную обработку. В отличие от традиционных сверточных нейронных сетей (от англ. convolutional neural networks, CNN), которые обрабатывают изображения с использованием локальных сверток, визуальные трансформеры обрабатывают все визуальные токены одновременно, что позволяет им учесть глобальные зависимости между различными частями изображения [25].

1.4 Выводы по разделу 1

1. Проведен обзор различных типов изображений, получаемых со спутников.

2. Рассмотрены различные типы задач, встречающиеся в процессе анализа спутниковых снимков.
3. Проанализированы различные методы и алгоритмы обработки снимков, в том числе с использованием новейших достижений в области нейросетевых технологий.

2 Осцилляторные нейронные сети

2.1 Исследование входных данных при разработке осцилляторных нейронных сетей

Этап предварительного исследования входных данных является важной частью процесса разработки искусственных нейронных сетей, в частности осцилляторных нейронных сетей (ОНС). В данном параграфе будут подробно рассмотрены только те аспекты входных данных, которые непосредственно влияют на выбор архитектуры ОНС, процесс обучения и полученный результат в контексте задачи обработки гиперспектральных изображений.

Рассмотрим некоторые особенности осцилляторных нейронных сетей:

- Информация в ОНС запоминается на основе фаз ее составных элементов, осцилляторов;
- Размер ОНС растет пропорционально количеству паттернов или объектов, которые могут быть классифицированы или сегментированы [26];
- Способность ОНС к адаптации, проявляющаяся в модификации параметров (весов) элементов или связей в качестве реакции на внешнюю среду.

Учитывая вышеизложенные особенности ОНС, а также специфику гиперспектральных изображений, описанную в блоке 1.1, стоит обратить внимание на следующие ограничения параметров входных данных:

- Невысокое количество различных типов целевых объектов;
- Не учитывается распределение одновременно пространственных и спектральных признаков (при использовании классического варианта ОНС);
- При высокой сложности целевых объектов потребуются использование многослойных ОНС, что может заметно увеличить их размер и время работы.

Принимая во внимание данные ограничения, можно сузить область эффективного применения ОНС, исключив из нее, в общем случае, задачи семантической сегментации. В частном случае, при малом количестве типов объектов на снимке, ОНС применимы к этому типу задач.

Несмотря на некоторые ограничения, ОНС могут обрабатывать широкий набор входных данных. Кроме того, они допускают изменение вида или формы

целевого объекта при незначительной потере точности, что позволяет несколько расширить множество входных данных.

2.2 Архитектуры осцилляторных нейронных сетей

Для успешного построения модели ОНС необходимо определить модель ее составной части, нейрона (осциллятора). Рассмотрим наиболее распространенные модели осцилляторов:

1. Модифицированный осциллятор Ван дер Поля. Он представляет собой релаксационный осциллятор, его динамика определяется двумя связанными переменными x и y и подчиняется двумерной динамической системе

$$\begin{aligned}\dot{x} &= f(x, y) + S + I + \eta, & f(x, y) &= 3x - x^3 + 2 - y, \\ \dot{y} &= \epsilon g(x, y), & g(x, y) &= \alpha(1 + \tanh(\frac{x}{\beta})) - y,\end{aligned}\quad (1)$$

где S определяет влияние (общий вклад) связей в ОНС, I представляет собой входные данные (определяется характеристиками изображения), η это Гауссовый шум, и ϵ, α, β это положительно определенные параметры, контролирующие форму предельного цикла осциллятора [27];

2. Осциллятор предельного цикла. Данный осциллятор описывается 2-мя нейронами, возбуждающим и тормозным. Его управляющая динамическая система может быть выписана в следующем виде

$$\begin{aligned}\dot{u} &= -u + \alpha h(u) - g(v) + I^+ \\ \dot{v} &= -v + \beta h(u) + I^-, \end{aligned}\quad (2)$$

где u и v отвечают за "потенциалы" возбуждающего и тормозного нейрона соответственно, $h(u)$ – это функция активации возбуждающего нейрона (в данном случае используется сигмоида или схожие с ней функции), $g(v)$ – функция активации тормозного нейрона (используется любая монотонная неубывающая функция) и I^\pm отражают признаки входных данных для возбуждающего и тормозного нейрона соответственно. Стоит заметить, что динамическая система обладает устойчивым предельным циклом,

при условии, что значения признаков входных данных I^\pm принадлежат ограниченному интервалу.

