

РАСПОЗНАВАНИЕ РЕКВИЗИТОВ БАНКОВСКИХ КАРТ В МОБИЛЬНЫХ УСТРОЙСТВАХ ПО ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯМ

© 2020 г. Х. Чен^{a,*}, Ш. Е^{a,**}, А. Курилович^{b,***},
Р. Богущ^{b,****}, С. Абламейко^{c,d,*****}

^a Университет Чжэцзян Шурэнь,
Ханчжоу, ул. Шурэнь, 8, 310009 Китай

^b Полоцкий государственный университет,
Новополоцк, ул. Блохина, 29, 211440 Республика Беларусь

^c Белорусский государственный университет,
Минск, п-т Независимости, 4, 220030 Республика Беларусь

^d Объединенный институт проблем информатики НАН Беларуси,
Минск, ул. Сурганова, 6, 220012 Республика Беларусь

*E-mail: eric.hf.chen@hotmail.com

**E-mail: zjsruysp@163.com

***E-mail: cfif921@yandex.ru

****E-mail: bogushr@mail.ru

*****E-mail: ablameyko@bsu.by

Поступила в редакцию 08.10.2019 г.

После доработки 05.11.2019 г.

Принята к публикации 05.12.2019 г.

Предлагается алгоритм для обнаружения и распознавания всех информационных полей лицевой стороны банковских карт по видеопоследовательностям, который предназначен для использования в мобильных устройствах. Данный алгоритм включает следующие основные шаги: обнаружение границ карты в кадре, сегментация информационных полей, улучшение изображений сегментов, выделение границ символов и распознавание блоков символов. Доля истинно положительной классификации предложенным алгоритмом всех трех полей данных на карте составляет 0.88, для номера и срока действия карты — 0.925. Представлены временные затраты на обработку видеопоследовательностей с различным разрешением кадра с применением iPhone7. Результаты экспериментов подтверждают эффективность предложенного подхода.

DOI: 10.31857/S0132347420060023

1. ВВЕДЕНИЕ

Развитие банковских технологий привело к широкому внедрению интернет-банкинга, т.е. дистанционного банковского обслуживания клиентов через сеть Интернет, а также мобильного банкинга, который предоставляет возможность управления банковскими счетами с помощью планшетного компьютера или смартфона. В настоящее время у одного клиента банка в наличии может находиться две и более банковские карты, которые он может использовать для оплаты. Заполнение всех полей данных при оплате услуг или осуществлении покупок с использованием мобильного банкинга требует временных затрат и внимательности. Поэтому, совершенствование мобильного банкинга предполагает разработку

алгоритмического и программного обеспечения для автоматизации ввода реквизитов банковской карты в систему на основе анализа ее изображения, полученного с использованием мобильного устройства.

Дизайн банковской карты является важным носителем бренда банка, что не накладывает практически никаких ограничений на яркостно-цветовые характеристики при его разработке. Таким образом, изображение на карте может быть неоднородным и достаточно сложным. Кроме того, пластик с глянцево-поверхностью, из которого, как правило, изготавливаются банковские карты, обладает сильными отражающими характеристиками, что при ярком освещении приводит к возникновению бликов и отражений на полу-

ченных изображениях, особенно при использовании вспышки в условиях недостаточной освещенности. При наличии теней или недостаточном освещении изображение карты может оказаться достаточно низкого качества.

Банковская карта является типичным примером документа с гибкой формой, поэтому при ее обработке могут быть использованы алгоритмы обнаружения и распознавания данных, которые используются на других типах документов с гибкой формой [1]. В работе [2] представлен подход для сегментации и распознавания текстовых символов на визитных картах с однотонным фоном. При этом используются такие основные шаги, как выделение краев оператором Собеля, утоньшение линий, проективное преобразование, адаптивная бинаризация, сегментация слов и текстовых символов с последующим их распознаванием. Применение данного подхода ограничивается возможностью работы только с простым фоном и достаточно значительными вычислительными затратами, что накладывает ограничение на его реализацию в мобильных устройствах. В работе [3] предложен алгоритм обработки визитных карт на первом этапе которого выполняется грубое удаление фона на основе блочного анализа входного изображения, на втором этапе классифицируются полученные связные компоненты для определения текстовых регионов и исключения изображений логотипов. Затем к результату применяется адаптивная бинаризация, которая позволяет отделить текст от фона для дальнейшего распознавания. Данный алгоритм также может быть использован только для изображений карт с монотонным фоном. Кроме этого, не предусмотрено проективное преобразование, что накладывает дополнительные требования к ориентации входного изображения относительно камеры смартфона.

Методы распознавания данных визитных карт для платформы Андроид представлены в [4] и [5]. В [4] используется программный пакет Matlab для предварительной обработки и сегментации текстовых символов, и библиотека Tesseract для их распознавания. Однако, как отмечают авторы, предварительная обработка используется ряд операций с большой вычислительной сложностью по улучшению входного изображения и сегментации символов, поэтому на мобильном устройстве реализован упрощенный вариант алгоритма с отсутствием определенных этапов предварительной обработки, что привело к значительному снижению качества его работы. На первом шаге метода из [5] выполняется уменьшение размера входного изображения, операции предобработки выполняются с использованием библиотеки OpenCV, распознавание символов осуществляется на основе Tesseract. Однако результаты экспериментов

в работах представлены для изображений визитных карт с однотонным фоном, на котором хорошо видны текстовые символы.

