

Анализ использования проблемно-ориентированных пакетов данных в научных исследованиях

В. В. Арлазаров^{1,II}

¹ Федеральное государственное учреждение «Федеральный исследовательский центр "Информатика и управление" Российской академии наук», г. Москва, Россия

^{II} ООО «Смарт Энджинс Сервис», г. Москва, Россия

Аннотация. В работе рассматриваются проблемы создания и использования открытых проблемно-ориентированных пакетов данных для проведения экспериментальных исследований с проверяемыми и воспроизводимыми результатами, на примере опыта создания пакетов семейства MIDV, содержащих изображения и видеопоследовательности идентификационных документов. Проведен анализ опубликованных научных работ в областях компьютерного зрения, обработки изображений и вычислительной лингвистики, использующих эти пакеты данных, описаны основные проблемы, с которыми сталкивались научные группы, и выявлены общие закономерности и принципы, которые могут быть использованы для создания пакетов данных такого класса и для расширения существующих.

Ключевые слова: распознавание текста, анализ документов, пакеты данных, воспроизводимость исследований, OCR, обработка изображений.

DOI 10.14357/20718632220302

Введение

Количество публикуемых исследовательских работ в таких областях, как распознавание образов, анализ изображений и распознавание документов растет с каждым годом. Одно из важных отдельных направлений в области обработки изображений документов является анализ и распознавание документов, удостоверяющих личность [1, 2]. Проблема, отмечаемая многими исследователями, работающими над методами и подходами обработки изображений идентификационных документов, является недостаток открытых пакетов данных, которые можно было бы использовать для обучения моделей и объективного анализа публикуемых алгоритмов. Для решения этой проблемы в 2018

году начала разрабатываться и публиковаться серия пакетов данных MIDV (Mobile Identity Document Video), предназначенная для объективной оценки алгоритмов распознавания документов, удостоверяющих личность, на изображениях, полученных при помощи современных устройств захвата.

В данной работе описана история создания и публикации открытых пакетов данных семейства MIDV и проанализированы исследовательские работы, использующие эти пакеты данных для обучения и тестирования. Целью работы является выработка общих принципов, согласно которым следует создавать новые и дополнять существующие открытые проблемно-ориентированные пакеты данных, предна-

значенные для объективной оценки исследований в конкретных областях знания.

1. Семейство пакетов данных MIDV

Открытый пакет данных MIDV-500 стал впервые доступен в июле 2018 года, а его подробное описание было опубликовано в 2019 году в журнале «Компьютерная оптика» [3]. Его создание послужило ответом на один из наиболее важных вопросов в области исследований методов и алгоритмов автоматического анализа документов, удостоверяющих личность — невозможность представлять и публиковать в научной литературе воспроизводимые результаты, поскольку не существует (и в общем виде не может существовать) открытых пакетов данных изображений документов, удостоверяющих личность, которые бы не нарушали требования по сохранности персональных данных. Для создания пакета данных MIDV-500 были использованы доступные примеры изображений 50-ти документов, удостоверяющих личность, выдаваемых различными государствами и межгосударственными организациями, которые распространялись в открытом доступе и репродукция которых не нарушала авторские права создателей.

Для создания пакета данных MIDV-500 исходные изображения документов предварительно регулировались (так, чтобы убрать «водяные

знаки», которых нет на действительных документах аналогичных типов и присутствие которых могло помешать алгоритмам обработки изображений документов), распечатывались с сохранением оригинального масштаба, и ламинировались, для моделирования отражающей поверхности документов. После этого для каждого из 50-ти физических примеров документов было снято 10 трехсекундных видеопоследовательностей — с использованием двух различных моделей мобильных камер (Apple iPhone 5 и Samsung Galaxy S3), в пяти различных условиях съемки: с относительно равномерным фоном (на столе), на фоне клавиатуры, в руках, на фоне разнообразных объектов, и съемка при которой на некоторых кадрах видеопоследовательности документ виден лишь частично, либо вообще не присутствует в кадре. Каждый из 500 видеороликов был разбит на отдельные кадры с частотой 10 кадров в секунду. Таким образом, общее количество кадров в пакете данных составило 15 000. Примеры кадров документов из пакета данных MIDV-500 в каждом из пяти условий представлены на Рис. 1.

Для каждого из 50-ти документов, на основе которых был подготовлен пакет данных, также были представлены координаты всех значимых текстовых полей документа и их истинные текстовые значения, а также координаты графических полей (подписи и фотографии владельца).



Рис. 1. Примеры кадров из пакета данных MIDV-500 [3]

- а) документ на фоне стола; б) документ на фоне клавиатуры; в) документ в руках; г) документ, частично скрытый за кадром; д) документ на фоне случайных объектов

Для каждого из 15 000 кадров также были предоставлены размеченные вручную координаты углов документа на изображении, таким образом, была предоставлена возможность вычислить координаты каждого значимого поля документа на каждом из кадров видеопоследовательности (для которых документ был виден в кадре).

В 2019 году пакет MIDV-500 был расширен дополнительными 200 видеопоследовательностями, которые были сняты в большем разрешении и в двух дополнительных режимах съемки (используя те же исходные физические шаблоны документов) — в режиме ограниченного освещения и с ярко выраженными проективными искажениями документа в кадре. Расширение получило название MIDV-2019 [4]. Примеры кадров видеопоследовательностей пакета MIDV-2019 представлены на Рис. 2.

Важным недостатком пакетов данных MIDV-500 и MIDV-2019 была низкая вариативность документов — хотя пакеты данных включали в себя 50 различных типов документов, для каждого типа существовал только один пример. Последующие пакеты данных MIDV-LAIT [5] и MIDV-2020 [6] были направлены, в том числе, на решение этой проблемы: пакет данных MIDV-2020 содержал 1000 видеопоследовательностей 10-ти типов документов, однако каждый из документов был создан с внесением уникальных текстовых данных, и с использованием искусственно сгенерированных изображений лица держателя документа. Пакет MIDV-LAIT содержал 180 уникальных синтезированных документов, со специализа-



Рис. 2. Примеры кадров из пакета данных MIDV-2019 [4]

- а) съемка в условиях низкого освещения;
б) съемка с ярко выраженными проективными искажениями

цией для Персо-Арабских, Тайских и Индийских языков. Примеры изображений исходных шаблонов, использованных при создании пакетов данных MIDV-LAIT и MIDV-2020, приведены на Рис. 3.

2. Использование пакетов данных семейства MIDV

Пакеты данных семейства MIDV использовались различными исследовательскими группами для проведения экспериментов в области анализа и распознавания документов, либо в качестве основных экспериментальных наборов данных, либо в комбинации с закрытыми пакетами или с другими, также публично доступными. Спектр задач, решаемых исследователя-



Рис. 3. Примеры уникальных шаблонов синтезированных документов
а) MIDV-LAIT [5]; б) MIDV-2020 [6]

ми, включал задачи детектирования и точной локализации документа на изображении, идентификацию типа изображения, распознавание текстовых реквизитов, точный поиск лица держателя документа, межкадровое комбинирование результатов распознавания, и другие.