3. Модифицированный осциллятор Гинзбурга-Ландау. Осциллятор задается двумя значениями $u_1, u_2 \in \mathbb{R}$. Тогда двумерная динамическая управляющая система представляется в виде одного уравнения с использованием комплексной переменной $u = u_1 + iu_2, u \in \mathbb{C}$

$$\dot{u} = f(u) + g(I), \quad f(u) = (p_0^2 + i\omega - |u - c|^2)(u - c), \quad (3)$$

где p_0 и $c = c_x + ic_y$ константы, задающие параметры предельного цикла осциллятора при $g(I) = 0$. Этот предельный цикл можно описать как круг радиуса p_0 , расположенный в комплексной плоскости (u_1, u_2) с координатами (c_x, c_y) при $g(I) = 0$. ω это собственная частота осциллятора. Функция $g(I)$ зависит от признаков входного изображения. При этом она контролирует размер и местоположение предельного цикла осциллятора в комплексной плоскости (u_1, u_2) . По размеру цикла можно определять яркость I пикселя, за который отвечает осциллятор, больше чем предопределенно заданное пороговое значение I^* . В этом случае мы можем считать состояние осциллятора как "активное". В противном случае предельный цикл осциллятора сходится в точку стабилизации (стабильный фокус). Состояние данного осциллятора определяется как "пассивное" или "неактивное".

Далее перейдем к описанию самих моделей.

Сеть фазовых осцилляторов Курамото-Сакагучи. Данная модель может применяться как в сетях с глобальными или локальными связями, так и в системах с центральным элементом [26]. В общем случае модель взаимодействия фазовых осцилляторов можно описать уравнением (4) (модель Курамото), а в частном случае (функция взаимодействия является синусоидальной, а также добавлены фазовые сдвиги) уравнением (5) (Модель Курамото-Сагучи)

$$\frac{\partial \theta_i}{\partial t} = \omega_i + \sum_{j=1}^n K_{ij} f(\theta_j - \theta_i), \quad i = 1, \dots, n, \quad (4)$$

$$\frac{\partial \theta_i}{\partial t} = \omega_i + \sum_{j=1}^n K_{ij} \sin(\theta_j - \theta_i - \alpha_{ij}), \quad i = 1, \dots, n, \quad (5)$$

где θ_i - фаза i -го осциллятора, ω_i - собственная частота осциллятора, K_{ij} можно интерпретировать как "веса" направленных связей между осцилляторами j и i (как сильно осциллятор j влияет на осциллятор i). f , в общем случае периодическая нечетная функция, определяет взаимодействие между осцилляторами.

Для описания модель с центральным осциллятором (ЦО) потребуется 2 уравнения

$$\frac{\partial \theta_0}{\partial t} = \omega_0 + \frac{A}{n} \sum_{j=1}^n f(\theta_j - \theta_0 + \gamma), \quad (6)$$

$$\frac{\partial \theta_i}{\partial t} = \omega_i + Bg(\theta_0 - \theta_i + \delta), \quad i = 1, \dots, n, \quad (7)$$

где A, B также можно интерпретировать как "веса" (параметры взаимодействия). В уравнении (6) записана модель для ЦО. Через этот осциллятор происходят все взаимодействия в системе. Уравнение (7) задает модель периферических осцилляторов (ПО). Следует отметить, что полученная модель для ЦО и ПО дает возможность выполнять подбор параметров для достижения определенного режима работы системы (полная или частичная синхронизация с ЦО)

Сеть фазовых осцилляторов Хонг-Строгаца. Данная модель является обобщением модели с ЦО и ПО в том случае, если присутствует синхронизирующее воздействие от ПО к ЦО и десинхронизирующем воздействии от ЦО к ПО.

Она состоит из двух популяций фазовых осцилляторов. Осцилляторы первой популяции получают из входных данных только синхронизирующие воздействия, а осцилляторы второй группы только десинхронизирующие воздействия соответственно. В результате такой организации взаимодействия между осцилляторами представители первой группы стремятся синхронизироваться с осцилляторами, которые оказывают на них воздействие, а вторая группа наоборот, старается работать в противофазе с воздействующими на них осцилляторами. Назовем первую группу фазовых осцилляторов конформистами, а вторую неконформистами.

Модель Хонг-Строгаца можно описать формулами, схожими с (4). Обозначим весь набор осцилляторов как J . Весь набор состоит из n элементов. Пусть J_1, J_2 – группы конформистов и неконформистов, соответственно. Теперь запи-

шем динамику модели

$$\frac{\partial \theta_i}{\partial t} = \omega + \frac{A}{n} \sum_{j=1}^n f(\theta_j - \theta_0 + \gamma), \quad (8)$$

$$\frac{\partial \theta_i}{\partial t} = \omega + \frac{A}{n} \sum_{j=1}^n f(\theta_j - \theta_0 + \gamma), \quad (9)$$

где k_1, k_2 , аналогично предыдущим моделям, параметры ("веса") связей. При этом $k_1 > 0, k_2 < 0$, что отражает различие в принципе работы двух групп осцилляторов.