Алгоритм распознавания только даты окончания действия кредитной карты и оценка его качественных характеристик с использованием смартфона iPhone 4S представлен в [6]. Его особенностью является то, что он не обеспечивает автоматизации ввода всех данных с изображения и, соответственно, требуется дальнейшее его развитие. Возможности алгоритмов распознавания номеров банковских карт для мобильных устройств под управлением операционной системы Android, рассмотренных в [7] и [8], также не включает распознавание имени и фамилии пользователя. В [7] точность распознавания достигает 80%, а в [8] использование предварительной обработки изображений карт и рекуррентной нейронной сети обеспечивает точность распознавания цифр до 90%. Следует также отметить, что данные алгоритмы предназначены для работы с эмбоссированными картами, у которых на лицевой стороне цифры и буквы выдавлены. В [9] приведен метод построения функции надежности распознавания изображений тисненых символов и показана его эффективность для повышения вероятности их правильного распознавания. Однако, широкое распространение получили неэмбоссированные карты, которые, могут выдаваться клиенту в день его обращения в банк. Но из-за отсутствия процедуры выдавливания буквенно-цифровой информации ее локализация и распознавание значительно сложнее и требует эффективной предварительной обработки для локализации символов и обеспечения возможности их правильного распознавания.

Существующие подходы используют в качестве входных данных статические изображения, на которых в ряде случаев могут возникать шумы, блики, значительные перепады яркости, иметь недостаточную резкость и т.д., что приводит к необходимости получения нового изображения, которое также может быть непригодно для правильного обнаружения карты, информационных полей на ней и (или) дальнейшего их распознавания. Повышение результативности обработки данных может быть обеспечено за счет использования видео, которое легко получается с помощью мобильного телефона. Однако, наличие не одного изображения, а их последовательности приводит к необходимости разработки других подходов к обработке. В работе [10] было предложено использовать видеопоследовательности для обнаружения краев документов мобильными устройствами. При этом, сегментация изображений выполняется в цветовом пространстве Lab с использованием морфологической обработки, связывание границ

отдельных сегментов на основе преобразования Хафа и обнаружение линий, ограничивающих документ в целом. Однако, для сокращения вычислительных затрат применяется уменьшение размера кадра входного изображения с 1280×720 до 180×100 пикселей. Такое уменьшение кадра не пригодно для решения задачи локализации и распознавания информации на банковской карте, так как вероятность распознавания символов будет низкой из-за их малых размеров.

В данной работе предлагается алгоритм для обнаружения и распознавания реквизитов банковских карт по видеопоследовательностям в мобильных устройствах. Алгоритм позволяет практически в реальном времени распознавать все информационные поля лицевой стороны банковской карты, включая номер, дату окончания действия, имя и фамилию владельца на латинице и на кириллице, для эмбоссированных и неэмбоссированных типов карт.

2. АЛГОРИТМ РАСПОЗНАВАНИЯ РЕКВИЗИТОВ БАНКОВСКИХ КАРТ ПО ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯМ

Разработанный алгоритм показан на рис. 1. Последовательность кадров, полученная с помощью видекамеры мобильного устройства, поступает в блок детектирования прямоугольных областей, который обнаруживает все прямоугольники в кадре и возвращает только один, удовлетворяющий параметрам банковской карты.

Для данной области выполняется преобразование в полутоновое изображение, которое передается в блок сегментации для выделения и индексации областей на изображении карты, которые соответствуют таким информационным полям, как номер карты, срок ее действия, имя и фамилия владельца. Для сегментированных блоков выполняется улучшение контраста фильтрации. После этого применяется адаптивная бинаризация и морфологическая обработка. Следующий шаг предназначен для уточнения границ областей с символами и использует метод скользящего окна. После этого, сегментированные области передаются в блок OCR для распознавания цифровых и текстовых символов. В завершение, блок оценки данных обрабатывает полученную информацию для отображения либо отклонения результата. Если полученные номер карты и дата срока ее окончания действия являются валидными, результат для них считается правильным. Распознавание данных поля имени владельца прекращается после первого полученного результата. Если на текущем кадре считана информация не со всех трех информационных полей, то на следующем кадре обрабатываются только недостающие.

2.1. Обнаружение карты

При разработке алгоритма обнаружения карты учитывается, что размер кадра входной видеопоследовательности в современных мобильных устройствах варьируется и может достигать разре-



Рис. 1. Обобщенная схема алгоритма.

шения 3840×2160 пикселей, также устройствами обеспечивается автоматическое удержание фокуса на объекте, корректировка яркости и баланс белого. Таким образом, полученное изображение будет, как правило, иметь приемлемые следующие характеристики: резкость, яркость и достаточное разрешение кадра для автоматического извлечения данных банковских карт.

На первом этапе осуществляется детектирование и локализация всех прямоугольных областей на изображении. С учетом возможного большого разрешения кадра, до 4 К, но ограниченных вычислительных ресурсов мобильных устройств, применяется метод поиска, основанный на быстрых алгоритмах Виолы–Джонса [11] и OverFeat [12], обеспечивающий также высокую результативность обнаружения. Далее, для определения контура, описывающего банковскую карту, используется анализ отношения размеров сторон полученных объектов. Поскольку высота m_o и ширина n_o карты стандартизированы, отношение размеров ее сторон является постоянной величиной, то целесообразно ее использовать в качестве критерия при нахождении контура карты из всех детектированных прямоугольных областей. Тогда прямоугольная область r_i с размерами $m_i \times n_i$ может быть отнесена к изображению карты в кадре, если выполняется условие:

$$\frac{n_o}{m_o} - e < \frac{n_i}{m_i} < \frac{n_o}{m_o} + e, \quad (2.1)$$

где e — коэффициент допустимого отклонения соотношения размеров сторон карты.