2.1. Поиск и ректификация документов

В работах [7-9] описан алгоритм поиска четырехугольника документа на фотографии или кадре видеопоследовательности, использующий анализ контуров объектов, присутствующих на изображении и ранжирование гипотез об истинном четырехугольнике исходя из известных заранее возможных соотношениях сторон искомого документа. В этих работах эксперимент проводился с использованием MIDV-500 в качестве основного пакета данных, в котором содержатся документы с тремя различными соотношениями сторон, наиболее присущими идентификационным документам. В работах [10, 11] пакеты данных MIDV-500 используются для анализа другого типа алгоритмов детектирования документов на изображении, уже зависящих непосредственно от графического представления документа. В этих работах применяется подход с предварительным выделением и дескрибированием ключевых точек на изображении, и поиском оптимального геометрического преобразования, которое бы переводило идеальный шаблон документа в область на изображении. В работе [10] рассматриваются вопросы построения ограничений в оптимизационной задаче поиска этого графического представления, а в работе [11] рассматривается система эффективного детектирования и локализации жестко структурированных документов (на примере документов, удостоверяющих личность), пригодная как для работы на фотографиях и кадрах видеопотока, так и на сканированных изображениях. Поскольку в оригинальном наборе данных MIDV-500 не содержалось фотографий и сканированных изображений, в работе [11] также публикуется соответствующее расширение, базирующееся на тех же исходных шаблонах документов и с форматом разметки, аналогичной оригинальному. В работе [12] исследуется полный набор потенциальных признаков (включая ключевые точки, исчезающие точки,

машиночитаемую зону, границы, углы, лицо держателя, логотипы, статические тексты и т. п.), которые могут использоваться для точной локализации и идентификации документа, удостоверяющего личность, на изображении. Авторы предлагают комбинированный метод, выбирающий из множества гипотез, построенных по всевозможным признакам, и используют пакет MIDV-500 как один из двух контрольных пакетов данных.

В работе [13] предлагается метод точной локализации и идентификации документа, основанный на предварительной грубой локализации при помощи поисковой нейронной сети, и последующего уточнения на основе контурных методов. В качестве одного из контрольных пакетов данных используется MIDV-500, однако в работе отмечается, что хотя пакеты данных MIDV-500 и MIDV-2019 достаточно хорошо описывают процесс мобильной съемки документа в процедурах удаленной идентификации личности, вариативность данных, представленная в них, недостаточна для обучения нейросетевых моделей (стоит отметить, что эта работа была проделана и опубликована до публикации пакета данных MIDV-2020, содержащего множество синтезированных примеров документов). Подобный недостаток также отмечается в работах по нейросетевому детектированию текста на произвольных изображениях [14]. Тем не менее, пакет данных MIDV-500 все же использовался для обучения и тестирования нейросетевых моделей поиска точек схода с использованием слоев нейронных сетей на основе преобразования Хафа в работах [15, 16] и для обучения нейросетевых моделей для поиска документа на изображении, выражающих задачу как задачу семантической сегментации [17] — в задачах, для которых вариативность текстовых данных и лица держателя документа не является критичным фактором. Метод ректификации изображения документа, использующий предварительный поиск точек схода также рассматривается в работе [18]. В работе [19] рассматривалась задача поиска произвольного документа на изображении (не обязательно документа, удостоверяющего личность) и пакет данных MIDV-500 использовался в качестве одного из пяти открытых пакетов данных, ис-

пользующихся для обучения глубокой нейросетевой модели DeepLab (в качестве тестовой выборки использовался отдельный, также публично доступный пакет данных [20], содержащий изображения документов произвольной природы).

2.2. Поиск отдельных объектов и геометрических примитивов

Помимо задач детектирования и локализации документов также ведутся исследования в области более частных задач, связанных с устойчивым и эффективным детектированием геометрических примитивов, таких как отдельные линии или отрезки [21], генерацией устойчивых описаний локальных особенностей изображений [22, 23], и использования множественных признаков [24] для локализации и идентификации документов. Разницы в постановках задач, которые ставят авторы таких исследований, приводят к различным требованиям к пакетам данных и к предоставляемой разметке. Так, в работе [23] авторы конструируют устойчивый бинарный дескриптор локальной области изображения, для использования в методах детектирования и точного поиска документов, удостоверяющих личность. Поскольку целевой задачей является именно поиск документов, удостоверяющих личность, пакеты данных MIDV-500 и MIDV-2019 были использованы как основные пакеты как для обучения, так и для тестирования, и авторы показывают значительное преимущество построенного дескриптора перед более «универсальными» аналогами, применительно к целевой задаче. Напротив, авторы работы [21], исследующие универсальные методы устойчивого детектирования прямых линий на произвольных изображениях, отмечают, что в рамках широкой постановки задачи пакеты данных MIDV-500 не могут быть использованы, поскольку не обладают достаточно подробной геометрической разметкой, и присутствующие в разметке прямые линии относятся только к границам документов. Это приводит к тому, что авторы вынуждены проводить экспериментальные исследования на полностью синтетических пакетах данных, моделируя изображения с заранее известным положением искомого объектов.

Различия в постановках задач меняет требования не только к глубине разметки, предоставляемой вместе с открытыми пакетами данных, но и к специфике определений размечаемых объектов. Так, в работе [25] пакет данных MIDV-500 был использован в качестве основного экспериментального набора данных для решения задачи поиска лиц на изображениях документов, удостоверяющих личность. Разметка лица владельца, присутствующая в пакетах данных MIDV-500 и MIDV-2019 представляла собой четырехугольник поля документа, в который вклеена или пропечатана фотография владельца, что имеет смысл с точки зрения построения систем анализа документов, предполагающих предварительное детектирование и ректификацию страниц документа [26], однако не является типичным среди исследователей, занимающихся детектированием лиц на произвольных изображениях — более типичным является разметка окаймляющего прямоугольника овала лица человека со сторонами, параллельными сторонам изображения. В связи с этим авторы работы [25] не могли использовать консистентные метрики качества детектирования лиц и были вынуждены готовить специальную разметку необходимых окаймляющих прямоугольников на первом и последнем кадре каждой видеопоследовательности пакета данных MIDV-500 с последующей интерполяцией. Высокая заинтересованность соавторов работы [25] в появлении объемных пакетов данных для объективного сравнения методов детектирования лиц на документах также мотивировало их принять участие в работе над пакетом данных MIDV-2020, в котором для каждого из 72 409 изображений документов представлена разметка как фотографии владельца как поля документа, так и окаймляющего прямоугольника овала лица (Рис. 4).

2.3. Распознавание текстовых реквизитов на изображениях и на видеопоследовательности

Одна из наиболее важных задач анализа документов является точное распознавание его текстовых реквизитов. В работе [27] пакет данных MIDV-500 используется в качестве основного тестового набора для построения и анализа метода распознавания текстовых полей



Рис. 4. Пример разметки шаблона документа в пакете данных MIDV-2020 [6]

1 — фотография лица держателя документа, 2 — подпись, 3-13 — значимые текстовые реквизиты, 14 — окаймляющий прямоугольник овала лица

идентификационных документов на изображениях, полученных с камеры мобильных устройств (с обучением на искусственно синтезированных данных). В работе [28] предлагается имплементация квантизованных 4-битных нейронных сетей для распознавания символов, обучаемая на синтезированных данных и тестируемая на текстовых полях пакета данных MIDV-500. Вопросы генерации синтетических наборов обучающих данных также находятся в центре внимания работы [29], где пакеты данных MIDV-500, MIDV-2019 и MIDV-LAIT используются в качестве обучающей выборки для построения модели генерации правдоподобных последовательностей искусственных обучающих изображений для распознавания машиночитаемых зон идентификационных документов.