После рассмотрения базовых моделей ОНС перейдем к построению многослойной архитектуры ОНС, которая сможет решать задачу сегментации целевых объектов на гиперспектральном изображении. Для решения этой задачи возможны два подхода, которые будут влиять на архитектуру многослойной ОНС. Обозначим первый подход как базовый, а второй как продвинутый:

1. Априорный выбор значимых (в рамках признаков целевого объекта) спектральных каналов и применение трехслойной модели сегментации для 2-х мерного изображения.
2. Добавление спектральных каналов как 3-ую размерность во входные данные и модификация модели из 1-го пункта для обработки 3-х мерных изображений.

Для базового подхода модель будет состоять из модуля селективного внимания, модуля выделения границ объекта (контуров) и модуля сегментации объектов. (написать побольше про модули)

2.3 Обучение осцилляторных нейронных сетей

В силу того, что ОНС по своей архитектуре являются рекуррентными нейронными сетями (РНС), то для их обучения могут применяться обучающие правила и подходы, разработанные для обучения без учителя. В данном параграфе будут рассмотрены обучающие правила Хеббиана, Сторки, а также принцип "победитель получает все" (ППВ).

Правило Хеббиана. Наиболее простая форма правила Хеббиана может быть описана как процесс изменения (эволюции) матрицы связи $\hat{W} = [W_{ij}]$ при в условиях

дискретного времени:

$$W_{ij}(t) = W_{ij}(t-1) + \Delta W_{ij}(t), \quad \Delta W_{ij}(t) = \eta y_i(t)x_j(t), \quad (10)$$

где $x_j(t)$ – состояние входного нейрона, $y_i(t)$ – состояние выходного нейрона и η ($\eta \geq 0$) параметр, отвечающий за скорость обучения (от англ. learning rate).

Правило (10) имеет недостаток в виде экспоненциального увеличения состояния $y_i(t)$ из-за повторяющихся обучающих серий, примененных к входному нейрону $x_j(t)$. Поэтому были предложены улучшенные версии этого правила, например:

$$\frac{\partial \hat{W}}{\partial t} = \eta(\mathbf{y}\mathbf{x}^T - \mathbf{y}^T\mathbf{y}\hat{W}). \quad (11)$$

Данное правило записано в условиях непрерывного времени.

Далее рассмотрим специальный вариант правила Хеббиана, а именно "соревновательное обучение". Смысл этого правила состоит в том, что выходные нейроны соревнуются между собой за входные нейроны. Выходной нейрон, который показал наибольший уровень активации, побеждает и его связь изменяется в то время как остальные нейроны остаются без изменений. Такую стратегию называют принципом "победитель получает все" (от англ. winner-takes-all) и он будет подробнее рассмотрен далее. Соревновательный алгоритм обучения, который подходит для использования в многослойных рекуррентных нейронных сетях может быть формализован следующим образом:

$$\Delta W_{ij}(t) = \begin{cases} \eta[x_i(t) - W_{ij}(t)] & \text{если } j\text{-ый нейрон выигрывает} \\ 0 & \text{иначе} \end{cases} \quad (12)$$

Существенной особенностью любого обучения без учителя (в частности, обучения Хеббиана и его обобщений) является автоматический, самоорганизующийся характер изменения структуры сетевой связности.

Правило Сторки. Данное правило схоже с правилом Хеббиана, но позволяет хранить больше запомненных образов и более устойчиво к зависимым образам [28]. Запишем формулу для этого правила в условиях дискретного време-

ни:

$$W_{ij}(t) = W_{ij}(t-1) + \frac{1}{n}(y_i(t)x_j(t) - y_i(t)h_{ji}(t) - h_{ij}(t)x_j(t)) \quad (13)$$

$$h_{ij}(t) = \sum_{k=1, k \neq i, j}^n W_{ik}(t-1)y_k(t) \quad (14)$$

Принцип ППВ. Принцип "победитель получает все"(от англ. Winner-Take-All, WTA) в контексте обучения нейронных сетей относится к ситуации, когда только один нейрон или группа нейронов активируется в ответ на входные данные. Этот нейрон или группа нейронов считается "победителем"и получает возможность обучаться на этих данных, в то время как остальные нейроны остаются неактивными. Этот принцип часто используется в моделях, основанных на конкуренции, где нейроны соревнуются за право реагировать на входные данные. Например, в самоорганизующихся картах Кохонена, нейроны соревнуются за то, чтобы стать "победителем и только победитель обновляет свои веса в процессе обучения.

