

**МЕТОД ЛОКАЛИЗАЦИИ ТЕКСТА НА ИЗОБРАЖЕНИИ С ОБУЧЕНИЕМ НА
ОСНОВЕ ВЕРОЯТНОСТНОЙ МОДЕЛИ СИМВОЛОВ**

Предложен метод локализации текста на изображении с обучением на основе вероятностной модели символов. Применение разработанных модели и метода позволит облегчить создание представительной выборки для обучения различных классификаторов для локализации текста. Экспериментальные исследования подтвердили дееспособность предложенной модели.

Ключевые слова: изображения, локализация текста, текстовая область, вероятностная модель символов, метод локализации текста.

S.G. ANTOSHCHUK, A.A. NIKOLENKO, TIEN T.K. NGUYEN
Odessa National Polytechnic University

**TEXT LOCALIZATION METHOD ON IMAGE WITH TRAINING BASED ON CHARACTERS PROBABILISTIC
MODELS**

Method of localization of the text on the image with the training on the basis of a probabilistic model of characters was discussed. Application of the developed models and methods will facilitate the creation of a representative sample of the different learning classifiers for localization of text. Experimental studies have confirmed the viability of the proposed model.

Keywords: images, text localization, text area, a probabilistic model of characters, text localization method.

Введение

Задача оптического распознавания текстов (Optical Character Recognition, OCR) имеет большое количество приложений в различных областях, таких как бухгалтерский документооборот, видеокорректорская и охранная деятельность, робототехника и др. Повышенный интерес к системам автоматического распознавания текстов вызван развитием человеко-машинных интерфейсов и систем искусственного интеллекта. Несмотря на то, что разработка методов распознавания текстов ведется уже несколько десятилетий, эта проблема по-прежнему далека от завершения для реальных изображений со сложным и неоднородным фоном, при отсутствии четких критериев отличия текста от фона [1-7]. Система оптического распознавания текстов на изображениях в таком случае, как правило, содержит две основные процедуры: локализация текстовой области (ТО) и непосредственное распознавание текста. Следует отметить, что успешное выполнение второй процедуры невозможно без качественных результатов первой.



Рис.1 Пример изображения с текстовыми областями на сложном фоне

Однако автоматическая локализация текстовых областей (ЛТО) является непростой задачей из-за переменчивых условий визуализации, связанных с освещением, разным положением текста по отношению к камере, из-за невозможности заранее предугадать, в какой области изображения расположен текст (особенно при наличии разнородной графической информации) и какое этот текст имеет искажение. Примерами текстовых областей на сложном фоне являются текстовые надписи на памятниках, номера домов и названия улиц на соответствующих указателях, надписи на рекламных плакатах, надписи на дорожных указателях, информационная текстовая строка на экране телевизора и т.д. (рис. 1).

Цель статьи

Целью статьи является разработка метода локализации текста на изображении с обучением на основе вероятностной модели символов. Использование предложенной модели упростит процесс получения обучающей выборки, обеспечив высокую достоверность классификации при локализации текста на изображениях со сложным фоном.

Решены следующие задачи: разработана обобщенная структура метода локализации, предложен метод генерации ТО на основе разработанной вероятностной модели символов, выполнены экспериментальные исследования для обоснования принятых решений.

Анализ литературных данных и постановка проблемы

Для решения задачи локализации ТО в научно-технической литературе предложено множество методов на основе корреляции, контурной и текстурной сегментации, нейронных сетей, преобразования Фурье, вейвлет-преобразования. В работе [1] приведен обзор методов, используемых при локализации текстовых областей на изображениях со сложным фоном. В работе [2] описывается автоматическая генерация признаков, которые будут использоваться для распознавания. Такие признаки предлагается

создавать с помощью машинного обучения, а в качестве базы для обучения предлагается использовать искусственные изображения, имеющие размер 8x8 пикселей, содержащие фрагменты текстовых символов. Для поиска ТО в реальных изображениях достаточно будет вычислить найденные признаки на нужных областях изображения.

В работе [3] предполагается, что буквы и слова на изображении, как правило, имеют постоянную толщину штриха. Поэтому для выявления таких объектов, по мнению авторов, перспективно использовать алгоритм SWT (stroke-width transformation). Толщину штриха можно использовать не только в качестве одного из признаков при классификации областей, но и в качестве признака при объединении областей в слова. Границы символов в рамках описанного подхода могут определяться, например, с помощью детектора границ Canny. Однако надо учитывать, что алгоритм SWT требует определенных дополнительных вычислительных затрат для борьбы с ошибками на углах и некоторыми другими специфическими для него эффектами.