В работе [30] рассматривается задача детектирования и распознавания машиночитаемой зоны документов, удостоверяющих личность. Авторы предлагают двухэтапный нейросетевой подход MRZNet и проводят сравнительное исследование семи различных методов (включая авторский) с использованием трех пакетов данных — собственный пакет данных, не являющийся публично доступным, синтетический пакет данных SyntheticMRZ [31] и пакет MIDV-500, который содержит изображения междуна-

родных паспортов и идентификационных карт, на которых присутствуют машиночитаемые зоны. Примечательно, что точность детектирования машиночитаемых зон на закрытом пакете данных, отмечаемая авторами, для собственного метода составила 100% (при точности ближайшего сравниваемого метода 74.15%), тогда как на пакете данных MIDV-500 точность авторского метода составила только 73.94%, и показала лишь третье место среди сравниваемых методов. Авторы объясняют такое несоответствие присутствием проективных и нелинейных искажений и бликов на изображениях пакета MIDV-500.

В работе [32] исследуется влияние предварительной бинаризации изображений документов на качество распознавания текстовых реквизитов, с использованием пакета данных MIDV-500. Авторами показано, что современные широко используемые методы распознавания текстовых строк показывают более высокую точность распознавания при обработке исходных изображений, нежели чем при обработке предварительно бинаризованных изображений, даже при использовании лидирующих методов бинаризации, однако при использовании идеальной бинаризации (при которой для каждого конкретного изображения порог бина-

ризации в каждом локальном окне подбирается человеком) точность методов распознавания все же возрастает. Для проведения такого исследования авторам пришлось подготовить специальную трудоемкую разметку идеальной бинаризации документов для исходных изображений-шаблонов, кадры видеопоследовательностей пакета данных MIDV-500 в работе не рассматривались.

Поскольку целевой задачей при подготовке пакетов данных семейства MIDV являлось распознавание документов, удостоверяющих личность, в видеопоследовательности, эти пакеты данных стали основными для ряда исследований по повышению точности распознавания текстовых реквизитов путем межкадрового комбинирования результатов распознавания [33, 34], анализа влияния на качество распознавания текстовых полей искажений отдельных кадров, таких, как дефокусировка [35] или поворот в трехмерном пространстве [36] и построения решающего правила, позволяющего автоматически останавливать процесс распознавания объектов в видеопотоке [37-40].

2.4. Анализ качества изображения и поиск компрометирующих признаков

В автоматических системах анализа и распознавания документов, удостоверяющих личность, помимо непосредственно распознавания текстовых полей и извлечения значимой графической информации (такой, как фотография лица держателя документа), важной задачей является детальный анализ качества изображения документа и выявление признаков, свидетельствующих о том, что изображение или сам документ могли быть подделаны.

В работе [41] рассматривается задача определения качества фотографии лица владельца, извлекаемой из изображения документа, и ее пригодности для дальнейшей обработки (к примеру, для автоматической сверки с фотографией лица, полученной из другого источника). Качество фотографии в рамках этого исследования определяется экспертом, что требует дополнительной разметки изображений пакета данных. В работе [42] рассматривается задача определения качества целикового изображения документа. В этой работе используется пакет данных MIDV-

2020 — исходные изображения шаблонов документов рассматриваются как «эталонные» изображения с максимальным качеством, что позволяет анализировать артефакты их сканирования и фотографирования.

В работах [43-45] рассматривается задача автоматического детектирования чувствительной информации (к примеру, персональных данных) в потоках изображений — задача информационной безопасности, требующая, в том числе, применение современных методов компьютерного зрения. В этих работах пакет данных MIDV-500 использовался в качестве положительной выборки — набора изображений, на которых заведомо присутствуют чувствительные данные, которые необходимо автоматически детектировать. Безусловное наличие на изображениях пакета данных MIDV-500 информации, структурированной аналогично персональным данным, также использовалось в работе [46] для построения интеллектуальных моделей лингвистического анализа платежной информации.

Пакет данных MIDV-500 также использовался в качестве позитивной выборки в работе [47] для построения автоматического метода детектирования нарушений в текстуре фона документов, удостоверяющих личность, для выявления подозрительных манипуляций над изображениями. Отрицательная выборка (т. е. выборка изображений, над которыми проводились манипуляции) генерировалась автоматически, используя изображения исходных шаблонов документов из MIDV-500. Напротив, в работе [48] пакет MIDV-500 используется в качестве отрицательной выборки — в данной работе авторы предлагают метод автоматической верификации присутствия голографических элементов защиты на идентификационном документе. В качестве положительной выборки авторы использовали закрытый пакет данных действительных документов, а в качестве «подделок» — пакет MIDV-500, в котором не содержится никаких элементов защиты бланка.

Исследование методов автоматического детектирования компрометирующих признаков является крайне важным направлением для повышения безопасности процесса удаленной идентификации личности, и для проведения та-

ких исследований необходимо наращивать базу пакетов данных самой различной природы. Как отмечают авторы работы [49], пакетов данных MIDV-500, MIDV-2019 и MIDV-2020 недостаточно для комплексного анализа систем верификации действительности идентификационных документов, поскольку в них не представлены примеры распространенных атак, таких как повторная съемка документа, отображаемого на экране. Одна из первых попыток представить вариант пакета данных MIDV-2020 с моделированием различных атак проделана в работе [50], в которой представлено расширение, включающее в себя неламинированные документы, повторную съемку с экрана, и съемку копий документов в градациях серого (Рис. 5).

3. Общие наблюдения и выводы

Для поддержания современных темпов развития методов и подходов к задачам компьютерного зрения, в особенности учитывая тренд использования глубоких нейросетевых моделей, требует создания больших объемов данных для обучения и тестирования. Применительно к конкретным задачам, таким как анализ и распознавание документов, удостоверяющих личность, это создает трудности — либо связанные с особенностями целевого объекта (такие, как соблюдение юридических и этических норм,

связанных с сохранностью чувствительных данных), либо связанные с ограниченным размером сообщества, занимающимся тем или иным рядом задач. Однако как показывает опыт использования пакетов данных семейства MIDV, даже в случае, если основной экспериментальный пакет данных не публикуется, ограниченный открытый пакет данных, такой как MIDV-500, может использоваться в качестве дополнительного набора данных для публикации объективной оценки качества того или иного метода, пригодного для сравнения с другими.

Авторы ряда открытых пакетов данных заранее разделяют публикуемые выборки на обучающие, валидационные и тестовые, однако, как можно заметить из практики использования пакетов семейства MIDV, такое разделение может не иметь смысла — разнообразие подходов к решению тех или иных задач настолько велико, что исследователи вынуждены самостоятельно выбирать протоколы разделения пакетов данных для обучения и проверки, а иногда, для решения более высокоуровневых задач, используют комбинацию нескольких пакетов данных из разных доменов (как, к примеру, в работе [19]).