Для реализации модулей внимания, упомянутых в предыдущем разделе, требуется разработать правило, которое учитывало бы особенности биологической модели внимания. В этом случае принцип ППВ отлично подходит на эту роль, так как он обеспечивает фокусировку внимания на одном объекте. При этом случаи отсутствия фокусировки (победителя не оказывается) возникают достаточно редко. Если перенести этот принцип на разработанную модель с ЦО и ПО, то он обеспечивает, что синхронизирован с ЦО будет лишь один ПО. Следует отметить, что для реализации принципа ППВ необходимо применять его в сети с динамическими (адаптирующимися) параметрами. В нашем случае собственная частота осциллятора, а также величины связи между нейронами будут динамическими переменными. Тогда динамику модели можно описать следующими уравнениями

$$\frac{\partial \theta_0}{\partial t} = \omega_0 + \frac{1}{n} \sum_{i=1}^n a_i f(\theta_i - \theta_0), \quad (15)$$

$$\frac{\partial \theta_i}{\partial t} = \omega_i + bg(\theta_0 - \theta_i), \quad i = 1, \dots, n, \quad (16)$$

$$\frac{\partial \omega_0}{\partial t} = \frac{\alpha}{n} \sum_{i=1}^n a_i f(\theta_i - \theta_0), \quad (17)$$

$$\frac{\partial a_i}{\partial t} = \beta(-a_i + c + \gamma h(\theta_i - \theta_0)), \quad i = 1, \dots, n, \quad (18)$$

где $\theta = (\theta_0, \theta_1, \dots, \theta_n)$, $\omega_0 \in \mathbb{R}$, $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ – переменные. Уравнение (17) задает адаптацию собственной частоты ЦО по направлению к его текущей частоте, α – параметр скорости адаптации. Параметр β отвечает за скорость изменения амплитуд, а параметры c и $c + \gamma$ задают минимальный и максимальный уровень амплитуды ПО.

Генетический алгоритм. Это метод оптимизации и поиска решений, который имитирует процессы естественной эволюции, такие как наследование, мутация, отбор и кроссовер (рекомбинация).

Генетический алгоритм начинается с генерации начальной популяции решений, обычно представленных в виде хромосом или строк кода. Каждое решение оценивается с помощью функции приспособленности, которая определяет, насколько хорошо оно решает поставленную задачу. Затем алгоритм входит в цикл, который повторяется до тех пор, пока не будет достигнут критерий остановки, например, достаточно хорошее решение или определенное количество поколений. В каждом поколении происходят следующие шаги [29]:

1. **Отбор.** "Особь" отбираются для создания нового поколения. Отбор происходит на основании значения функции фитнеса. Эта функция показывает, насколько пригодна данная "особь" для решения конкретной задачи. Выбор корректной функции фитнеса является одним из ключевых этапов разработки генетического алгоритма. Для отбора есть несколько различных методов, например рулетка или турнирный.
2. **Кроссинговер (рекомбинация).** С помощью двух ДНК (генов) выбранных "особей" создается новая особь с измененным набором генов. Полученные ей гены зависят от методов, которыми осуществляется рекомбинация. Различают дискретную рекомбинацию и кроссинговер. Кроссинговер, в свою очередь, делится на одноточечный, двухточечный и многоточечный. Также важную роль играет способ выбора родителей. Их можно выбрать случайным образом, либо наиболее близких или далеких по набору генов. Расстояние между наборами генов считается по метрике, применимой к пространству ДНК "особей".
3. **Мутация.** Это заключительный этап в рамках одного поколения. Мутация позволяет существенно изменить набор генов отдельных особей. Это до-

бавляет вариативность в популяцию и помогает преодолевать локальные минимумы. Стоит отметить, что существует множество способов выбора особей для мутации, а также самих методов изменения генов.

В контексте нашей задачи генетический алгоритм обладает наиболее широкой областью применения, но проигрывает приведенным правилам и принципам в области обучения модулей ОНС. Поэтому будет рассмотрена возможность применения генетического алгоритма для выбора спектральных каналов, для которых будет производится дальнейшая обработка.

2.4 Выводы по разделу 2

- 1.

ЗАКЛЮЧЕНИЕ

Все супер пупер работает

ПРИЛОЖЕНИЕ А