



МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ НАУЧНЫЙ ЦЕНТР НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ СИСТЕМНЫХ ИССЛЕДОВАНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК»
(ФГУ ФНЦ НИИСИ РАН)

Нахимовский просп. 36-1, Москва, 117218 Телефон: (499) 124-97-44, Факс: (495) 719-76-81
E-mail: niisi@niisi.msk.ru <http://www.niisi.ru> ОКПО 05825395; ОГРН 1027700384909; ИНН/КПП 7727086772/772701001

УТВЕРЖДАЮ

Директор федерального государственного
учреждения "Федеральный научный центр
Научно-исследовательский институт
системных исследований Российской
академии наук"

д.т.н., С.Е. Власов

«28»



2023 г.



ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

**на диссертационную работу Арлазарова Владимира Викторовича «Мобильное
распознавание и его применение к системе ввода идентификационных
документов», представленной на соискание ученой степени доктора
технических наук по специальности 2.3.1 – «Системный анализ, управление и
обработка информации, статистика»**

Актуальность темы диссертации связана с использованием технологий распознавания образов и ввода идентификационных документов, в заданных форматах, реализованной на технической базе мобильных устройств. Мобильные устройства и переносные вычислительные комплексы все более интегрируются в повседневную жизнь людей и различные сферы правовых и экономических отношений. Обще признанным фактом является переход от стационарных систем идентификации личности к мобильным комплексам и дистанционным системам идентификации, которые могут использоваться для различных целей. Современные мобильные устройства обладают значительной вычислительной эффективностью и значительным арсеналом компонентов, реализующих ввод информации по различным каналам, фото-, видеокамеры, датчики освещенности, дальнометры и т.д.

Необходимо отметить, что существует множество готовых мобильных программных продуктов которые обеспечивают взаимодействие и доступ к различным услугам пользователей с государственными учреждениями, банками и т.д., при этом эффективность дистанционной идентификации личности и документов не достаточно развита в настоящий момент.

Существующие технологии мобильного распознавания идентификационных документов имеют свои ограничения, такие как низкая точность распознавания и длительное время обработки. Основные проблемы распознавания документов, удостоверяющих личность связаны в первую очередь особенностью их изготовления, призванную максимально усложнить возможность фальсификации документа, сложные, уникальные шрифты, ламинирование, различные способы нанесения текста и т.д. Кроме этого, в реальных условиях не возможно обеспечить идеальные условия фото-видео фиксации по яркости, контрастности, неравномерность освещения. Дополнительными факторами являются множество разнообразного оборудования, интегрируемого в мобильные устройства, матрицы фото-видео фиксации имеют различное соотношение сигнал/шум, разрешение, алгоритмы первичной обработки изображений. Все эти аспекты требуют системного анализа и выработки методов решения сложной и многокритериальной задачи.

В диссертации исследователь предлагает новый метод мобильного распознавания идентификационных документов, которые учитывают представленные выше ограничения и обеспечивают более высокую точность, эффективность распознавания и более быструю обработку данных.

Тематика диссертации имеет широкое практическое применение в различных сферах, таких как финансы, банковское дело, медицина и государственное управление. Разработка системы мобильного распознавания идентификационных документов может способствовать улучшению качества обслуживания клиентов, ускорению процессов проверки идентификации, а также повышению безопасности и защищенности данных.

Таким образом, Мобильное распознавание и его применение к системе ввода идентификационных документов на основе **новых научных результатов Арлазарова В.В.** имеет высокую теоретическую и практическую

значимость и позволяет перейти к созданию эффективных передовых разработок в области мобильного распознавания идентификационных документов и применению их в различных областях народного хозяйства.