Гораздо более важным аспектом построения открытых пакетов данных является структурирование пакетов таким образом, чтобы исследователи могли выделять строго определенным образом подмножества, объединяемые теми

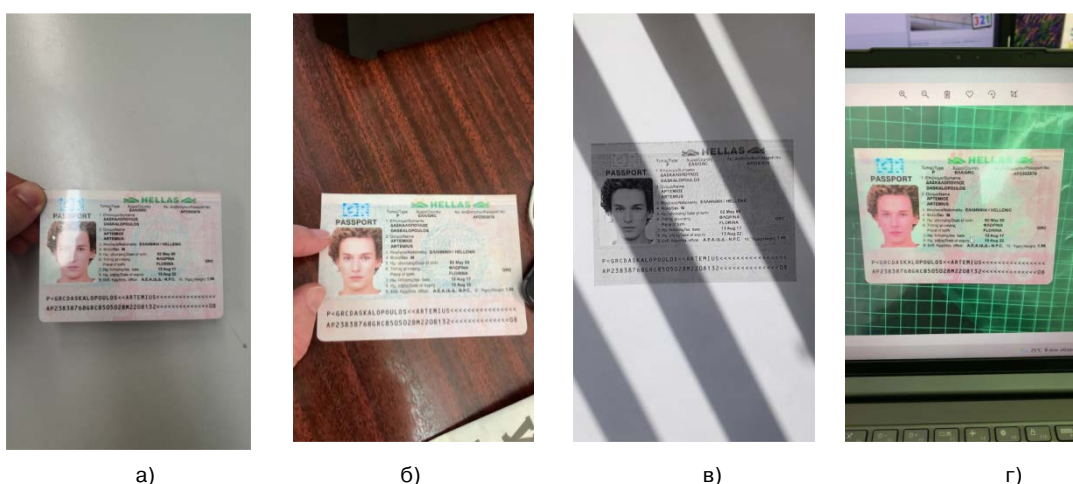


Рис. 5. Примеры изображений пакета данных DLC-2021 [50]

а) оригинальный документ MIDV-2020; б) неламинированная цветная копия;
в) неламинированная полутоновая копия; г) повторная съемка документа с экрана

или иными признаками. К примеру, авторы проанализированных работ отдельно анализировали подмножества пакетов данных по освещенности и степени выраженности проективных искажений [33], по количеству видимых в кадре углов документа [7, 9], по наличию и целостности определенных видов текстовых полей или машиночитаемых зон [30, 37] и т. п. Поскольку одной из целей публикации открытых пакетов данных является предоставление возможности различным исследовательским группам сравнивать разрабатываемые методы и подходы на общей экспериментальной базе, необходимо предоставить возможность в явном виде указывать те подмножества пакета данных, которые используются в том или ином эксперименте.

Наиболее важным вопросом в построении и развитии открытых пакетов данных является вопрос их расширения либо новыми данными (с целью увеличения общего объема, либо с целью добавления данных, обладающих новыми, не представленными ранее признаками), либо добавлением специфической, проблемно-ориентированной разметки (к примеру, окаймляющих прямоугольников овалов лиц [25], идеальной бинаризации изображений документов [32], экспертной разметкой качества изображений или их фрагментов [41] и др.), либо производных сгенерированных изображений (таких, как изображения с нарушениями текстур фона документа [47]). Зачастую такие расширения используются авторами работ для конкретных исследований, но не выкладываются в публичное пространство для повторного воспроизведения результатов или для сравнительного анализа. В связи с этим важным аспектом открытых пакетов данных является предоставление инструментария для пользовательского расширения этих пакетов данных новыми изображениями, либо новой разметкой так, чтобы другие исследовательские группы могли выбирать не только интересующие их подмножества опубликованных данных, но и специфические производные изображения или уже подготовленную разметку, тем самым уменьшая несоответствия при сравнении методов и упрощая методологию исследований.

Заключение

В работе были проанализированы исследовательские работы, направленные на создание и анализ методов и подходов решения задач автоматической обработки документов, удостоверяющих личность, с использованием открытых пакетов данных семейства MIDV. Анализ работ показывает общие тренды использования проблемно-ориентированных пакетов данных, включая различающиеся среди коллективов требования к признакам, которыми должны обладать изображения и идеальная разметка, и были выявлены принципы, использование которых при дальнейшем создании и расширении открытых пакетов данных позволит использовать их более эффективным образом и повышать качество публикуемых исследований. Работа над открытыми пакетами данных важна не только для того, чтобы давать возможность исследователям проводить объективное сравнение и анализ методов решения тех или иных задач, но и для создания и развития новых научных связей между международными исследовательскими группами, объединенных интересом к общему кругу научных проблем.

Литература

1. V. L. Arlazarov, V. V. Arlazarov, K. B. Bulatov, T. S. Chernov, D. P. Nikolaev, D. V. Polevoy, A. V. Sheshkus, N. S. Skoryukina, O. A. Slavin and S. A. Usilin, "Mobile ID Document Recognition-Coarse-to-Fine Approach," *Pattern Recognit. Image Anal.*, vol. 32, no 1, pp. 89-108, 2022, DOI: 10.1134/S1054661822010023.
2. A. Chandra and R. Stefanus, "An End-to-End Optical Character Recognition Pipeline for Indonesian Identity Card," 2021 9th International Conference on Information and Communication Technology (ICoICT), 2021, pp. 307-312, DOI: 10.1109/IcoICT52021.2021.9527436.
3. V. V. Arlazarov, K. Bulatov, T. Chernov and V. L. Arlazarov, "MIDV-500: A Dataset for Identity Document Analysis and Recognition on Mobile Devices in Video Stream," *Computer Optics*, vol. 43, no 5, pp. 818-824, 2019, DOI: 10.18287/2412-6179-2019-43-5-818-824.
4. K. Bulatov, D. Matalov and V. V. Arlazarov, "MIDV-2019: Challenges of the Modern Mobile-Based Document OCR," *ICMV 2019*, vol. 11433, pp. 114332N1-114332N6, 2020, DOI: 10.1117/12.2558438.
5. Y. S. Chernyshova, E. V. Emelianova, A. V. Sheshkus and V. V. Arlazarov, "MIDV-LAIT: a challenging dataset for recognition of IDs with Perso-Arabic, Thai, and Indian scripts," *ICDAR 2021*, vol. 12822, pp. 258-272, 2021, DOI: 10.1007/978-3-030-86331-9_17.