Однозначным соответствием считается прямоугольник с наибольшей длинной стороны m_{max} из всех детектированных. Найденные на данном этапе размеры прямоугольной области и ее расположение относительно всего изображения $p_i(x_i, y_i)$ используются для извлечения области карты из входного кадра.

2.2. Сегментация областей интереса

На полученном изображении карты I с размерами $m_I \times n_I$ необходимо найти области, содержащие информацию о номере банковской карты (I_C), дате истечения срока ее действия (I_D) и владельце (I_E). Размер и расположение этих областей определено ISO/IEC 7811-5.4:2018, что может быть использовано для сегментации этих полей, формально представлено как:

$C(x_C, y_C, m_C, n_C)$ — номер карты;

$D(x_D, y_D, m_D, n_D)$ — дата окончания срока действия карты;

$E(x_E, y_E, m_E, n_E)$ — имя и фамилия держателя карты.

Поскольку значения m_I и n_I могут варьироваться в процессе видеосъемки, необходимо предусмотреть нормализацию, чтобы корректно определять C , D и E . Для этого используются коэффициенты масштабирования, которые умножаются на параметры этих областей, по ширине ($c_{ми}$) и по высоте ($c_{мв}$):

$$c_{ми} = \frac{n_I}{n_o}, \quad c_{мв} = \frac{m_I}{m_o}. \quad (2.2)$$

Номер банковской карты содержит 16 цифр, распределенных на 4 равные группы по 4 цифры в каждой. Исходя из этого выделенное на предыдущем этапе изображение I_C с областью номера карты C_I , разделяется на 4 равные области C_{I1} , C_{I2} , C_{I3} , C_{I4} :

$$C_{Ij} \left(x_{CI} + \left(\frac{n_{CI}}{4} \times (j-1) \right), y_{CI}, \frac{n_{CI}}{4}, m_{CI} \right), \quad (2.3)$$

$$j = 1, 2, 3, 4.$$

Выделив рассчитанные области из I_C , получим изображения сегментов групп номера карты — I_{C1} , I_{C2} , I_{C3} , I_{C4} .

2.3. Улучшение изображения сегментов

Для уменьшения количества обрабатываемой информации полученные изображения (I_D , I_E , $I_{C1} - I_{C4}$) преобразуются в полутоновые с дальнейшим повышением их контраста методом нормализации гистограммы, который обеспечивает растяжку не всего диапазона изменения интенсивностей, а только его наиболее информативного участка, что усиливает эффект контрастности за счет потери шумовых областей с редко встречающимися интенсивностями.

При таком подходе выходной уровень яркости g для пикселя определяется следующим образом:

$$g_{x,y} = \frac{255 \cdot (Y_{x,y} - Y_{c_{min}})}{Y_{c_{max}} - Y_{c_{min}}}, \quad (2.4)$$

где $Y_{x,y}$ — уровень яркости входного пикселя с координатами x, y ; $Y_{c_{min}}$, $Y_{c_{max}}$ — заданные значения уровней пикселей входного изображения, минимальное и максимальное соответственно.

Для улучшения результатов обработки $Y_{c_{min}}$, $Y_{c_{max}}$ отличны от минимального и максимального уровней яркости входного изображения, что позволяет пренебречь незначительным количеством пикселей по краям гистограммы.

Таким образом, значения $Y_{c_{min}}$, $Y_{c_{max}}$ могут быть определены как:

$$Y_{c_{\min}}(i) = \begin{cases} Y_{c_{\min}}(i+1), & \text{если } \sum_{j=0}^i f_i(Y_j) < T; \\ Y_i, & \text{в противном случае} \end{cases} \quad (2.5)$$

$$Y_{c_{\max}}(i) = \begin{cases} Y_{c_{\max}}(i-1), & \text{если } \sum_{j=255}^i f_i(Y_j) < T, \\ Y_i, & \text{в противном случае} \end{cases}$$

где f_i – количество пикселей на изображении с яркостью Y_j ; $T = m \cdot n \cdot e$; e – коэффициент допустимого отклонения (%), определяемый экспериментально.

Далее, с использованием операций математической морфологии, осуществляется подчеркивание границ символов на полученном с предыдущего этапа полутоновом изображении. Для того чтобы определить цвет фона и символов рассчитывается среднее значение яркости. Если результат больше 127, то цвет фона светлый, а цвет символов темный, и наоборот.

Если цвет символов определен как светлый, для изображения выполняется морфологическое преобразование WhiteTopHat, которое вычитает замкнутое изображение из исходного, что обеспечивает подчеркивание деталей контуров светлых символов. Преобразование BlackTopHat используется, чтобы подчеркнуть детали границ темных символов путем вычитания исходного изображения из замкнутого. Структурирующий элемент b для ядер обоих фильтров имеет прямоугольную форму и его размер $n_b \times m_b$ с учетом экспериментально определенных коэффициентов масштабирования рассчитывается как:

$$n_b = 0.06n_l, \quad m_b = 3n_b, \quad (2.6)$$

где n_l , ширина поступившего для преобразования изображения.