В работе [4] используется понятие так называемого «контекста окружения» (environmental context). Авторы предлагают для обнаружения ТО анализировать фон, на котором находится ТО. Они предполагают, что вероятность наличия текста на не подходящем фоне (на травяном покрове, на небе и т.д.) – низкая. Следует отметить, что такое предположение не всегда оправдывается, что снижает качество локализации ТО.

В работе [5] предлагается использовать метод SURF для поиска особых точек изображения и создания их дескрипторов. Метод ищет особые точки с помощью матрицы Гессе. Детерминант матрицы Гессе достигает экстремума в точках максимального изменения градиента яркости. Предполагается, что если на изображении присутствуют буквы, то они будут плотно покрыты такими точками. Вдобавок к этому вычисляется визуальная заметность (visual saliency) в предположении, что символы текста всегда должны быть заметны на окружающем фоне. В результате формируется $(128 + 1)$ -мерный вектор признаков, совокупность которых и используется для обучения классификатора AdaBoost.

В работе [6] локализация ТО сводится к задаче поиска точечного текста, но для выявления точек, составляющих буквы, применяется хорошо известный алгоритм FAST. Затем проводится эвристическая фильтрация ложных кандидатов, объединения точек в буквы, букв – в слова, после чего применяется классификатор SVM [7] для детектирования текстовых областей.

В работе [8] локализация ТО основана на модели текстовой строки, однако не учитывает пространственные связи между строками, имеющиеся в изображениях символов, что отрицательно сказывается на результате локализации.

Проведенный анализ показал, что методы локализации, приведенные в большинстве из опубликованных работ, основаны при поиске текстов на анализе обособленных пикселей изображения без учета их пространственной взаимосвязи. Это не обеспечивает хорошие результаты локализации ТО на реальных изображениях, при наличии слитного текста (номера домов, рекламные щиты, вывески магазинов, автомобильные номера и т.п.). Очевидно, что для эффективной работы систем локализации метод локализации ТО на изображении должен учитывать особенности представления данных, а именно – строгую двумерную структуру изображения, в которой существует пространственная зависимость между значениями соседних пикселей.

Таким образом, можно сделать вывод, что наиболее перспективным подходом для решения задачи локализации текстовых областей на изображениях со сложным и неоднородным фоном является подход с учетом пространственного представления символов.

В работах авторов предложено для локализации ТО на изображении использовать в качестве классификатора сверточную нейронную сеть (СНС), которая по сравнению с классическими нейронными сетями обладает следующими преимуществами: возможностью учета пространственной структуры, снижением сложности архитектуры и обучения, устойчивостью к искажениям символов [9-13].

Обобщенная схема метода локализации текста на изображении

Разработана обобщенная схема метода локализации ТО (рис. 2) с учетом особенностей задачи на основе многоуровневого подхода и классификатора на основе СНС. Это позволило усовершенствовать метод ЛТО на изображениях со сложным фоном в направлении повышения достоверности обнаружения ТО.

Предлагаемый метод локализации ТО на изображении использует многомасштабное представление изображения на основе двумерного дискретного вейвлет-преобразования (вейвлет-пирамида), классификатор, построенный на сверточной СНС для отбора ТО-кандидатов [11-13]. Метод реализуется следующим алгоритмом:

Шаг 1. Из исходного цветного изображения (цветовое пространство RGB) выделяется прямоугольный фрагмент.

Шаг 2. Формируется пирамида изображений для выделенного фрагмента. Для этого осуществляется преобразование значений интенсивности фрагмента изображения к диапазону $[-1, 1]$ и выполняется дискретное вейвлет-преобразование с разложением изображения на два уровня по каждой составляющей каналов цвета R, G, B.

Шаг 3. Полученный набор пирамид изображений поступает на предварительно

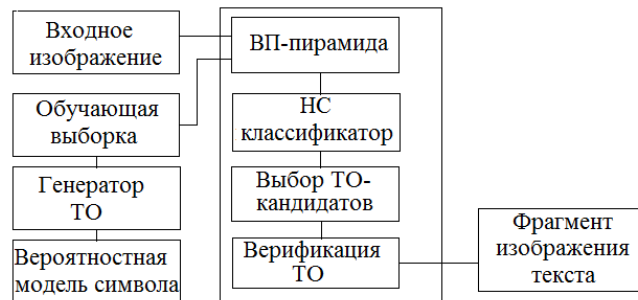


Рис. 2 Обобщенная схема реализации метода локализации ТО

обученную сверточную нейронную сеть, которая

классифицирует данный участок изображения как содержащий текст (текстовую область) или не содержащий текст — проводится выбор ТО-кандидатов.

