

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

кандидата технических наук Ронкина Михаила Владимировича на диссертационную работу Дмитриева Никиты Владимировича на тему «Повышение точности и производительности алгоритмов анализа и обработки сложноструктурных изображений», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

1. Актуальность темы исследования

Задача автоматического получения информации о пространственно-распределённых объектах на растровых изображениях, сочетающих в себе числовые и графические данные (например, диаграммы, чертежи, схемы, архитектурные проекты, карты местностей), является актуальной. Особенности таких изображений являются: большое количество «семантических слоёв» и необходимость максимального точного учета небольших деталей. Указанные изображения охарактеризованы автором как сложноструктурные изображения (ССИ). В настоящее время одни из наиболее перспективных способов анализа ССИ основываются на использовании современных методов машинного обучения, в том числе глубоких нейронных сетей. Однако при подробном анализе указанных методов часто возникают задачи обеспечения точности и других характеристик получения информации в отдельных задачах и для отдельных объектов на изображениях, например, объектах типа «цифра». Также на практике часто возникает задача достижения компромисса точность-вычислительная сложность, что часто не рассматривается в научной литературе. Исследованию данных вопросов посвящена рецензируемая работа.

Таким образом, можно заключить, что диссертационное исследование Дмитриева Н. В. выполнено на актуальную тему, обладает новизной и имеет существенное теоретическое, методологическое и практическое значение.

2. Научная новизна полученных результатов

В диссертационной работе получены следующие новые научные результаты:

1. На основе проведённого анализа современного состояния и развития подходов обработки цифровых изображений, предложена классификация сложноструктурных изображений (ССИ), отличающихся использованием явных и неявных правил отображения распределённых образов различной размерности, и методика построения алгоритмического и программного обеспечения анализа и обработки таких изображений (соответствует п. 7 паспорта специальности).

2. Разработаны алгоритмы сегментации, локализации, распознавания и группировки точечных, линейных и площадных объектов на сложноструктурных изображениях, основанные на комплексном использовании известных методов анализа и обработки изображений (соответствует пп. 4, 5 и 12 паспорта специальности).

3. Разработан алгоритм комплексного анализа распознанных числовых отметок высот и горизонталей цифровых топографических карт, отличающийся представлением контурной карты в виде орграфа и способный устранять ошибки предшествующего этапа распознавания (соответствует пп. 4, 5, 12 паспорта специальности).

Результат анализа содержания диссертационной работы и публикаций автора подтверждает обоснованность научной новизны представленных положений.

3. Обоснованность и достоверность научных положений, сформулированных в диссертации

Достоверность полученных результатов, научных положений и выводов, изложенных в работе, подтверждается использованием известных математических методов, адекватной постановкой цели и задач исследования, а также их согласованностью с результатами, полученными другими авторами. Автором изучены и проанализированы достижения ученых, опубликованных в работах различных изданий, представлено состояние предметной области и сравнения достоинств и недостатков существующих методов. Список использованных источников состоит из 275 наименований.

По теме диссертации опубликовано 17 работ, из них 5 статей опубликованы в рецензируемых научных изданиях, определенных ВАК РФ и Аттестационным советом УрФУ, включая 1 статью в издании, индексируемом в международной цитатно-аналитической базе Scopus. Основные результаты диссертации докладывались и обсуждались на конференциях и семинарах различного уровня.

Корректность работы разработанного автором программного комплекса подтверждается использованием современных методов и средств при проектировании, разработке и тестировании программного комплекса. На разработанный программный комплекс автором работы получено 2 свидетельства о регистрации программы для ЭВМ.

4. Практическая значимость результатов диссертации

Практическая значимость подтверждена внедрением полученных результатов в деятельность АО «Урало-Сибирская Гео-Информационная компания» и ФГБОУ ВО «Уральский государственный университет путей сообщения».

Дмитриевым Н.В. был разработан программный комплекс для преобразования сканированных изображений топографических карт в цифровую модель рельефа, в котором реализованы созданные алгоритмы обработки объектов сложноструктурных изображений различных размерностей.