Актуальность представленной диссертационной работы заключается в том, что представлен комплексный системный подход к разработке методов распознавания идентификационных документов, представлена авторская универсальная архитектура системы распознавания документов, удостоверяющих личность. Автором впервые проведен масштабный системный анализ методов распознавания документов, удостоверяющих личность, разработана и представлена новая вероятностная модель, предоставляющая возможность оценить результаты распознавания на кадрах видеопотока. Одним из ключевых результатов, представленных в диссертационной работе методы построения коллекций видеоизображений для создания баз данных, направленных на решение задачи обучения систем распознавания документов, идентифицирующих личность. Представленные в работе теоретические результаты были использованы в программном комплексе распознавания документов, идентифицирующих личность. На базе представленного комплекса создана серия прикладных систем, внедренных в организациях различного уровня. что является актуальной задачей в данной области исследований. Диссертационная работа Арлазарова Владимира Викторовича «Мобильное распознавание и его применение к системе ввода идентификационных документов» является актуальной.

Обоснованность и достоверность полученных результатов и выводов работы обеспечена использованием обоснованных теоретических выводов, фундаментальных информационных и математических методов, алгоритмов, включая общепризнанные информационные модели, методы математической статистики. Результаты работы подтверждены данными экспериментов и результатами внедрения.

Научная и практическая значимость.

Научные исследования и наиболее существенные научные результаты:

1. Диссертация содержит новую универсальную архитектуру системы распознавания документов, которая использует современные методы обработки изображений. С помощью алгоритмов, основанных на методах математической морфологии, Виолы-Джонса и RANSAC, система может эффективно выделять изображение документа на сложном фоне, находить соответствующий шаблон, выделять и распознавать отдельные поля документа, а также проверять подлинность документа.
2. В работе проведен обширный анализ методов использования видеопотока для распознавания документов, удостоверяющих личность. Исследователь разработал новую вероятностную модель, основанную на распределении Дирихле, которая адекватно описывает оценки результатов распознавания на кадрах видеопоследовательности и соответствует эмпирическим данным по критерию Андерсона-Дарлинга. Эта модель может быть использована для получения комбинированного результата распознавания объектов в видеопоследовательности, а также для принятия решения об остановке процесса распознавания.
3. В работе впервые построены базы видеопотока изображений документов, удостоверяющих личность, в объемах, достаточных для проведения обучения алгоритмов и контроля программных комплексов. Исследователь разработал оригинальную систему аугментации, которая позволяет создавать различные варианты изображений документов, используя различные технологии, такие как получение бумажных псевдодокументов и их предъявление в видеопотоке. Созданные базы данных видеопотока обеспечивают необходимый объем данных для обучения алгоритмов распознавания документов, удостоверяющих личность, и для тестирования программных комплексов.
4. В работе представлен разработанный уникальный программный комплекс распознавания идентификационных документов. На основе комплекса создана серия прикладных систем, которые были

внедрены и внедряются в организации различного уровня. Системы используются как в России, так и за рубежом. Разработанный комплекс обеспечивает высокую точность распознавания документов, удостоверяющих личность, и может быть использован в различных областях, где требуется быстрое и точное распознавание идентификационных документов.

С учетом поставленной цели, решаемых задач и полученных результатов рассматриваемая диссертационная работа соответствует областям исследования паспорта специальности 2.3.1 «Системный анализ, управление и обработка информации, статистика», а именно пункту 1 «Теоретические основы и методы системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта», пункту 2 «Формализация и постановка задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта», пункту 4 «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта», пункту 5 «Разработка специального математического и алгоритмического обеспечения для решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта».

Научная и практическая значимость.

В работе предложена новая постановка задачи и новые подходы к построению систем распознавания документов, которые могут быть использованы в различных областях, где требуется точное и быстрое распознавание идентификационных документов. Теоретическое обоснование разработанных методов и подходов применимо не только в рамках задач, рассматриваемых в работе, но и в других активно развивающихся направлениях распознавания документов. Это дает возможность для дальнейшего развития и улучшения методов распознавания документов, а также для создания новых систем и приложений, которые могут быть использованы в различных отраслях экономики и общественной жизни. Кроме

того, предложенные подходы и методы могут быть адаптированы для работы с различными типами документов, что расширяет область их применения.

