

ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Арлазарова Владимира Викторовича

«Мобильное распознавание и его применение к системе ввода идентификационных документов», представленную на соискание ученой степени доктора технических наук по специальности

2.3.1 – «Системный анализ, управление и обработка информации, статистика»

Актуальность

Системы автоматизированного ввода документов на сегодняшний день широко применяются на практике. Среди них выделяются системы, работающие с идентификационными документами, содержащими персональные данные, поскольку к ним предъявляются повышенные требования по качеству работы и информационной безопасности. Диссертация В.В. Арлазарова посвящена теоретическим и практическим принципам и решениям для распознавания идентификационных документов на мобильных устройствах. Распознавание непосредственно на пользовательских устройствах позволяет решить большую часть проблем с безопасностью персональных данных, однако значительно усложняет задачу построения таких систем, поскольку качество изображений, снятых в произвольных неконтролируемых условиях, варьируется в широких пределах. Кроме того, мобильные устройства, как правило, обладают ограниченными вычислительными ресурсами и требуют использования эффективных алгоритмов обработки изображений и распознавания образов.

Таким образом, тема диссертации В.В. Арлазарова является крайне актуальной и в научной, и прикладной сферах.

Содержание работы

Диссертационная работа состоит из введения, 5 глав, заключения и 2 приложений. Текст изложен на 358 страницах, список литературы содержит 350 наименований.

Во введении раскрывается тема диссертационной работы, обосновывается ее актуальность, формулируются цели и задачи исследования.

В первой главе выполнен анализ существующих подходов к построению распознающих систем; проведено исследование современных алгоритмов обработки изображений, применяемых в задачах поиска, классификации и анализа документов. Проведен обзор существующих средств для оценки качества работы систем распознавания идентификационных документов и сформулирована проблема отсутствия репрезентативных открытых данных для экспериментов и тестирования.

Во второй главе ставится задача распознавания идентификационных документов, и проводится анализ архитектуры программных комплексов, предназначенных для ее решения. Систематизируются особенности и проблемы распознавания на мобильных устройствах в неконтролируемых условиях и рассматриваются пути их решения. Кроме того, в главе ставится задача проверки подлинности документа, которая возникает при распознавании на пользовательском устройстве, и предлагается ряд методов ее решения. Например, в случае паспортов РФ и некоторых других документов ключевую информацию о подлинности можно получить с помощью анализа подписей и печатей. Для этого В.В. Арлазаров предлагает метод на основе локализации оттисков печатей, их нормализации и последующего распознавания визуальной информации. Другой предложенный способ заключается в проверке способа нанесения текстовой информации. Как правило, для идентификационных документов помимо стандартных вариантов печати используются и специализированные способы, такие как лазерная гравировка, лазерная перфорация, эмбоссирование символов и т.д. В.В. Арлазаров предложил новый способ выявления аномалий начертания текста на основе нейронных сетей.

Предложенная в диссертации оригинальная архитектура системы распознавания содержит три больших этапа. На первом этапе выполняется оценка качества изображения, поиск документа на изображении и определение конкретного типа документа в кадре. Этот этап опирается на алгоритмы

обработки изображений: нормализацию изображения, нахождение особых точек, их дескрипторов и поиск стабильного сопоставления точек изображения и шаблона документа. На втором этапе выполняется анализ найденного документа: на нем выделяются зоны, содержащие целевые объекты, например, поля, содержащие текстовую информацию. При этом особое внимание уделяется не определению границ и точного местоположения объектов, а поиску зоны документа целиком.

Финальный этап работы системы включает в себя определение значения найденных объектов. Сюда относятся алгоритмы использования избыточной информации, интеграция результатов распознавания видеопотока, и распознавание элементов, специфических для ID документов, таких как подписи и печати.

Таким образом, во второй главе диссертации приводятся ключевые результаты работы и предлагается ряд новых алгоритмов и методов.