5. Оценка содержания диссертации и её оформления

Диссертационная работа содержит 200 страниц основного текста (всего 222 с.), 35 рисунков и 40 таблиц. Работа состоит из введения, четырех глав, заключения, списка сокращений и обозначений, списка литературы, 3 приложений.

Во введении приводится анализ актуальности, состояния области исследования, цели и задач работы, формулируются положения, выносимые на защиту, и их новизна, указываются теоретическая и практическая значимость, а также результаты внедрения.

В первой главе представлен анализ современного состояния предметной области, основанный на результатах исследований отечественных и зарубежных учёных. Автором рассмотрены основные определения и понятия, используемые в диссертационной работе. Описаны ССИ, представляющие собой объект исследования, и приведена их классификация, включая подклассы. Показано, что объекты исследования относятся к подклассу «Слабо формализованные документы: топографические карты». Для исследуемой категории приведено определение как растровое изображение, которое содержит объекты, относящиеся к различным семантическим планам, и которое обладает набором явных и неявных правил отображения этих объектов на визуальный план с помощью различных техник в виде распределённых образов, подразумевающих возможность однозначного восстановления смысла исходных объектов и их атрибутов. Также особенностями исследуемых изображений являются «артефакты», вызванные как физическими изменениями в документе (появление потёртостей, пятен), так и сканированием (например, шумов от фотодетектора, «скачков» цветовых векторов между соседними пикселями из-за дискретизации, засвеченности и т. д.).

В качестве известных этапов анализа изображений отмечены операции цифровой предобработки изображений, статистической фильтрации (удаление шумов), кластеризации изображений (с целью семантической сегментации), распознавания образов на основе их дескрипторов. Отмечается, что возможны два подхода к распознаванию образов, названные в работе «оптический» и «интеллектуальный».

Среди оптических методов распознавания образов выделяются методы на основе признаков. Отмечается, что такие подходы наиболее стабильны при влиянии шумов. В том числе к данной группе автор относит методы на основе нейронных сетей, в том числе

сверточных нейронных сетей. Отмечается, что также возможны гибридные подходы, заключающиеся, например, в объединении результатов работы нескольких нейронных сетей. Автор отмечает возможности создания гибридных подходов при помощи таких приемов, как бустинг, генетические алгоритмы или алгоритма имитации отжига.

В главе также перечислены основные определения искусственных нейронных сетей. Приведен анализ известных в научной литературе методов ручного признакового описания топографических изображений и методов принятия решений на их основе. А также приведен обзор известных в научной литературе методов на основе искусственных нейронных сетей и других алгоритмов. Отдельный раздел обзора посвящен анализу сцен, то есть пониманию отношений объектов на изображениях.

Также в первой главе приводится раздел «Оценка алгоритмов». Отмечается, что в рабочих условиях качество распознавания средства измерения может быть оценено только на последнем этапе обработки изображений как отношение правильно распознанных объектов к количеству объектов на карте. Отмечается, что при анализе «качества распознавания» следует анализировать как т.н. матрицу ошибок, так и ее разброс «от запуска к запуску» обучения – то есть стабильность обучения нейронной сети. Также отмечается, что последний параметр может быть улучшен при помощи, например, метода беггинга. Также автор указывается на такие известные метрики работы с объектами на изображениях, как мера Жаккара и Intersection over Union (IoU), а также на необходимость измерения времени работы алгоритмов.

В результате проведенного литературного анализа автор делает выводы о том слабой разработанности темы анализа изображений типа топографических карт, которые характеризуются наличием «графики» различного типа и гетерархичности семантических планов. Автор указывает на отсутствие разработанности в научной литературе проблем совместной обработке образов, относящихся к различным семантическим полям, а также проблем наличия неявных правил отображения объектов или анализа влияния нарушения этих правил. Также не разработанной является проблема достижения компромисса «точность-качество работы» для исследуемого класса изображений. При этом открытым является вопрос стабильности качества известных в литературе «интеллектуальных моделей». Стабильными по определению являются линейный регрессор, k-NN классификатор и SVM, для всех типов ANN анализ стабильности является открытой проблемой. Также автор отмечает недостаточное наличие в известной литературе наборов данных типа топографические изображения. При этом указывается, что известные в литературе наборы данных часто имеют не точную разметку.