6. K. B. Bulatov, E. V. Emelyanova, D. V. Tropin, N. S. Skoryukina, Y. S. Chernyshova, A. V. Sheshkus, S. A. Usilin, Z. Ming, J. Burie, M. Luqman and V. V. Arlazarov, "MIDV-2020: A Comprehensive Benchmark Dataset for Identity Document Analysis," *Computer Optics*, vol. 46, no 2, pp. 252-270, 2022, DOI: 10.18287/2412-6179-CO-1006.
7. D. V. Tropin, A. M. Ershov, D. P. Nikolaev and V. V. Arlazarov, "Advanced Hough-based method for on-device document localization," *Computer Optics*, vol. 45, no 5, pp. 702-712, 2021, DOI: 10.18287/2412-6179-CO-895.
8. D. V. Tropin, S. A. Ilyukhin, D. P. Nikolaev and V. V. Arlazarov, "Approach for document detection by contours and contrasts," *ICPR 2020*, pp. 9689-9695, 2021, DOI: 10.1109/ICPR48806.2021.9413271.
9. D. V. Tropin, I. A. Konovalenko, N. S. Skoryukina, D. P. Nikolaev and V. V. Arlazarov, "Improved algorithm of ID card detection by a priori knowledge of the document aspect ratio," *ICMV 2020*, vol. 11605, 116051F, pp. 116051F1-116051F9, 2021, DOI: 10.1117/12.2587029.
10. N. Skoryukina, I. Farajev, K. Bulatov and V. V. Arlazarov, "Impact of geometrical restrictions in RANSAC sampling on the ID document classification," *ICMV 2019*, vol. 11433, pp. 1143306-1-1143306-7, 2020, DOI: 10.1117/12.2559306.
11. N. Skoryukina, V. V. Arlazarov and D. P. Nikolaev, "Fast method of ID documents location and type identification for mobile and server application," *ICDAR 2019*, pp. 850-857, 2020, DOI: 10.1109/ICDAR.2019.00141.
12. G. Chiron, N. Ghanmi and A. M. Awal, "ID documents matching and localization with multi-hypothesis constraints," 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 3644-3651, DOI: 10.1109/ICPR48806.2021.9412437.
13. Chiron, G., Arrestier, F., Awal, A.M. Fast End-to-End Deep Learning Identity Document Detection, Classification and Cropping. In: Lladós, J., Lopresti, D., Uchida, S. (eds) *Document Analysis and Recognition – ICDAR 2021. Lecture Notes in Computer Science*, vol 12824. Springer, Cham, 2021. DOI: 10.1007/978-3-030-86337-1_23
14. O. Naparstek, O. Azulai, D. Rotman, Y. Burshtein, P. Staar, U. Barzelay. BusiNet — a Light and Fast Text Detection Network for Business Documents. *ArXiv preprints* (2207.01220), URL: <https://arxiv.org/abs/2207.01220>
15. A. V. Sheshkus and D. Nikolaev, "Transfer of a high-level knowledge in HoughNet neural network," *ICMV 2019*, vol. 11433, ISSN 0277-786X, ISBN 978-15-10636-44-6, vol. 11433, pp. 1143322-1-1143322-6, 2020, DOI: 10.1117/12.2559454.
16. A. Sheshkus, A. Ingacheva, V. Arlazarov and D. Nikolaev, "HoughNet: neural network architecture for vanishing points detection," *ICDAR 2019*, pp. 844-849, 2020, DOI: 10.1109/ICDAR.2019.00140.
17. A. V. Sheshkus, D. P. Nikolaev and V. L. Arlazarov, "Houghencoder: neural network architecture for document image semantic segmentation," *IEEE ICIP 2020*, pp. 1946-1950, 2020, DOI: 10.1109/ICIP40778.2020.9191182.
18. J. Shemiakina, I. Konovalenko, D. Tropin and I. Faradjev, "Fast projective image rectification for planar objects with Manhattan structure," *ICMV 2019*, vol. 11433, pp. 114331N1-114331N9, 2020, DOI: 10.1117/12.2559630.
19. Baniadamdzaj, S. Localization Using DeepLab in Document Images Taken by Smartphones. In: *Digital Interaction and Machine Intelligence. MIDI 2021. Lecture Notes in Networks and Systems*, vol 440. Springer, Cham, 2022. DOI: 10.1007/978-3-031-11432-8_6.
20. S. B. Dizaj, M. Soheili and A. Mansouri. A New Image Dataset for Document Corner Localization. *International Conference on Machine Vision and Image Processing (MVIP)*, 2020, pp. 1-4, DOI: 10.1109/MVIP49855.2020.9116896.
21. L. Teplyakov, K. Kaymakov, E. Shvets and D. Nikolaev, "Line detection via a lightweight CNN with a Hough Layer," *ICMV 2020*, vol. 11605, pp. 116051B1-116051B10, 2021, DOI: 10.1117/12.2587167.
22. A. Sheshkus, A. Chirvonaya and V. L. Arlazarov, "Tiny CNN for feature point description for document analysis: approach and dataset," *Computer Optics*, vol. 46, no 3, pp. 429-435, 2022, DOI: 10.18287/2412-6179-CO-1016.
23. D. P. Matalov, E. E. Limonova, N. S. Skoryukina and V. V. Arlazarov, "RFDoc: memory efficient local descriptors for ID documents localization and classification," *ICDAR 2021*, vol. 12822, pp. 209-224, 2021, DOI: 10.1007/978-3-030-86331-9_14.
24. Buonanno, A.; Nogarotto, A.; Cacace, G.; Di Gennaro, G.; Palmieri, F.A.N.; Valenti, M.; Graditi, G. Bayesian Feature Fusion Using Factor Graph in Reduced Normal Form. *Appl. Sci.*, vol. 11, p. 1934, 2021 DOI: 10.3390/app11041934.
25. S. Bakkali, M. M. Luqman, Z. Ming and J. Burie, "Face Detection in Camera Captured Images of Identity Documents Under Challenging Conditions," 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), 2019, pp. 55-60, DOI: 10.1109/ICDARW.2019.30065.
26. K. B. Bulatov, P. V. Bezmaternykh, D. P. Nikolaev and V. V. Arlazarov, "Towards a unified framework for identity documents analysis and recognition," *Computer Optics*, vol. 46, no 3, pp. 436-454, 2022, DOI: 10.18287/2412-6179-CO-1024.
27. Y. S. Chernyshova, A. V. Sheshkus and V. V. Arlazarov, "Two-step CNN framework for text line recognition in camera-captured images," *IEEE Access*, vol. 8, pp. 32587-32600, 2020, DOI: 10.1109/ACCESS.2020.2974051.
28. A. V. Trusov, E. E. Limonova, D. G. Slugin, D. P. Nikolaev and V. V. Arlazarov, "Fast Implementation of 4-bit Convolutional Neural Networks for Mobile Devices," *ICPR 2020*, pp. 9897-9903, 2021, DOI: 10.1109/ICPR48806.2021.9412841.
29. P. K. Zlobin, Y. S. Chernyshova, A. V. Sheshkus and V. V. Arlazarov, "Character Sequence Prediction Method for Training Data Creation in the Task of Text Recognition," *ICMV 2021*, vol. 12084, pp. 120840R1-120840R8, 2022, DOI: 10.1117/12.2623773.
30. Liu, Y., James, H., Gupta, O. et al. MRZ code extraction from visa and passport documents using convolutional neural networks. *IJDAR* vol. 25, pp. 29–39, 2022. DOI: 10.1007/s10032-021-00384-2.