Третья глава посвящена развитию средств распознавания на мобильных устройствах. Такие устройства могут получать в качестве входных данных последовательность кадров с изображением одного и того же объекта – видеопоток. Эти данные В.В. Арлазаров предлагает использовать для повышения качества распознавания. Такой подход позволяет компенсировать блики и анализировать голографические элементы на поверхности документов, а также улучшить результаты поиска и распознавания другой информации. В этой главе рассмотрены две задачи: задача оптимального комбинирования результатов распознавания разных кадров, а также задачу остановки распознавания. Отмечу, что полученные новые теоретические и практические результаты решения этой задачи имеют значение не только при распознавании идентификационных документов, но также могут применяться при распознавании других типов объектов в видеопотоке.

В четвертой главе рассмотрен процесс создания наборов тестовых данных для систем распознавания идентификационных документов. Проблема состоит в том, что реальные документы не могут использоваться для открытой публикации

и ограниченно доступны для обучения и проведения научных экспериментов. При этом есть ряд документов с близкими шаблонами, для корректной настройки которых требуется много различных обучающих примеров. Поэтому В.В. Арлазаров предлагает алгоритмы создания реалистичных искусственных документов, которые позволяют добиться высокого итогового качества распознавания. Полученные наборы данных были опубликованы в открытом доступе: MIDV-500, MIDV-2019, MIDV-2020, MIDV-LAIT, MIDV-Holo.

В пятой главе приводятся: описание реализованного программного комплекса для распознавания идентификационных документов Smart ID Engine, его структура, а также экспериментальная оценка вычислительной эффективности на различных мобильных устройствах и стационарных отечественных платформах.

Степень обоснованности научных положений, выводов и рекомендаций, их достоверность

Положения и выводы диссертации детально описаны и обоснованы. Достоверность результатов подтверждается использованием методов системного анализа, математической статистики, а также результатами численных экспериментов, которые находятся в согласии с результатами других авторов. Результаты работы изложены в 40 публикациях, 30 из которых – в изданиях, индексируемых Web of Science и Scopus, причем 10 работ опубликованы в журналах Q1 и Q2. В журналах из списка ВАК опубликовано 20 работ. Результаты диссертации докладывались и обсуждались на профильных международных конференциях.

Новизна полученных результатов

Диссертационной работе В.В. Арлазарова оригинальна, в ней представлен авторский взгляд на решение проблемы распознавания и интерпретации видеопотока цифровых изображений идентификационного документа, полученного с использованием мобильных вычислительных устройств (в т.ч. смартфонов).

Основными научными результатами оппонент считает следующие:

1. Системный анализ процессов регистрации, распознавания и интерпретации видеопотока цифровых изображений идентификационного документа.
2. Системный подход к созданию систем распознавания и интерпретации идентификационных документов на основе алгоритмов компьютерного зрения на мобильных устройствах, учитывающий их ограниченные вычислительные возможности.
3. Алгоритмы решения задач локализации, идентификации и интерпретации цифровых изображений идентификационных документов, полученных с камер(ы) мобильных вычислительных устройств в неконтролируемых условиях съемки.
4. Метод использования видеопотока для повышения качества распознавания и интерпретации цифровых изображений идентификационного документа, стратегии агрегирования результатов покадрового распознавания и решение проблемы останова процесса распознавания и интерпретации.
5. Системный подход к синтезу пакетов данных для обучения и тестирования систем распознавания и интерпретации идентификационных документов, разработанные MIDV пакеты данных для оценки качества систем распознавания идентификационных документов.
6. Алгоритмы обнаружения подделок и фальсификаций в идентификационных документах.
7. Прикладные программные средства распознавания идентификационных документов (Smart ID Reader, Smart ID Engine).

Замечания по содержанию и оформлению

В диссертационной работе В.В. Арлазарова можно отметить следующие недостатки:

1. В работе отсутствует унифицированная система обозначений, обозначения одних и тех же понятий меняются в процессе изложения, присутствуют

одинаковые обозначения для различных понятий, что затрудняет чтение и понимание работы. Например:

- стр.106: A – алфавит символов, a – символ алфавита, s – слово, p и $P(.)$ – вероятности; величина j – используется как индекс в (2.33), а в (2.34) и ниже – как символ алфавита A ;
- стр.107 в формуле (2.24) величина a обозначает вероятность;
- стр. 148: Z – алфавит символов, s – символ алфавита, A^k – знакоместо, a_i – некоторый индекс, X_i – вероятности.