2.4. Локализация границ

Поскольку на изображениях банковских карт, несмотря на их небольшой размер, яркостно-контрастные характеристики фрагментов могут значительно отличаться, уровни яркостей пикселей участков фона, на которых расположены символы, также изменяются. С учетом данной особенности и необходимости сохранения границ символов, используется метод бинаризации изображений с адаптивным порогом на основе анализа локально-го региона согласно выражению:

$$bin_{x,y} = \begin{cases} 255, & \text{если } Y_{x,y} > T(x,y) \\ 0, & \text{в противном случае} \end{cases} \quad (2.7)$$

где $T(x, y)$ – адаптивный порог бинаризации, рассчитываемый для каждого пикселя как $T(x, y) = corr(I, G) - c$; $corr(I, G)$ – взаимная корреляция локального фрагмента изображения размером $k \times k$ с окном Гаусса; c – константа. Матрица коэффициентов окна Гаусса с размером $k \times k$ определяется как [13]:

$$G_i = \alpha \cdot e^{-\frac{(i - \frac{k-1}{2})^2}{2\sigma^2}}, \quad (2.8)$$

$$\sigma = 0.3 \cdot ((k-1) \cdot 0.5 - 1) + 0.8,$$

где $i = 0 \dots k-1$, α – коэффициент масштабирования, выбранный таким образом, что $\sum_i G_i = 1$

Для уменьшения количества шумов и удаления из изображения неинформативных деталей, а также для грубого отделения символов от фона изображения применяются последовательно морфологические операции замыкания и эрозии.

Форма структурирующего элемента b размером $n_b \times m_b$ ($n_b = \frac{w_{\text{symp}}}{3}$) ядер обоих фильтров задана с учетом особенностей символов шрифта согласно ISO/IEC 7811-7.1:2018 в виде эллипса. Толщина символа w_{symp} не зависит от типа обрабатываемого фрагмента изображения карты и определена для OCR-B. Проекция значения ширины шрифта (мм) на размер в пикселях, относительно размера области банковской карты, является константой и рассчитывается на основе ее полной ширины:

$$w_{\text{symp}} = 0.004884 \cdot n_0. \quad (2.9)$$

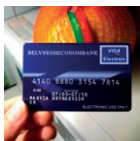
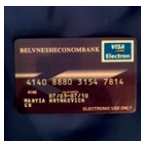
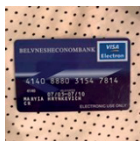
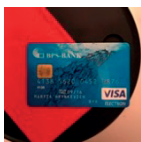
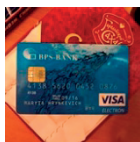
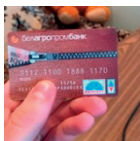
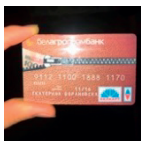
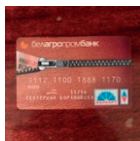
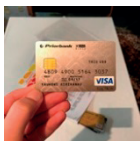
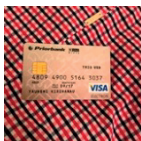
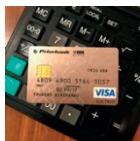
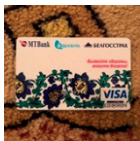
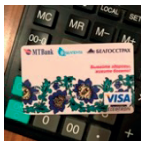
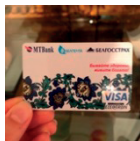
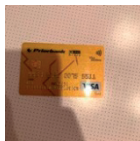
Далее выполняется уточнение краев области символов на обработанном изображении на базе метода скользящего вертикального окна. Высота окна варьируется по размеру шрифта символов. $H_{\text{нк}}$ – высота шрифта номера карты (4.0 мм), $H_{\text{дси}}$ – высота шрифта даты срока истечения карты (2.85 мм), $H_{\text{увк}}$ – высота шрифта имени владельца карты (2.65 мм). Тогда их проекции на высоту в пикселях относительно размеров области полной карты $h_{\text{нк}}$, $h_{\text{дси}}$, $h_{\text{увк}}$. Общий вид расчета представлен формулой:

$$h = \frac{H \cdot m_0}{54.0}, \quad (2.10)$$

где 54.0 мм – высота карты (ISO/IEC 7811-6.3:2018), а m_0 – высота детектированной области карты (пикс).

Полученная область символов из предыдущего шага, черный и белый списки символов, а также

Таблица 1. Результаты тестирования алгоритма

Кадр из видеоряда				Карта обнаружена, (число кадров)	Число кадров с правильным распознаванием всех реквизитов/общее число кадров	Общее время, затраченное на обработку (мс)
1	2	3	№			
			1	22	9/90	1217
			2	—	—/90	—
			3	26	10/90	1221
			1	23	10/90	1284
			2	26	11/90	1149
			3	31	13/90	1265
			1	24	—/90	—
			2	21	7/90	1344
			3	—	—/90	—
			1	30	18/120	852
			2	37	21/120	924
			3	36	26/120	892
			1	39	—/135	—
			2	44	—/135	—
			3	41	—/135	—
			1	34	24/90	1324
			2	35	28/90	1238
			3	32	25/90	1287

языковой идентификатор последовательно передаются в систему распознавания Tesseract.

Белый список, состоящий из набора цифр от 0 до 9 и языковой идентификатор “eng”, передаются в качестве параметров для изображения номера банковской карты. Для изображения с датой срока ее истечения: цифры от 0 до 9 и символ “/” образуют белый список, а “eng” — идентификатор языка. Для имени держателя карты: “rus” и “eng” являются языковыми идентификаторами, а набор символов, включая знаки препинания, специальные символы и цифры, составляют черный список. Такой подход способствует увеличению

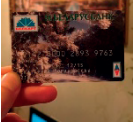
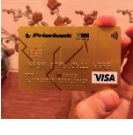
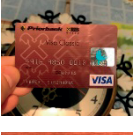
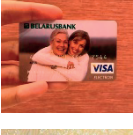
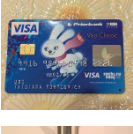
скорости распознавания символов библиотекой Tesseract и ускорению всего процесса обработки в целом.

3. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

3.1. Программная реализация

Программная реализация алгоритма выполнена на языке Objective-C с применением библиотеки низкоуровневой обработки изображений OpenCV, системы оптического распознавания символов Tesseract, фреймворков iPhone SDK: CoreMedia и AV-

Таблица 2. Результаты сравнения подходов для распознавания информационных полей банковских карт

Исходные данные		CardIO			Предложенный алгоритм			
Изображение карты	Эмбоссинг символов	Время обработки при успешном распознавании	Результат распознавания информационного поля*		Время обработки при успешном распознавании	Результат распознавания информационного поля*		
			Номер карты	Срок действия		Номер карты	Срок действия	Держатель карты
	да	0.343	1	1	1.018	1	1	1
	нет	—	0	0	0.896	1	1	1
	да	0.314	1	1	1.059	1	1	0
	да	0.916	1	0	0	0	0	0
	нет	—	0	0	1.01	1	1	1
	да	0.864	1	1	0.971	1	1	1
	да	0.48	1	1	0.861	1	1	0
	нет	—	0	0	—	0	0	0
	да	0.478	1	0	0.957	1	1	1
	нет	—	0	0	0.986	1	1	1

* Примечание: 1 — соответствует успешному распознаванию информационного поля; 0 — соответствует неверному распознаванию.



Рис. 2. Внешний вид мобильного приложения, реализующего разработанный алгоритм: 1 — область просмотра изображения; 2 — метка успешного детектирования карты; 3 — область вывода распознанных данных.

Foundation — менеджмент медиаданных; UIKit — работа с интерфейсом приложения; CoreGraphics — низкоуровневая обработка 2D изображений на базе Quartz.

Фрэймворк AVFoundation используется для захвата видеопотока с основной камеры iOS-устройства в режиме реального времени с частотой 30 кадров в секунду. Кадры размещаются в очередях в отдельных потоках при помощи объектов класса `NSOperationQueue`. Первый поток получает кадры и размещает их в последовательной очереди для обработки, если последняя не занята, в противном случае кадры игнорируются. Второй поток последовательно извлекает кадры из очереди захвата и, передавая их на обработку, занимает очередь на период, пока не будут получены результаты. После завершения алгоритма, обработанный кадр удаляется, очередь освобождается, и итерация обработки запускается для следующего кадра.

Библиотека Tesseract выполняет обработку асинхронно в фоновом режиме. После окончания работы алгоритма полученный результат распознавания при помощи функции обратного вызова поступает для отображения пользователю в главном потоке. Интерфейс мобильного приложения с результатами распознавания показан на рис. 2.

3.2. Результаты экспериментов

Тестирование качественных характеристик разработанного алгоритма выполнено с использованием iPhone 7 и созданной базы данных из 180 банковских карт (рис. 3), которые размещались на сложном фоне, информационные поля реквизитов не были перекрыты посторонними предметами и визуально распознавались.

Значения коэффициентов, которые используются при вычислениях, были экспериментально получены с использованием методики ROC-анализа [14]. Вычисление допустимого коэффициента отклонения соотношения размеров сторон карты банковской карты (e) выполнено с учетом точных размеров ее согласно ISO/IEC 7811-5.4:2018 на выборке из 214 изображений, 178 из которых имели объекты карт, 36 — карт не имело, либо их границы были смазаны или искажены. В результате эксперимента для каждого коэффициента e , который изменялся от 0 до 0.02 с шагом в 0.0002, были рассчитаны доли истинного (TPR) и ложного (FPR) обнаружения. Для $e = 0.011$ получено максимальное значение $TPR = 0.974194$ и минимальное $FPR = 0.111111$.

Определение размера локального региона для бинаризации выполнено с учетом коэффициента масштабирования, так как данный параметр зависит от входных размеров детектированной области банковской карты (ее ширины (W_I , пикс.)):