В результате проведенного анализа автор формулирует цель работы как повышение точности и производительности алгоритмов анализа и обработки сложноструктурных изображений с помощью объединения различных подходов обработки изображений.

Вторая глава посвящена анализу ССИ. Предложена общая методика обработки ССИ и разработан набор алгоритмов, выполняющих отдельные операции данной методики. Разработаны критерии анализа качества реализации алгоритмов с точки зрения скорости и точности. В главе приводится попытка формализации проблемы наличия в ССИ «неявных правил» (многозначных отношений). Для упрощения работы с изображениями предлагается их декомпозиция с целью уменьшения их рассматриваемых размеров и, следовательно, влияния их друг на друга. Для решения указанной проблемы предлагает следующий набор действий, представляющий собой «методику обработки ССИ»:

1. Классификация документа – разные типы могут иметь разную обработку.
2. Предобработка документа – денойзинг, устранение искажений и т.д.
3. Сегментация – автор предлагает использование подхода многоклассовой семантической сегментации.

4. Локализация точечных, линейных и площадных объектов. Локализацию точечных и линейных объектов предложено проводить путем подбора операций морфологической обработки изображений. Подбор предложено проводить при помощи генетических алгоритмов. Локализацию площадных объектов автор предлагает реализовывать с использованием гистограмм.

5. Объединение объектов. Для точечных объектов типа цифры автор предлагает архитектуру на основе нейронной сети с одним скрытым слоем. Цифры предлагается объединять в числа при помощи кластеризации. Для линейных объектов типа контуры предлагается оригинальный алгоритм на основе эвристических свойств анализируемых изображений.

6. Внутрисемантический анализ.
7. Межсемантический анализ с возможной корректировкой значений.
8. Восстановление данных.
9. Постобработка и консолидация данных.

Этапы 6-9 в главе 2 автором не описаны.

В конце главы автор отмечает, что разработаны гибридные и итеративные алгоритмы, которые позволяют сочетать несколько подходов и уходить от узкой специализации методов. При этом указывается, что алгоритмы не способны полностью автоматизировать обработку ССИ без ошибок; некоторые алгоритмы требуют

дополнительных исследований для подбора операций, в том числе для группировки точечных и линейных объектов подбор операций остаётся прерогативой разработчика.

В третьей главе рассмотрена обработка нескольких классов ССИ и проведён экспериментальный подбор параметров алгоритмов и интеллектуальных моделей, обеспечивающий оптимизацию по критериям качества и быстродействия. Автор отмечает отсутствие в научной литературе подходящих наборов данных для тестирования релевантных точечных и линейных объектов. По этой причине автор описывает синтез собственного набора данных для решения указанных задач. Основное внимание при исследовании точечных и линейных объектов уделяется задаче распознавания символов. В ходе экспериментов по обработке площадных объектов используется открытый набор данных, для которого выбрана собственная архитектура нейронной сети.

В четвертой главе автор ставит своей задачей получение цифровой модели рельефа, максимально соответствующей рельефу реальной местности, используя изображения цифровых топографических карт. Для достижения цели автором предложен алгоритм комплексного анализа данных на основе представления линий рельефа графами и интерполяции линий. Алгоритм реализован в виде программного комплекса преобразования изображений топографических карт в цифровую модель рельефа. Работоспособность комплекса проверена на 20 примерах цветных отсканированных топографических картах.

В заключении делается вывод о достижении основных результатов согласно заявленным задачам в работе, а также даются рекомендации и перспективы развития темы исследования.