Шаг 4. Повторяются шаги алгоритма 1 – 3 (для всего изображения).

Шаг 5. Отобранные области подвергаются верификации по форме и размерам. Формируется список координат текстовых областей, содержащих символы.

Одной из основных проблем, с которой сталкиваются при реализации предложенного метода, является то, что качество этапа классификации (в том числе и с использованием сверточных нейронных сетей) существенно зависит от объема и представительности обучающей выборки. В качестве таких выборок сегодня используются специально подобранные базы изображений, например, база данных конкурса алгоритмов искусственного интеллекта ICDAR 2011 Robust Reading Competition (<http://www.cvc.uab.es/icdar2011competition/>). Следует отметить, что количество и вид изображений, включенных в обучающую выборку, т. е. участвующих в обучении классификатора, определяют скорость обучения и достоверность классификации. При этом разработчики сталкиваются с противоречием: большой объем обучающей выборки, позволяющий учесть максимальное количество реальных вариантов ТО, обеспечивает улучшение работы классификатора, но значительно увеличивает время обучения, что часто является недопустимым. Кроме того подбор таких баз, как правило, определяется опытом разработчика, т. е. является субъективным.

Поэтому предложено в рамках проводимых исследований по разработке метода локализации текстовых областей такие ТО формировать автоматически с использованием генератора ТО и вероятностной модели символа.

Получение модели текстовой области

Будем считать, что изображение текстовой области представляет собой упорядоченный определенным образом набор изображений символов. При этом изображение символа можно рассматривать как двумерный массив пикселей, в котором определенным образом (в зависимости от шрифта, размеров и др.) расположены информативные и фоновые пиксели (ФП).

Формально изображение символа может быть представлено вероятностной моделью символа (ВМС) в виде кортежа

$$S = \{H, Wh, P\}, \quad (1)$$

где **H** – множество возможных значений высоты h текстового символа (в пикселях), **Wh** – множество возможных значений ширины wh текстового символа (в пикселях), **P** – множество матриц, соответствующих вероятности появления информационного пикселя в изображениях символа текста размерами $h \times wh$, $h \in H, wh \in Wh$.

Для получения ВМС проведен анализ изображений текстовых символов, включающих английские буквы ('a', 'b', ..., 'z') и десятичные цифры ('0', '1', ..., '9') различных типов (Arial, Times New Roman, Calibri, Cambria, Courier New, Adobe Garamond Pro, ...) и размеров (14, 15, ..., 20) шрифтов. Изображение символов – бинарное, информативный пиксель имеет белый цвет (значение пикселей равно 1), а фон – черный (значение пикселей равно 0).

Изображения всех символов разбиты на 4 группы в зависимости от заполнения информативными символами предложенной формы (рис. 3, а), учитывающей соотношения сторон символов.

В первой группе содержатся символы букв ('a', 'c', 'e', 'm', 'n', 'o', 'r', 's', 'u', 'v', 'w', 'x', 'z'), информативные пиксели (ИП) которых расположены в зоне 1 (рис. 3, а). Во второй – 'b', 'd', 'f', 'h', 'i', 'k', 'l', 't', '0', '1', '2', '3', '4', '5', '6', '7', '8', '9', ИП расположены одновременно в зонах 1 и 2 (рис. 3, б). В третьей – 'g', 'p', 'q', 'i', ИП занимают зоны 1 и 3 (рис. 3, в). В четвертой – 'j', ИП занимают все три зоны. Для каждой группы символов определены параметры модели (1): множество возможных значений высоты **H** и множество возможных значений ширины **Wh** текстового символа, множество матриц **P**, соответствующих вероятности появления информативного пикселя в изображениях символов. Такие матрицы составляют базу вероятностных моделей символов и могут быть использованы для создания изображений псевдосимволов при формировании обучающей выборки.

Для получения матрицы **p** из множества **P** для каждой группы символов определяется вероятность появления ИП в каждой точке изображения $p_{i,j} = n_{i,j} / N_{gr}$, где $n_{i,j}$ – количество ИП во всех изображениях символов группы, занимающих позицию с координатами (i,j) , N_{gr} – количество символов в группе. Пример матрицы **p** вероятности появления ИП пикселя в каждой точке изображения для символов первой группы размера 13x9 приведен в табл. 1.