Диссертационная работа имеет высокую практическую значимость, так как предлагаемые концепции и методы успешно применены для создания практических систем распознавания идентификационных документов. Это доказывается представленным программным инструментарием на базе решений, которые были разработаны в работе и успешно применены в большом количестве прикладных систем, работающих в организациях, использующих распознавание паспортов РФ, пластиковых карт, машиночитаемых зон, водительских удостоверений и других документов. Это говорит о том, что результаты, полученные в работе, имеют высокую практическую ценность.

Кроме того, практическая значимость работы подтверждается получением 2 патентов в США, 5 патентов на изобретение в РФ, 16 патентов на полезную модель и 4 свидетельств о государственной регистрации программы для ЭВМ, которые включают в себя результаты, содержащиеся в диссертации. Это доказывает, что результаты данной работы имеют высокую научную и практическую ценность и вызывают интерес у научного сообщества и бизнес-сектора. Полученные патенты и свидетельства о государственной регистрации программы для ЭВМ подтверждают оригинальность и новизну предложенных решений и подходов в области распознавания идентификационных документов.

Диссертация состоит из введения, 5 глав и заключения. Полный объем диссертации составляет 358 страниц, включая 104 рисунка и 27 таблиц. Список литературы содержит 350 наименований.

Во введении автор диссертационной работы обосновывает актуальность темы и новизну темы в научном контексте. Также разделе формулируются цели и задачи, которые были поставлены в рамках работы, и обосновывается теоретическая и практическая значимость, представлены основные положения, выносимые на защиту.

Первая глава посвящена обзору литературы и методов обработки изображений с детальным изучением аспекта распознавания документов,

идентифицирующих личность. Подробно рассмотрены современные методы обработки изображений документов с использованием компьютерного зрения. В рамках обзора рассмотрены задачи классификации, идентификации, локализации, анализа структуры, распознавания отдельных объектов и постобработки результатов распознавания. Кроме того, описаны сценарии применения этих методов в различных задачах, таких как извлечение значимых атрибутов, сравнение и проверка документов и автоматический ввод данных удостоверяющих личность документов.

Проведенный анализ показал, что в литературе отсутствует системное рассмотрение всех аспектов задачи распознавания документов, удостоверяющих личность. Также отмечается, что слабо изучены вопросы работы систем распознавания на современных вычислительных платформах, таких как мобильные устройства. При этом, особенной проблемой является отсутствие открытых и репрезентативных пакетов экспериментальных данных, которые бы позволяли исследователям верифицировать результаты экспериментов и промышленные образцы систем.

На основании проведенного анализа в первой главе диссертации сформулированы основные задачи работы. В целом, глава представляет собой важный обзор современного состояния исследований в области обработки изображений документов методами компьютерного зрения, а также выделяет нерешенные проблемы в этой области.

Во второй главе диссертации рассматривается проблема распознавания идентификационных документов, таких как удостоверяющие личность документы или ID-документы. Указывается, что такие документы содержат различную персональную информацию, такую как фамилию, имя, отчество, дату рождения, пол, государственный идентификационный номер, а также могут содержать специальные кодифицированные зоны и фотографию владельца. Представлен спектр проблем возникающих при распознавании документов, полученных с фото и видео устройств, например, такие как наличие разнородного фона, что существенно затрудняет процесс обработки изображений.

Проводится системный анализ архитектуры комплексов распознавания документов, удостоверяющих личность, в котором выявляются особенности документов, оказывающие влияние на процесс распознавания и особенности формирования изображений этих документов. Подчеркивается важность разработки систем распознавания документов, удостоверяющих личность, способных эффективно обрабатывать различные типы изображений и учитывать особенности таких документов.