6. Замечания и вопросы по диссертации

Вместе с этим следует отметить некоторые замечания и вопросы по содержанию диссертационной работы:

1. Автор указывает на слабую разработанность темы анализа сложноструктурных изображений типа топографических карт в научной литературе, однако на стр. 7 и 152 также указывает на наличие коммерческого программного обеспечения для такого анализа. В главе 1 явных указаний на какое-либо рассмотрение доступного программного обеспечения не приведено. Возможно ли исследование и ограничения указанного программного обеспечения? Например, остается неясным, что автор считает оптическими, что интеллектуальными, а что статистическими методами. Например, на стр. 51 «ТО тяготеют к более интеллектуальным методам, ПЛО – к более статистическим». При этом на стр. 14

говорится, что в первой главе рассмотрены только оптические и интеллектуальные алгоритмы. Какие определения указанным подходам дает автор?

2. Не совсем понятен тезис о том, что автор указывает на отличия понятий паноптическая сегментация и многометочная сегментация (раздел 2.2). Тут я полагаю, что многоклассовая сегментация, более традиционная для задач паноптической и семантической сегментации, подразумевает, что каждый пиксель относится только к одному классу. Многометочная классификация подразумевает, что каждый пиксель может содержать информацию сразу о нескольких классах. Отмечу, что в современных архитектурах решения задач семантической и экземплярной сегментации многометочный подход также распространен и не предполагает нестандартной терминологии. Таким образом, возникает вопрос, какую задачу все-таки решает автор и в чем уникальность ее постановки?
3. Также возникает вопрос о причинах не рассмотрения задачи сегментации контуров. Такой подход, вероятно, позволил бы закрыть часть задач автору. Можно предположить, что комбинация подходов сегментации контуров в комбинации с архитектурой для поиска ключевых точек позволил бы решить проблемы обнаружения в совокупности точечных, линейных и площадных объектов на изображениях.
4. Необоснованным кажется тезис автора о том, что алгоритмы типа R-CNN и YOLO «локальны» (стр. 54), то есть могут применяться исключительно для нахождения точечных объектов. Например, алгоритмы Mask-R-CNN или YOLO в режиме instance segmentation принципиально работают для задач, вероятно характеризуемых автором как ПЛО.
5. Из текста на стр. 54 «сегментации, представляющую собой классификацию отдельных пикселей» можно сделать вывод о том, что автор все же рассматривает семантическую сегментацию. А выражение 2.3 явно указывает на многоклассовую, а не многометочную классификацию. Как это соотносится с тезисом из вопроса 2? Возможно, что термины многометочная классификация (в оригинале multilabel classification) и многоклассовая классификация (multiclass classification) перепутаны.
6. Не понятен тезис (стр. 55) о том, что «из-за большого размера цифровых топографических карт провести полностью операцию разметки автору не представляется возможным, поэтому необходимо обеспечить

репрезентативность выборки. Она должна отражать все возможные проявления изображения условных знаков: на разном фоне, под разными углами знаки должны иметь разное значение. Особое внимание следует уделить областям, где пересекаются объекты разных классов». Кажется, что изображение большого размера может быть разделено на несколько частей. Возможно, автор имеет в виду проблему того, как нужно разделить изображение?