$$k = W_I \cdot k'. \quad (3.1)$$

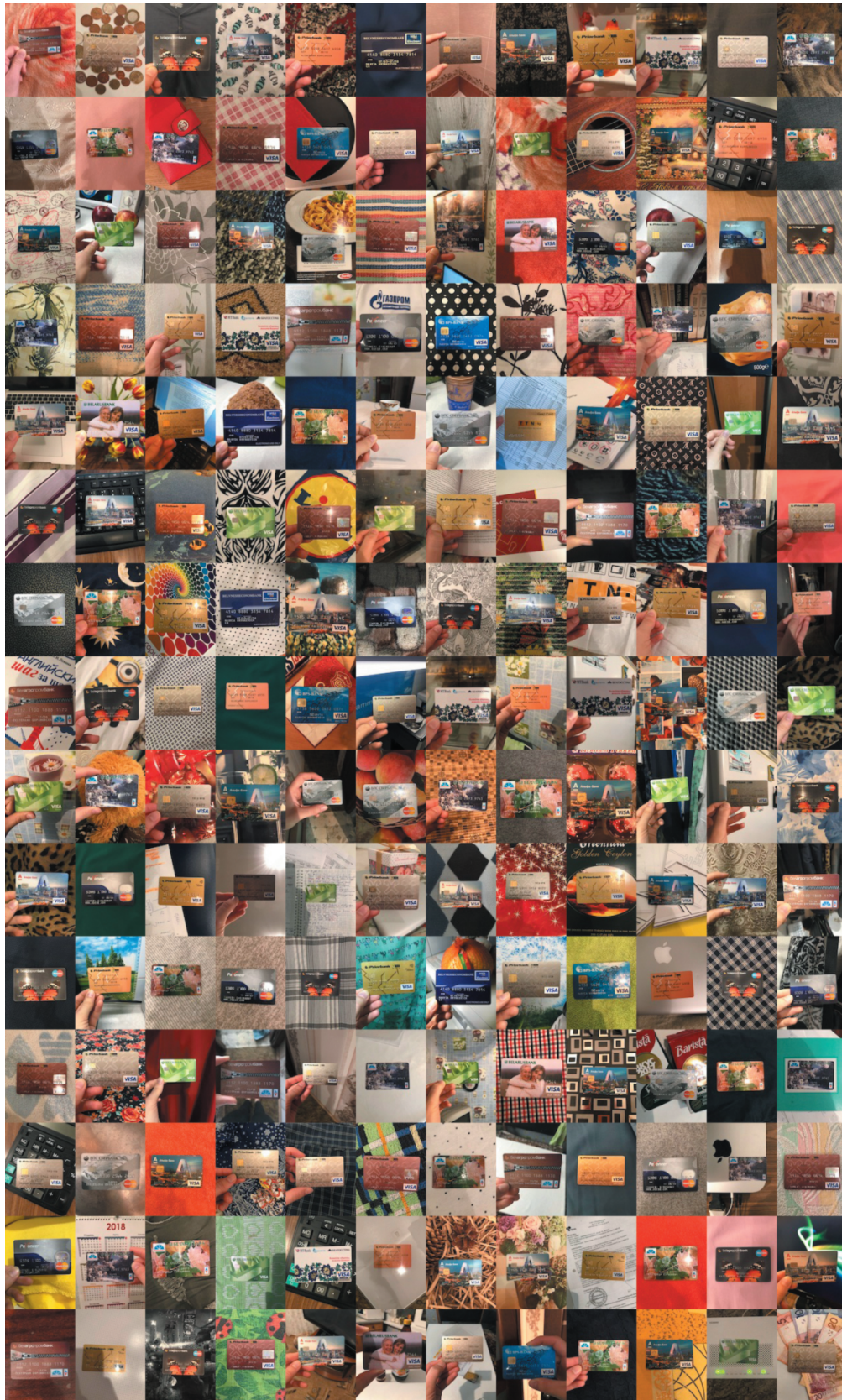


Рис. 3. Изображения используемых для тестирования банковских карт.

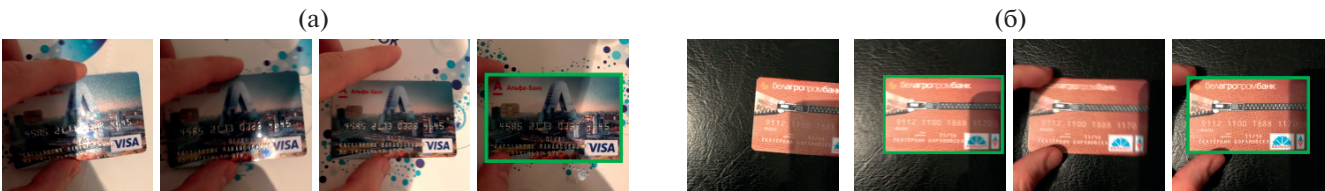


Рис. 4. Пример кадров видеопоследовательности.

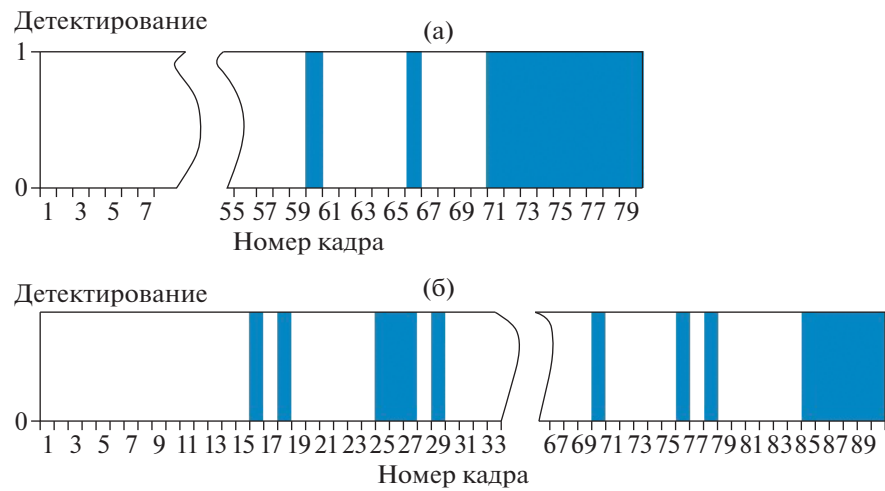


Рис. 5. Диаграмма детектирования карты на видеопоследовательности: а) для рис. 4а; б) для рис. 4б.