Показано, что детектирование и локализация документов на изображениях может быть затруднено разнородным и неконтролируемым фоном. Фон, который имеет много высококонтрастных линий или локальных областей, может привести к тому, что алгоритм распознавания документа "ошибочно" распознает часть фона за часть документа, что усложняет задачу точного детектирования и локализации. Неконтролируемые условия освещения могут также создавать проблемы для детектирования и локализации документов на изображениях, а также для анализа структуры документа, распознавания текста. Кроме того, геометрическое положение документа на изображении является крайне важным фактором распознавания. На изображении, полученном с помощью камеры мобильного устройства, документ может быть повернут во всех проекциях, что усложняет задачу предварительной локализации документа. В главе рассматриваются подходы, основанные на поиске ключевых точек и сопоставлении их дескрипторов, которые могут помочь решить проблему детектирования и локализации документов на изображениях при наличии перечисленных выше ограничений. Эти методы называют "Lowe-based segmentation", и они основываются на поиске ключевых точек с помощью алгоритмов, которые ищут локальные экстремумы в изображении и описывают их с помощью дескрипторов.

В диссертации показано, что на практике стало обычным использование веб-камер и мобильных устройств для получения изображений документов, что привело к возникновению нового типа входных данных - последовательности видеок кадров вместо одного статического снимка. Использование нескольких входных изображений одного и того же объекта дает дополнительные преимущества при распознавании и анализе документов,

такие как возможность применения методов фильтрации и уточнения для повышения точности обнаружения и локализации объекта, применение методов суперразрешения для получения изображений более высокого качества, а также улучшение результатов распознавания текста путем накопления покадровых результатов распознавания и их интеграции в один наиболее достоверный результат. Однако наличие в системе распознавания дополнительного элемента - модуля оценки качества входного изображения, является ключевым аспектом данной технологии. Этот модуль позволяет оценить качество входного изображения и принять решение о том, можно ли использовать его для распознавания и анализа. Разработанные и представленные методы проверки подлинности охватывают большой круг возможных фальсификаций и подделок. В результате, использование последовательности видеок кадров вместо одной фотографии дает значительные преимущества для распознавания и анализа документов, и эта технология может быть применена в различных областях, где требуется высокая точность и достоверность данных. В главе исследована вероятностная модель распознавания текстовых строк.

В диссертации изучен способ проверки подлинности документа, основанный на контроле способа нанесения текста. Удостоверяющие документы используют различные техники для обеспечения долговечности и защиты от фальсификации, включая лазерную гравировку, перфорацию и эмбоссирование символов в текстовых полях.

В заключительной части главы описывается метод создания всесторонней системы распознавания идентификационных документов с проверкой на подлинность. Этот метод основан на использовании компьютерного зрения и анализа трехмерных сцен, а также учитывает особенности как документов, так и процессов их создания.

Третья глава описывает модель системы распознавания, использующей последовательность кадров, полученных с видеокамеры, в качестве входных данных. Это позволяет решать такие задачи, как компенсация бликов, улучшение детектирования объектов и анализ объектов с изменяющимися оптическими характеристиками, например, голографических элементов

защиты. В работе с видеопотоком возникают новые задачи, такие как предварительная оценка кадров, межкадровое комбинирование и принятие решения об остановке распознавания. В главе представлены постановки этих задач, основанные на минимизации функционала убытка.

Рассмотрены различные подходы к межкадровому комбинированию результатов распознавания текстовых строк при работе с документами, удостоверяющими личность. Сравнительный анализ показал преимущества комбинирования результатов по сравнению с распознаванием единственного изображения. Выбор лучшего кадра с помощью оценки фокусировки дает наилучшие результаты в условиях близких к идеальным съемки, а метод ROVER предпочтительнее при наличии помех.

Также в главе построена новая вероятностная модель, описывающая оценки принадлежности распознаваемых символов к тем или иным классам, изменяющиеся в видеопоследовательности. Предложены три модели на основе распределения Дирихле и его вариантов и описаны методы вычисления их внутренних параметров. Рассмотрена задача остановки процесса распознавания в видеопоследовательности, как с ограничением количества обрабатываемых кадров, так и с учетом реального времени обработки, методы остановки на основе анализа характеристик результатов и вероятностных моделей. Использование правил остановки позволяет уменьшить уровень ошибки при ограничении на количество обрабатываемых кадров.