31. Hartl, A., Arth, C., Schmalstieg, D. Real-time Detection and Recognition of Machine-Readable Zones with Mobile Devices. *International Conference on Computer Vision Theory and Applications*, pp. 79–87, 2015.
32. R. Sánchez-Rivero, P. V. Bezmaternykh, A. Morales-González and K. B. Bulatov, "Assessing the relationship between binarization and OCR in the context of deep learning-based ID document analysis," *IWAIPR 2021*, vol. 13055, pp. 134-144, 2021, DOI: 10.1007/978-3-030-89691-1_14.
33. O. O. Petrova, K. B. Bulatov, V. V. Arlazarov and V. L. Arlazarov, "Weighted combination of per-frame recognition results for text recognition in a video stream," *Computer Optics*, vol. 45, no 1, pp. 77-89, 2021, DOI: 10.18287/2412-6179-CO-795.
34. K. B. Bulatov, "A Method to Reduce Errors of String Recognition Based on Combination of Several Recognition Results with Per-Character Alternatives," *Bulletin of the South Ural State University, Series: Mathematical Modelling, Programming and Computer Software*, vol. 12, no 3, pp. 74-88, 2019, DOI: 10.14529/mmp190307.
35. T. S. Chernov, S. A. Ilyuhin and V. V. Arlazarov, "Application of dynamic saliency maps to video stream recognition systems with image quality assessment," *ICMV 2018*, vol. 11041, pp. 110410T1-110410T8, 2019, DOI: 10.1117/12.2522768.
36. Y. A. Shemyakina, E. E. Limonova, N. S. Skoryukina, V. V. Arlazarov and D. P. Nikolaev, "A method of image quality assessment for text recognition on camera-captured and projectively distorted documents," *Mathematics*, vol. 9, no 17, pp. 1-22, 2021, DOI: 10.3390/math9172155.
37. K. Bulatov, N. Razumnyi and V. V. Arlazarov, "On optimal stopping strategies for text recognition in a video stream as an application of a monotone sequential decision model," *IJDAR*, vol. 22, no 3, pp. 303-314, 2019, DOI: 10.1007/s10032-019-00333-0.
38. K. Bulatov, B. Savelyev and V. V. Arlazarov, "Next integrated result modelling for stopping the text field recognition process in a video using a result model with per-character alternatives," *ICMV 2019*, vol. 11433, pp. 114332M1-114332M6, 2020, DOI: 10.1117/12.2559447.
39. K. B. Bulatov and V. V. Arlazarov, "Determining optimal frame processing strategies for real-time document recognition systems," *ICDAR 2021*, vol. 12822, pp. 273-288, 2021, DOI: 10.1007/978-3-030-86331-9_18.
40. K. B. Bulatov, N. V. Fedotova and V. V. Arlazarov, "Fast Approximate Modelling of the Next Combination Result for Stopping the Text Field Recognition in a Video Stream," *ICPR 2020*, pp. 239-246, 2021, DOI: 10.1109/ICPR48806.2021.9412574.
41. D. V. Polevoy, M. A. Aliev and D. P. Nikolaev, "Choosing the best image of the document owner's photograph in the video stream on the mobile device," *ICMV 2020*, vol. 11605, 116050F, pp. 116050F1-116050F9, 2021, DOI: 10.1117/12.2586939.
42. M. Al-Ghadi, P. Gomez-Kramer, J.-C. Burie. CheckScan: a reference hashing for identity document quality detection. *ICMV 2021*, 120840J, DOI: 10.1117/12.2623887.
43. E. Myasnikov and A. Savchenko, "Detection of Sensitive Textual Information in User Photo Albums on Mobile Devices," 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), 2019, pp. 0384-0390, DOI: 10.1109/SIBIRCON48586.2019.8958325.
44. L. Kopeykina and A. V. Savchenko, "Automatic Privacy Detection in Scanned Document Images Based on Deep Neural Networks," 2019 International Russian Automation Conference (RusAutoCon), 2019, pp. 1-6, DOI: 10.1109/RUSAUTOCON.2019.8867614.
45. H. Ahmed, I. Traore, S. Saad, M. Mamun. Automated detection of unstructured context-dependent sensitive information using deep learning. *Internet of Things*, vol. 16, 100444, 2021, DOI: 10.1016/j.iot.2021.100444.
46. A. Startseva, A. Vulfin, V. Vasilyev, A. Nikonov and A. Kirillova, "Analysis of Financial Payments Text Labels in the Dynamic Client Profile Construction," 2020 International Conference on Information Technology and Nanotechnology (ITNT), 2020, pp. 1-10, DOI: 10.1109/ITNT49337.2020.9253280.
47. M. Al-Ghadi, Z. Ming, P. Gomez-Kramer, J.-C. Burie. Identity documents authentication based on forgery detection of guilloche pattern. *ArXiv preprints*, 2206.10989, URL: <https://arxiv.org/abs/2206.10989v1>.
48. Kada, O., Kurtz, C., van Kieu, C., Vincent, N. Hologram Detection for Identity Document Authentication. In: El Yacoubi, M., Granger, E., Yuen, P.C., Pal, U., Vincent, N. (eds) *Pattern Recognition and Artificial Intelligence. ICPRAI 2022. Lecture Notes in Computer Science*, vol 13363. Springer, Cham, 2022. DOI: 10.1007/978-3-031-09037-0_29.
49. C. Chen, L. Zhao, J. Yan, H. Li. A distortion model-based pre-screening method for document image tampering localization under recapturing attack. *Signal Processing*, vol. 200, 108666, 2022. DOI: 10.1016/j.sigpro.2022.108666.
50. D. V. Polevoy, I. V. Sigareva, D. M. Ershova, V. V. Arlazarov, D. P. Nikolaev, M. Zuheng, M. L. Muhammad and J. Burie, "Document Liveness Challenge dataset (DLC-2021)," *J. Imaging*, vol. 8, no 7, pp. 181-1-181-12, 2022, DOI: 10.3390/jimaging8070181.

Арлазаров Владимир Викторович. Федеральное государственное учреждение «Федеральный исследовательский центр "Информатика и управление" Российской академии наук», г. Москва, Россия. Заведующий отделом, кандидат технических наук. Количество печатных работ: 130. Область научных интересов: искусственный интеллект, машинное обучение, системы распознавания, информационные технологии. E-mail: vva777@gmail.com

Analysis of the Usage of Problem-Oriented Datasets in Scientific Research

V. V. Arlazarov^{1,II}

¹Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

^{II}Smart Engines Service LLC, Moscow, Russia

Abstract. In this paper we consider the problem of creating and using open problem-oriented datasets to facilitate verifiable and reproducible research, based on the study of the usage of MIDV family of datasets, which contain images and video sequences of identity documents. An analysis is presented of published scientific works in the fields of computer vision, image processing, and computational linguistics, which use these datasets. Main problems are described which were tackled by the research groups, and general principles are formulated, which could be used for creating and expanding the datasets of this class.

Keywords: text recognition, document analysis, datasets, reproducible research, OCR, image processing.