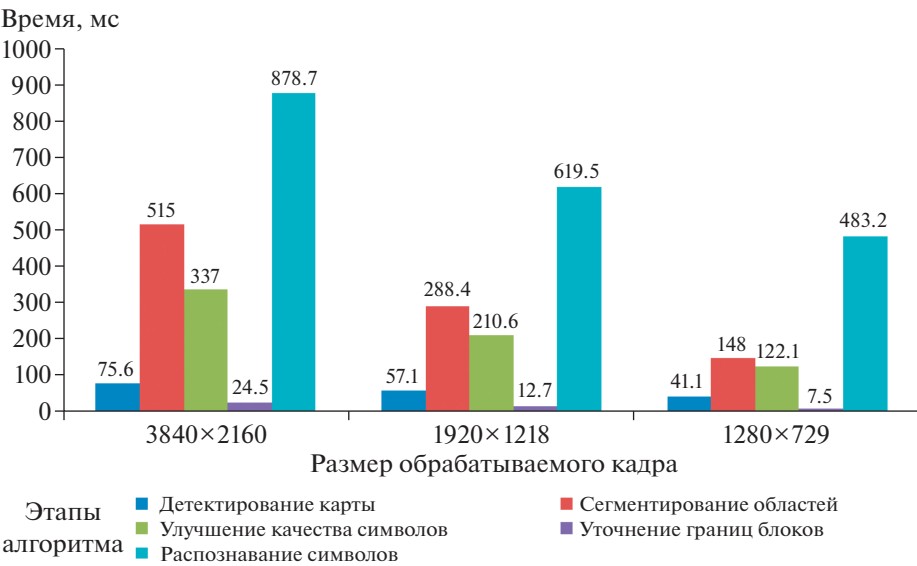


Рис. 6. Распределение временных затрат алгоритма.

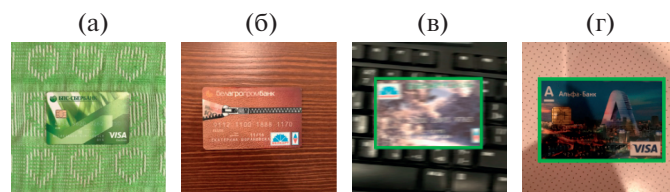


Рис. 7. Примеры изображений карт, которые не обнаруживаются или данные на них не распознаются.

Значения k' варьировались от 0.002 до 0.03 с шагом 0.001. На данном этапе максимальное значение доли истинного распознавания символов $TPR = 0.83447$ и доли ложного распознавания символов $FPR = 0.140864$ получено для $k' = 0.016$. Аналогично определена константа $c = 13$.

В табл. 1 представлены некоторые результаты выполненных исследований по оценке эффективности распознавания реквизитов карт на видеопоследовательностях, полученных в различных условиях, с применением разработанного подхода.

Согласно полученным с помощью ROC-анализа результатам, после установки оптимальных коэффициентов и пороговых значений в реализации алгоритма, доля истинно положительной классификации всех полей данных на карте $TPR = 0.88$, а доля ложноположительной классификации $FPR = 0.14$. На рис. 4–6 приводятся примеры детектирования банковских карт из набора видеопоследовательностей, с указанием ключевых кадров и диаграмм детектирования. Обнаружению карты на диаграмме соответствует – 1, отсутствию – 0. Захват видео осуществлялся при постоянной частоте 30 кадр/с.

Анализ кадров видеопоследовательности, представленной на рис. 4а, свидетельствует, что изображение карты имеет сильные яркостные перепады, тени и блики, поэтому устойчивое ее обнаружение (начиная с 71-го кадра, рис. 5а) выполняется когда пользователь переносит карту в зону более стабильного освещения. На другой видеопоследовательности (рис. 4б) первоначально карта не имеет полных границ в кадре и не обнаруживается. Далее она успешно детектируется (рис. 5б), но имеет смазанный контур и нечеткие символы, что не позволит распознавать данные.

Временные затраты для каждого этапа обработки алгоритма одного кадра видеоряда различного разрешения с использованием устройства iPhone 7 показаны на рис. 6.

Как видно из графиков (рис. 6), время, затрачиваемое на операцию распознавания символов, имеет нелинейную зависимость от размера обра-

батываемого кадра, в отличие от остальных операций.

3.3. Сравнение и обсуждение

Алгоритмы, предложенные в [7] и [8], характеризуются точностью распознавания цифр 80% и 90% соответственно. Однако, следует отметить, результаты получены на разных базах данных. Для обеспечения сравнения результатов работы представленного алгоритма с известными подходами в одинаковых условиях на одной и той же базе данных использовался реализованный модуль CardIO SDK [15].

Эксперименты выполнены для 180 карт при различных условиях съемки, из которых для 104 случаев на картах присутствовало тиснение символов. В табл. 2 представлены примеры распознавания 10 изображений карт алгоритмом CardIO и предложенным алгоритмом.

Анализ результатов распознавания по всем картам используемой базы данных показывает, что разработанный подход значительно улучшает распознавание на картах с неэмбоссированными символами по сравнению с CardIO, который не предусматривает распознавание имени владельца карты. В общем, доля истинно положительной классификации предложенным алгоритмом всех трех полей данных на карте составляет $TPR = 0.88$, для номера и срока действия карты – $TPR = 0.925$. Модуль CardIO обеспечивает долю истинно положительной классификации цифровых полей

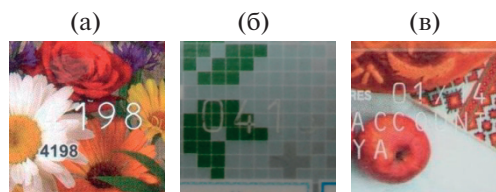


Рис. 8. Примеры фрагментов карт с поврежденными символами.