В главе уделено внимание трудоемкости алгоритмов обработки изображений и компьютерного зрения в системах распознавания на мобильных устройствах, описываются методы оптимизации вычислений в задачах обработки изображений, используемых в системах распознавания, с использованием SIMD-инструкций ARM NEON, доступных на большинстве современных мобильных устройствах.

В четвертой главе подробно описано - создание и использование пакетов данных для обучения и тестирования систем распознавания документов, содержащих удостоверяющую информацию. Рассмотрены методы синтеза данных для обучения и настройки алгоритмов распознавания, включая формирование обучающих множеств в процессе эксплуатации

системы, а также проблемы создания искусственных обучающих выборок. Описан подход аугментации образов объектов для получения репрезентативных выборок, включая алгоритм аугментации оттисков круглых печатей для построения соответствующих классификаторов и детекторов.

Помимо этого, рассмотрены методологии оценки качества работы систем распознавания документов и их компонентов. Открытые пакеты данных являются ключевым аспектом оценки их качества, так как они позволяют научному сообществу объективно сравнивать методы и алгоритмы. Однако, уже существующие пакеты данных, содержащие изображения документов, удостоверяющих личность, не всегда репрезентативны и могут быть улучшены.

В диссертации представлен системный подход к созданию таких пакетов данных, и были изготовлены и опубликованы 6 пакетов, предназначенных для анализа различных аспектов систем распознавания:

- пакет данных MIDV-500, представляющий собой коллекцию аннотированных видеопоследовательностей идентификационных документов из открытых источников;
- пакет данных MIDV-2019, расширяющий набор условий съемки документов, представленных в MIDV-500;
- пакет данных MIDV-2020, содержащий полностью синтетические данные, и представляющий собой обширную коллекцию видеопоследовательностей, фотографий и сканов идентификационных документов, снятых в различных условиях;
- пакет данных MIDV-LAIT, содержащий синтетические данные, предназначенный для использования в задачах распознавания документов с текстовыми полями, выполненные с использованием персидско-арабской, тайской и индийской письменностью;
- пакет данных MIDV-Holo, содержащий видеопоследовательности искусственно созданных документов с пленочными дифракционными оптическими элементами защиты (голограммами);

- пакет данных DLC-2021, содержащий модели атак на предъявление документов, включая случаи предъявления цветных неламинированных копий, полутонных копий и съемки изображения документа с экрана.

В главе представлены результаты, полученные на основе опыта создания и использования пакетов данных для анализа и распознавания документов, удостоверяющих личность, сформулированы основные этапы и методологические принципы создания подобных пакетов. Принципы учитывают особенности создания и адаптации исходных шаблонов документов, разметки, и генерации значений для текстовых реквизитов и графических полей без использования персональных данных. Показано, работа над открытыми пакетами важна для того, чтобы исследователи могли объективно сравнивать и анализировать методы решения задач.

Пятая глава посвящена описанию архитектуры промышленной системы распознавания удостоверяющих документов Smart ID Engine, реализующей высокоточное и безопасное программное обеспечение для распознавания данных более 2000 типов удостоверяющих документов 210 юрисдикций мира. Компоненты системы разделены на три категории:

- компоненты, которые обрабатывают входные изображения или видеокadres, выполняя поиск всех шаблонов и определение их координат;
- компоненты, которые обрабатывают каждый отдельный шаблон документа;
- компоненты, которые комбинируют результаты распознавания шаблонов в логические представления документов, выполняют постобработку и выводят результат распознавания.

Показано, что быстродействие систем распознавания документов является ключевым параметром для эффективности массового ввода документов и работы на мобильных устройствах. Обычно такие системы используют высокопроизводительные серверы с процессорами различной архитектуры, включая архитектуры с принципом VLIW, такие как Эльбрус. Для оценки производительности системы в диссертации измерено среднее время

распознавания шести типов документов, включая паспорта, водительские удостоверения и идентификационные карты, в режимах «клиентского» и «серверного» распознавания и представлены результаты сравнения.