7. При прочтении раздела 2.3.2 «Алгоритм для классификации площадных объектов» возникает вопрос, если автор изначально проводит семантическую сегментацию, то в чем заключается необходимость проведения повторной сегментации площадных объектов? Разделы 2.3 и 2.4 имеют скорее реферативный стиль, нежели показывают результаты исследований. Автор априорно-синтетически описывает некоторые алгоритмы решения набора задач для постобработки сегментированных изображений. При этом наиболее понятные практики использования т.н. state-of-the-art архитектур нейронных сетей в этих разделах отвергаются, ссылаясь на некоторые источники литературы или из тех же априорно-синтетических умозаключений. Современное машинное обучение является скорее экспериментальной наукой, чем сугубо теоретической. Поэтому утверждённые результаты автора в данных разделах должны доказываться путем численных экспериментов для целевых наборов данных.
8. Так, например, в разделе 2.4.1 автор предлагает свою архитектуру нейронной сети для распознавания цифр, ссылаясь на низкую точность модели Tesseract. Однако данная модель не описывает современные достижения в области задач распознавания символов на изображениях. Поэтому обоснование предложенной архитектуры должно быть иным.
9. Для проверки алгоритмов обработки точечных объектов (раздел 3.2) автор указывает на синтез всего 20 изображений. Непонятно, почему автор выбирает изображения морского дна. Также возникает вопрос о репрезентативности результатов на такой маленькой выборке. Более того, автор не показывает код для синтеза изображений и не выкладывает набор данных в открытый доступ. Также текст диссертации не отвечает на вопрос о том, почему нельзя взять открытые данные из интернета и не собрать свой набор данных. Таким образом, возникает сомнение в репрезентативности экспериментальных результатов.
10. В разделе 3.1 автор указывает: «В качестве моделей для сегментации и распознавания ТО предлагается исследовать логистическую регрессию (LR), линейный дискриминантный анализ (LDA), гауссов наивный байесовский...».

Возникает вопрос о том, каким образом автор использовал указанные алгоритмы для решения указанных задач сегментации и распознавания ГО.

11. Не ясна фраза «В дополнение к экспериментальному исследованию всего алгоритма 2.5 в целом производилось тестирование только его второй части, отвечающей за постобработку, используя более мощное оборудование для реализации ресурсозатратной архитектуры DeepLabV3+ [129] в качестве интеллектуальной модели для первичной сегментации». Алгоритм 2.5 (стр. 79) не содержит деления на части, не предполагает использование нейронных сетей и не содержит этап первичной сегментации.
12. Результаты, представленные в таблицах 3.5 – 3.7, вызывают сомнения. Не совсем понятно, для какого объема выборки получены указанные результаты и имеются ли статистически значимые различия в цифрах для разных алгоритмов.
13. В разделе 3.3 – таблица 3.9 указывает на то, что разработанная автором архитектура нейронной сети имеет значение метрики IoU 56%. Однако, следует также дать обоснование достаточности этого значения для решения прикладных задач.
14. В разделе 4.5 автор показывает результаты работы разработанного программного комплекса на 20 примерах цифровых отсканированных карт. Однако, автор не указывает на то, как была проведена разметка карт для их тестирования.

7. Общая характеристика работы

Исследования проведены в значительном объеме, логически выстроены, включают как теоретические аспекты, так и экспериментальную проверку.

Сформулированные вопросы и замечания имеют дискуссионный характер и не снижают общего впечатления от проделанной работы.

В целом диссертационная работа Дмитриева Никиты Владимировича хорошо и логично структурирована, соответствует требованиям, предъявляемым к научным работам и паспорту специальности 2.3.1 «Системный анализ, управление, обработка информации и статистика». Иллюстрации выполнены на высоком научном и оформительском уровне. Автореферат соответствует диссертации.

8. Заключение по работе

Исследовательская работа изложена грамотным научно-техническим языком, в полной мере отвечает требованиям по актуальности, научной новизне, практической значимости, личному вкладу автора, отражению результатов в публикациях, а также полностью соответствует п. 9-14 Положения о присуждении ученых степеней в УрФУ и специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Автор диссертации Дмитриев Никита Владимирович заслуживает присуждения учёной степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Официальный оппонент:

Кандидат технических наук

Доцент кафедры информационных технологий и систем управления Института радиоэлектроники и информационных технологий,

Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет им. Первого Президента России Б.Н. Ельцина»,

Тел.: +7 (343) 375-45-81

e-mail: m.v.ronkin@urfu.ru

Адрес: 620062, г. Екатеринбург, ул. Мира, 32

Ронкин Михаил Владимирович

16.01.2025



(подпись) (дата)

М.П.

ПОДПИСЬ
ЗАВЕРЯЮ.

Ронкина М.В.

УЧЕНЫЙ
МОРЗО