TPR = 0.68. Однако, только для карт с тиснением используемой базы данных CardIO характеризуется TPR = 0.935.

Если на входе разработанного алгоритма будут яркостно-цветовые характеристики изображения банковской карты очень схожи с общим фоном (рис. 7а, 7б), входное изображение размыто (рис. 7в), значительная область покрыта тенью, в результате чего цвет шрифта полностью совпадает с его фоном (рис. 10г), карты не могут быть им детектированы (рис. 7а, 7б) или данные на них не будут распознаны (рис. 7в, 7г). Кроме этого, невыразительный, не создающий контрастности с фоном карты контур символов, а также его повреждение или дефективность приводят к невозможности распознавания данных (см. рис. 8).

Анализ полученных результатов при проведении исследований показывает, что ошибки детектирования и распознавания возникают из-за влияния совокупности факторов: расположение объекта карты на сложном фоне, схожем с изображением самой карты; наличия бликов искусственного освещения или существенный перепад яркости; слишком удаленное ее расположение относительно камеры; смазывание контуров карты если камера или карта движется при видеосъемке. Дальнейшее улучшение работы алгоритма планируется за счет совершенствования алгоритма распознавания символов.

4. ЗАКЛЮЧЕНИЕ

Распознавание информационных полей банковских карт с использованием мобильных устройств является актуальной задачей. В данной статье представлен алгоритм для распознавания реквизитов лицевой стороны банковских карт с использованием мобильных устройств. Повышение результативности обнаружения карты и распознавания данных по сравнению с существующими подходами обеспечивается за счет анализа последовательности изображений, получаемой из видеоряда. Предложенный подход включает обнаружение границ карты в кадре на основе алгоритма Виолы–Джонса и метода OverFeat, сегментацию информационных полей с учетом их расположения на карте, улучшение изображений сегментов с применением нормализации гистограммы и морфологической обработки, определение границ блоков символов на основе адаптивной бинаризации и морфологической обработки, их уточнение с использованием скользящего вертикального окна, распознавание символов с применением библиотеки Tesseract. Программная реализация разработанного алгоритма выполнена с примене-

нием iPhone SDK, библиотек OpenCV и Tesseract. Тестирование предложенного алгоритма проведено с использованием мобильного устройства iPhone 7. При этом, доля истинно положительной классификации предложенным алгоритмом всех трех полей данных на карте составляет TPR = 0.88, для номера и срока действия карты – TPR = 0.925.

СПИСОК ЛИТЕРАТУРЫ

1. *Sheshkus A., Nikolaev D., Ingacheva A., Skoryukina N.* Approach to the recognition of flexible forms on the example of the credit card date recognition. Proceedings SPIE 9875, Eighth International Conference on Machine Vision (ICMV 2015), Barcelona, Spain, 8 December 2015, session 3. P. 83–88.
2. *Christian T., Gustavsson D.* Content recognition of business cards. Summer Project, IT University of Copenhagen, Copenhagen, Denmark, 5 September 2007.
3. *Mollah A., Basu S., Das N., Sarkar R., Nasipuri M., Kundu M.* Text Region Extraction from Business Card Images for Mobile Devices. Proceedings of International Conference on Information Technology and Business Intelligence (ITBI-09), Nagpur, India, 6–8 November 2009. P. 227–235.
4. *Bhaskar S., Lavassar N., Green S.* Implementing on the Android Operating System for Business Cards. IEEE 2010, EE 368 Digital Image Processing, Stanford, CA 7 June 2010. P. 1–5.
5. *Sharma P., Fujii K.* Automatic Contact Importer from Business Cards for Android. Stanford University, Stanford, CA, 2013.
6. *Hua G., Liu Z., Zhang Z., Wu Y.* Automatic business card scanning with a camera. International Conference on Image Processing, Atlanta, GA, USA 8–11 October 2006. P. 373–376.
7. *Cai S., Wen J., Xu H., Chen S., Ming Z.* Bank Card and ID Card Number Recognition in Android Financial APP, Lecture Notes in Computer Science. 2017. V. 10135. P. 205–213.
8. *Liu L., Huang L., Xue J.* Identification of serial number on bank card using recurrent neural network, Ninth International Conference on Graphic and Image Processing (ICGIP), Qingdao, China 14–16 October 2017. P. 319–325.
9. *Арлазаров В.В., Булатов К.Б., Карпенко С.М.* Метод определения надежности распознавания тисненных символов. Труды Института системного анализа РАН. 2013. Т. 63. № 3. С. 117–122.
10. *Puybureau E., Geraud T.* Real-Time Document Detection in Smartphone Videos. Proceedings of the 24th IEEE International Conference on Image Processing (ICIP), France, October 2018. P. 1498–1502.
11. *Viola P., Jones M.* Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, USA 8–14 December 2001, Section 1. P. 511–519.

12. *Sermanet P., Eigen D., Zhang X., Mathieu M., Fergus R., LeCun Y.* OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. International Conference on Learning Representations (ICLR), Banff, Canada 14–16 April 2014. P. 1055–1061.
13. Copyright 2011–2014, *OpenCV dev team*, 2019. OpenCV документация / Фильтрация изображений. <https://docs.opencv.org/2.4/modules/imgproc/doc/filtering.html#getgaussiankernel>
14. *Mason S.J., Graham N.E.* Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. / S.J. Mason, N.E. Graham // Q. J. R. Meteorol. Soc. 2002. V. 128. P. 2145–2166.
15. CardIO official site. Available: <https://www.card.io>