В заключении сформулированы основные результаты, полученные в ходе диссертационного исследования.

Замечания по содержанию диссертации.

1. Разработанные автором методы распознавания документов в современных условиях применимы не только к достаточно узкой задаче распознавания документов, удостоверяющих личность, но и других, "обычных" бизнес-документов, которые сейчас также часто сканируются при помощи мобильных устройств. В диссертации не уделено внимание потенциальному расширению применимости разработанных методов на другие типы документов.

2. В главе 3 одним из факторов использования видеопотока для распознавания является использование системы на мобильных устройствах в реальном времени. При этом кажется естественным, что использование аппаратных компонентов, таких как акселерометр и гироскоп, позволили бы как улучшить точность поиска объектов на изображениях при анализе трехмерной сцены, так и более эффективно анализировать видеопоследовательность.

3. В пакетах данных, созданных автором, присутствуют значительное количество документов различных стран, при это Российские документы представлены только паспортом РФ. Отсутствуют другие документы, которые имеют серьезную значимость в РФ: заграничный паспорт, водительские удостоверения РФ, военный билет и т. п., что ограничивает возможности применения пакетов для исследования документов, удостоверяющих личность, Российской Федерации. Кроме того, в пакетах содержатся только заполненные печатным способом документы, хотя часть паспортов РФ заполнены от руки.

4. В работе представлена оценка быстродействия промышленной системы распознавания в задаче мобильного распознавания и задаче массового ввода документов. При этом, в качестве серверного

вычислительного оборудования рассматриваются только вычислительные устройства на базе процессоров семейства Эльбрус, неясно насколько это позволяет оценить характеристики производительности в общем случае. Хотелось бы видеть оценку производительности на других российских вычислительных платформах, например платформе КОМДИВ.

Заключение.


Работа является законченным диссертационным исследованием, выполнена на высоком научном уровне, а ее результаты имеют научно-практическую значимость. Основные результаты по теме диссертации изложены в 40 печатных работах, 20 из которых изданы в журналах, рекомендованных ВАК, 30 работ индексируется Web of Science и Scopus (включая 10 работ, опубликованных в журналах Q1 и Q2). Основные результаты работы докладывались и обсуждались на профильных международных конференциях. В рамках работы по диссертации было получено 2 патента США, 5 патентов на изобретение РФ, 16 патентов на полезные модели и 4 свидетельства о государственной регистрации программы для ЭВМ. Автореферат точно отражает основное содержание диссертации.

Результаты диссертационного исследования могут быть рекомендованы к использованию в научно-исследовательских, проектных и коммерческих организациях для решения соответствующих практических задач.

Диссертационная работа Арлазарова Владимира Викторовича на тему «Мобильное распознавание и его применение к системе ввода идентификационных документов» по актуальности, научной новизне, основным положениям и выводам соответствует пунктам 9-11, 13, 14 «Положения о порядке присуждения ученых степеней», утвержденным постановлением Правительства РФ от 24.09.13 № 842, а ее автор **Арлазаров Владимир Викторович заслуживает присуждения ученой степени доктора технических наук по специальности 2.3.1 – «Системный анализ, управление и обработка информации, статистика».**

Отзыв обсужден и утвержден на совместном семинаре Центра оптико-нейронных технологий и отдела учебной информатики Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук», протокол № 2 от 28.04.2023.

Главный научный сотрудник-руководитель
Центра оптико-нейронных технологий
д.ф.-м.н., профессор, чл.-корр. РАН

 /Крыжановский Б.В.
« 28 » 04 2023 г.

Федеральное государственное учреждение
"Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук"

Адрес: 117218, Москва, Нахимовский
просп., 36, к.1.

Тел.: +7 495-718-21-10. Факс: +7 495-719-
76-81.

Электронная почта: niisi@niisi.msk.ru.

Адрес сайта: <https://niisi.ru>