2. Присутствуют формальные ошибки в ряде выражений. Например:
 - пропали условия в выражениях для условных вероятностей в формулах (2.18) и (2.19);
 - в левой части равенства (2.22) не должно быть “ $=s$ ” (выше и ниже в изложении оно корректно не указано);
 - в (2.39) первое равенство следует заменить на “ $\leq$ ” в соответствии с (2.36);
 - потеряна переменная суммирования в знаменателе формулы (3.26).
3. Подраздел 2.9 назван «Модель универсальной системы распознавания документов...», однако сама модель в нем не представлена - указаны особенности систем распознавания документов (п.2.9.1) и «список предложений, которым следует придерживаться» (п.2.9.2).
4. Выражение (3.26) названо «аналогичным» выражению (3.20). Требуется пояснить, в чем заключается аналогия, поскольку эти выражения идентичны (при устранении указанной выше ошибки в (3.26)).
5. Одним из результатов подраздела 3.3.2 «Моделирование потока результатов...» является представленный в таблице 11 «Результат проверки согласия моделей...для двух... альтернатив...», в котором указаны 12 чисел с точностью до 4 знаков после запятой. В то же время в этом подразделе не представлена информация о том, как именно ставился эксперимент (процесс получения данных) и какой объем данных был использован.
6. В диссертации автор предлагает интегрировать данные с кадров видеопотока посредством интеграции только результатов распознавания/интерпретации этих кадров. При этом не уделено должного

внимания альтернативному решению, заключающемуся в реконструкции и повышении качества (в том числе, повышения разрешения) статичного цифрового изображения идентификационного документа, которое могло бы существенно упростить процесс его распознавания/интерпретации.

Указанные недостатки не снижают общей высокой оценки диссертационной работы, ее научную ценность и практическую значимость.

Общая оценка работы

Диссертационная работа В.В. Арлазарова является завершенным научным исследованием, хорошо структурирована и обладает внутренним единством. Приведенные результаты обоснованы, а положения, выносимые на защиту, обладают научной новизной и отражают результаты работы.

Результаты диссертации опубликованы в высокорейтинговых научных журналах и трудах международных конференций. Разработанные подходы и алгоритмы реализованы в программных решениях для распознавания документов, используемых в ряде организаций. Следует отметить впечатляющие масштабы разработки и внедрения (в приложении к диссертационной работе представлены 9 актов внедрения; 23 патента, включая 2 американских). Комплекс способен распознать более 2000 типов идентификационных документов на различных языках, демонстрируя достаточно высокое качество. На его основе разработано множество реально используемых систем, установленных в крупнейших банках, у операторов связи, в аэропортах, кассах железных дорог и т.п. Реализация идей, заложенных в диссертации, и их практическое внедрение находится на высочайшем уровне.

Работа соответствует паспорту специальности 2.3.1 – «Системный анализ, управление и обработка информации, статистика», автореферат в полной мере отражает содержание диссертации.

Заключение

Считаю, что диссертация В.В. Арлазарова «Мобильное распознавание и его применение к системе ввода идентификационных документов» является завершенной научно-квалификационной работой, отвечает критериям, предъявляемым к диссертация на соискание ученой степени доктора наук, содержит новые научно обоснованные технические, технологические или иные решения, внедрение которых вносит значительный вклад в развитие цифровой

профессор кафедры геоинформатики и информационной безопасности федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» (Самарский университет), д.ф.-м.н., доцент

Маси

«11» мая 2023 г.

+7 (846) 267-43-70

Начальник отдела сопровождения деятельности советов Самарского
университета

Беларусь

Бояркина У. В.

« 11 »

05

2023 г.