DOI 10.14357/20718632220302

References

1. V. L. Arlazarov, V. V. Arlazarov, K. B. Bulatov, T. S. Chernov, D. P. Nikolaev, D. V. Polevoy, A. V. Sheshkus, N. S. Skoryukina, O. A. Slavin and S. A. Usilin, "Mobile ID Document Recognition-Coarse-to-Fine Approach," *Pattern Recognit. Image Anal.*, vol. 32, no 1, pp. 89-108, 2022, DOI: 10.1134/S1054661822010023.
2. Chandra and R. Stefanus, "An End-to-End Optical Character Recognition Pipeline for Indonesian Identity Card," 2021 9th International Conference on Information and Communication Technology (ICoICT), 2021, pp. 307-312, DOI: 10.1109/IcoICT52021.2021.9527436.
3. V. V. Arlazarov, K. Bulatov, T. Chernov and V. L. Arlazarov, "MIDV-500: A Dataset for Identity Document Analysis and Recognition on Mobile Devices in Video Stream," *Computer Optics*, vol. 43, no 5, pp. 818-824, 2019, DOI: 10.18287/2412-6179-2019-43-5-818-824.
4. K. Bulatov, D. Matalov and V. V. Arlazarov, "MIDV-2019: Challenges of the Modern Mobile-Based Document OCR," *ICMV 2019*, vol. 11433, pp. 114332N1-114332N6, 2020, DOI: 10.1117/12.2558438.
5. Y. S. Chernyshova, E. V. Emelianova, A. V. Sheshkus and V. V. Arlazarov, "MIDV-LAIT: a challenging dataset for recognition of IDs with Perso-Arabic, Thai, and Indian scripts," *ICDAR 2021*, vol. 12822, pp. 258-272, 2021, DOI: 10.1007/978-3-030-86331-9_17.
6. K. B. Bulatov, E. V. Emelyanova, D. V. Tropin, N. S. Skoryukina, Y. S. Chernyshova, A. V. Sheshkus, S. A. Usilin, Z. Ming, J. Burie, M. Luqman and V. V. Arlazarov, "MIDV-2020: A Comprehensive Benchmark Dataset for Identity Document Analysis," *Computer Optics*, vol. 46, no 2, pp. 252-270, 2022, DOI: 10.18287/2412-6179-CO-1006.
7. D. V. Tropin, A. M. Ershov, D. P. Nikolaev and V. V. Arlazarov, "Advanced Hough-based method for on-device document localization," *Computer Optics*, vol. 45, no 5, pp. 702-712, 2021, DOI: 10.18287/2412-6179-CO-895.
8. D. V. Tropin, S. A. Ilyukhin, D. P. Nikolaev and V. V. Arlazarov, "Approach for document detection by contours and contrasts," *ICPR 2020*, pp. 9689-9695, 2021, DOI: 10.1109/ICPR48806.2021.9413271.
9. D. V. Tropin, I. A. Konovalenko, N. S. Skoryukina, D. P. Nikolaev and V. V. Arlazarov, "Improved algorithm of ID card detection by a priori knowledge of the document aspect ratio," *ICMV 2020*, vol. 11605, 116051F, pp. 116051F1-116051F9, 2021, DOI: 10.1117/12.2587029.
10. N. Skoryukina, I. Farajev, K. Bulatov and V. V. Arlazarov, "Impact of geometrical restrictions in RANSAC sampling on the ID document classification," *ICMV 2019*, vol. 11433, pp. 1143306-1-1143306-7, 2020, DOI: 10.1117/12.2559306.
11. N. Skoryukina, V. V. Arlazarov and D. P. Nikolaev, "Fast method of ID documents location and type identification for mobile and server application," *ICDAR 2019*, pp. 850-857, 2020, DOI: 10.1109/ICDAR.2019.00141.
12. G. Chiron, N. Ghanmi and A. M. Awal, "ID documents matching and localization with multi-hypothesis constraints," 2020 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 3644-3651, DOI: 10.1109/ICPR48806.2021.9412437.
13. Chiron, G., Arrestier, F., Awal, A.M. Fast End-to-End Deep Learning Identity Document Detection, Classification and Cropping. In: Lladós, J., Lopresti, D., Uchida, S. (eds) Document Analysis and Recognition – ICDAR 2021. Lecture Notes in Computer Science, vol 12824. Springer, Cham, 2021. DOI: 10.1007/978-3-030-86337-1_23
14. Naparstek, O. Azulai, D. Rotman, Y. Burshtein, P. Staar, U. Barzelay. BusiNet — a Light and Fast Text Detection

- Network for Business Documents. ArXiv preprints (2207.01220), URL: <https://arxiv.org/abs/2207.01220>
15. V. Sheshkus and D. Nikolaev, "Transfer of a high-level knowledge in HoughNet neural network," *ICMV 2019*, vol. 11433, ISSN 0277-786X, ISBN 978-15-10636-44-6, vol. 11433, pp. 1143322-1-1143322-6, 2020, DOI: 10.1117/12.2559454.
 16. Sheshkus, A. Ingacheva, V. Arlazarov and D. Nikolaev, "HoughNet: neural network architecture for vanishing points detection," *ICDAR 2019*, pp. 844-849, 2020, DOI: 10.1109/ICDAR.2019.00140.
 17. V. Sheshkus, D. P. Nikolaev and V. L. Arlazarov, "Houghencoder: neural network architecture for document image semantic segmentation," *IEEE ICIP 2020*, pp. 1946-1950, 2020, DOI: 10.1109/ICIP40778.2020.9191182.
 18. J. Shemiakina, I. Konovalenko, D. Tropin and I. Faradjev, "Fast projective image rectification for planar objects with Manhattan structure," *ICMV 2019*, vol. 11433, pp. 114331N1-114331N9, 2020, DOI: 10.1117/12.2559630.
 19. Baniadamdizaj, S. Localization Using DeepLab in Document Images Taken by Smartphones. In: *Digital Interaction and Machine Intelligence. MIDI 2021. Lecture Notes in Networks and Systems*, vol 440. Springer, Cham, 2022. DOI: 10.1007/978-3-031-11432-8_6.
 20. B. Dizaj, M. Soheili and A. Mansouri. A New Image Dataset for Document Corner Localization. *International Conference on Machine Vision and Image Processing (MVIP)*, 2020, pp. 1-4, DOI: 10.1109/MVIP49855.2020.9116896.
 21. L. Teplyakov, K. Kaymakov, E. Shvets and D. Nikolaev, "Line detection via a lightweight CNN with a Hough Layer," *ICMV 2020*, vol. 11605, pp. 116051B1-116051B10, 2021, DOI: 10.1117/12.2587167.
 22. Sheshkus, A. Chirvonaya and V. L. Arlazarov, "Tiny CNN for feature point description for document analysis: approach and dataset," *Computer Optics*, vol. 46, no 3, pp. 429-435, 2022, DOI: 10.18287/2412-6179-CO-1016.
 23. D. P. Matalov, E. E. Limonova, N. S. Skoryukina and V. V. Arlazarov, "RFDoc: memory efficient local descriptors for ID documents localization and classification," *ICDAR 2021*, vol. 12822, pp. 209-224, 2021, DOI: 10.1007/978-3-030-86331-9_14.
 24. Buonanno, A.; Nogarotto, A.; Cacace, G.; Di Gennaro, G.; Palmieri, F.A.N.; Valenti, M.; Graditi, G. Bayesian Feature Fusion Using Factor Graph in Reduced Normal Form. *Appl. Sci.*, vol. 11, p. 1934, 2021 DOI: 10.3390/app11041934.
 25. S. Bakkali, M. M. Luqman, Z. Ming and J. Burie, "Face Detection in Camera Captured Images of Identity Documents Under Challenging Conditions," 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), 2019, pp. 55-60, DOI: 10.1109/ICDARW.2019.30065.
 26. K. B. Bulatov, P. V. Bezmaternykh, D. P. Nikolaev and V. V. Arlazarov, "Towards a unified framework for identity documents analysis and recognition," *Computer Optics*, vol. 46, no 3, pp. 436-454, 2022, DOI: 10.18287/2412-6179-CO-1024.
 27. Y. S. Chernyshova, A. V. Sheshkus and V. V. Arlazarov, "Two-step CNN framework for text line recognition in camera-captured images," *IEEE Access*, vol. 8, pp. 32587-32600, 2020, DOI: 10.1109/ACCESS.2020.2974051.
 28. V. Trusov, E. E. Limonova, D. G. Slugin, D. P. Nikolaev and V. V. Arlazarov, "Fast Implementation of 4-bit Convolutional Neural Networks for Mobile Devices," *ICPR 2020*, pp. 9897-9903, 2021, DOI: 10.1109/ICPR48806.2021.9412841.
 29. P. K. Zlobin, Y. S. Chernyshova, A. V. Sheshkus and V. V. Arlazarov, "Character Sequence Prediction Method for Training Data Creation in the Task of Text Recognition," *ICMV 2021*, vol. 12084, pp. 120840R1-120840R8, 2022, DOI: 10.1117/12.2623773.
 30. Liu, Y., James, H., Gupta, O. et al. MRZ code extraction from visa and passport documents using convolutional neural networks. *IJDAR* vol. 25, pp. 29-39, 2022. DOI: 10.1007/s10032-021-00384-2.
 31. Hartl, A., Arth, C., Schmalstieg, D. Real-time Detection and Recognition of Machine-Readable Zones with Mobile Devices. *International Conference on Computer Vision Theory and Applications*, pp. 79-87, 2015.
 32. R. Sánchez-Rivero, P. V. Bezmaternykh, A. Morales-González and K. B. Bulatov, "Assessing the relationship between binarization and OCR in the context of deep learning-based ID document analysis," *IWAIPR 2021*, vol. 13055, pp. 134-144, 2021, DOI: 10.1007/978-3-030-89691-1_14.
 33. O. O. Petrova, K. B. Bulatov, V. V. Arlazarov and V. L. Arlazarov, "Weighted combination of per-frame recognition results for text recognition in a video stream," *Computer Optics*, vol. 45, no 1, pp. 77-89, 2021, DOI: 10.18287/2412-6179-CO-795.
 34. K. B. Bulatov, "A Method to Reduce Errors of String Recognition Based on Combination of Several Recognition Results with Per-Character Alternatives," *Bulletin of the South Ural State University, Series: Mathematical Modelling, Programming and Computer Software*, vol. 12, no 3, pp. 74-88, 2019, DOI: 10.14529/mmp190307.
 35. T. S. Chernov, S. A. Ilyuhin and V. V. Arlazarov, "Application of dynamic saliency maps to video stream recognition systems with image quality assessment," *ICMV 2018*, vol. 11041, pp. 110410T1-110410T8, 2019, DOI: 10.1117/12.2522768.
 36. Y. A. Shemyakina, E. E. Limonova, N. S. Skoryukina, V. V. Arlazarov and D. P. Nikolaev, "A method of image quality assessment for text recognition on camera-captured and projectively distorted documents," *Mathematics*, vol. 9, no 17, pp. 1-22, 2021, DOI: 10.3390/math9172155.
 37. K. Bulatov, N. Razumnyi and V. V. Arlazarov, "On optimal stopping strategies for text recognition in a video stream as an application of a monotone sequential decision model," *IJDAR*, vol. 22, no 3, pp. 303-314, 2019, DOI: 10.1007/s10032-019-00333-0.
 38. K. Bulatov, B. Savelyev and V. V. Arlazarov, "Next integrated result modelling for stopping the text field recognition process in a video using a result model with per-character alternatives," *ICMV 2019*, vol. 11433, pp. 114332M1-114332M6, 2020, DOI: 10.1117/12.2559447.
 39. K. B. Bulatov and V. V. Arlazarov, "Determining optimal frame processing strategies for real-time document recognition systems," *ICDAR 2021*, vol. 12822, pp. 273-288, 2021, DOI: 10.1007/978-3-030-86331-9_18.

40. K. B. Bulatov, N. V. Fedotova and V. V. Arlazarov, "Fast Approximate Modelling of the Next Combination Result for Stopping the Text Field Recognition in a Video Stream," ICPR 2020, pp. 239-246, 2021, DOI: 10.1109/ICPR48806.2021.9412574.
41. D. V. Polevoy, M. A. Aliev and D. P. Nikolaev, "Choosing the best image of the document owner's photograph in the video stream on the mobile device," ICMV 2020, vol. 11605, 116050F, pp. 116050F1-116050F9, 2021, DOI: 10.1117/12.2586939.
42. M. Al-Ghadi, P. Gomez-Kramer, J.-C. Burie. CheckScan: a reference hashing for identity document quality detection. ICMV 2021, 120840J, DOI: 10.1117/12.2623887.
43. E. Myasnikov and A. Savchenko, "Detection of Sensitive Textual Information in User Photo Albums on Mobile Devices," 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), 2019, pp. 0384-0390, DOI: 10.1109/SIBIRCON48586.2019.8958325.
44. L. Kopeykina and A. V. Savchenko, "Automatic Privacy Detection in Scanned Document Images Based on Deep Neural Networks," 2019 International Russian Automation Conference (RusAutoCon), 2019, pp. 1-6, DOI: 10.1109/RUSAUTOCON.2019.8867614.
45. H. Ahmed, I. Traore, S. Saad, M. Mamun. Automated detection of unstructured context-dependent sensitive information using deep learning. Internet of Things, vol. 16, 100444, 2021, DOI: 10.1016/j.iot.2021.100444.
46. Startseva, A. Vulfin, V. Vasilyev, A. Nikonov and A. Kirillova, "Analysis of Financial Payments Text Labels in the Dynamic Client Profile Construction," 2020 International Conference on Information Technology and Nanotechnology (ITNT), 2020, pp. 1-10, DOI: 10.1109/ITNT49337.2020.9253280.
47. M. Al-Ghadi, Z. Ming, P. Gomez-Kramer, J.-C. Burie. Identity documents authentication based on forgery detection of guilloche pattern. ArXiv preprints, 2206.10989, URL: <https://arxiv.org/abs/2206.10989v1>.
48. Kada, O., Kurtz, C., van Kieu, C., Vincent, N. Hologram Detection for Identity Document Authentication. In: El Yacoubi, M., Granger, E., Yuen, P.C., Pal, U., Vincent, N. (eds) Pattern Recognition and Artificial Intelligence. ICPRAI 2022. Lecture Notes in Computer Science, vol 13363. Springer, Cham, 2022. DOI: 10.1007/978-3-031-09037-0_29
49. Chen, L. Zhao, J. Yan, H. Li. A distortion model-based pre-screening method for document image tampering localization under recapturing attack. Signal Processing, vol. 200, 108666, 2022. DOI: 10.1016/j.sigpro.2022.108666.
50. V. Polevoy, I. V. Sigareva, D. M. Ershova, V. V. Arlazarov, D. P. Nikolaev, M. Zuheng, M. L. Muhammad and J. Burie, "Document Liveness Challenge dataset (DLC-2021)," J. Imaging, vol. 8, no 7, pp. 181-1-181-12, 2022, DOI: 10.3390/jimaging8070181.

Arlazarov V. V. Head of the Department for the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russia. PhD. Number of articles: 130. Research interests: artificial intelligence, machine learning, recognition systems, information technology, queueing theory. E-mail: vva777@gmail.com