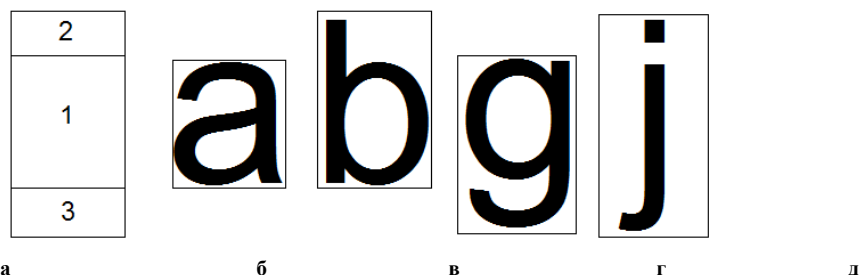


Рис. 3 Разбиение символов на группы: а – форма для разбиения изображения на группы, б,в,г,д – изображения различных символов

| | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.194 | 0.3 | 0.377 | 0.467 | 0.55 | 0.358 | 0.141 | 0.03 | 0 |
| 0.203 | 0.609 | 0.505 | 0.386 | 0.436 | 0.332 | 0.201 | 0.03 | 0 |
| 0.256 | 0.415 | 0.259 | 0.274 | 0.368 | 0.273 | 0.274 | 0.077 | 0 |
| 0.252 | 0.415 | 0.259 | 0.274 | 0.368 | 0.272 | 0.274 | 0.077 | 0 |
| 0.223 | 0.403 | 0.308 | 0.329 | 0.358 | 0.371 | 0.236 | 0.076 | 0 |
| 0.188 | 0.352 | 0.349 | 0.259 | 0.455 | 0.339 | 0.277 | 0.087 | 0.017 |
| 0.209 | 0.335 | 0.247 | 0.279 | 0.37 | 0.306 | 0.285 | 0.111 | 0 |
| 0.242 | 0.332 | 0.271 | 0.294 | 0.444 | 0.226 | 0.296 | 0.162 | 0.017 |
| 0.209 | 0.288 | 0.244 | 0.292 | 0.385 | 0.192 | 0.238 | 0.173 | 0.017 |
| 0.223 | 0.320 | 0.189 | 0.265 | 0.402 | 0.341 | 0.242 | 0.133 | 0 |
| 0.227 | 0.509 | 0.415 | 0.406 | 0.471 | 0.506 | 0.409 | 0.206 | 0.017 |
| 0.139 | 0.379 | 0.365 | 0.289 | 0.342 | 0.43 | 0.311 | 0.147 | 0 |
| 0.12 | 0.149 | 0.168 | 0.117 | 0.126 | 0.192 | 0.123 | 0.106 | 0.05 |

Использование ВМС для генерации изображения символа

Изображение псевдосимвола формируется в соответствии со следующим алгоритмом:

Шаг1. Проводится выбор номера группы ($i=1,...,4$) и инициализация размеров псевдосимвола (высоты h_i и ширины Wh_{ij}) в соответствии с имеющейся базой вероятностных моделей символов.

Шаг 2. Генерируется матрица случайных чисел TR_{rand} размера $h_i \times Wh_{ij}$, распределенных по равномерному закону в диапазоне $[0...p_{imax}]$, где p_{imax} – максимальное значение вероятности появления ИП в изображениях символов i -й группы.

Шаг3. Сравняется поэлементно матрица TR_{rand} с матрицей p из базы вероятностных моделей символов вероятности в соответствии со следующим выражением:

$$\text{Если } TR_{rand}(i,j) \geq p(i,j), \text{ то } TR(i,j)=ИП, \text{ иначе } TR(i,j)=ФП.$$

Бинарное изображение псевдосимвола получают в результате визуализации полученной матрицы.

В случае необходимости получения полутонового изображения, псевдосимвола, полученное бинарное изображение обрабатывают фильтром Гаусса. Если необходимо иметь цветное изображение псевдосимвола в цветовом пространстве RGB, то полутоновое изображение принимают за R-составляющую, а составляющие G и B получают поэлементным умножением элемента r_{ij} R-составляющей на величину $(1+\epsilon_{ij})$, где ϵ_{ij} – случайные числа, распределенные по равномерному закону в диапазоне $[-0.2...0.2]$.

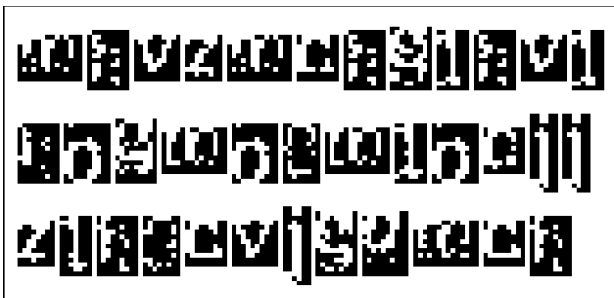


Рис. 4. Результат генерации ТО.

Изображение текстовой области может быть сгенерировано как упорядоченный определенным образом набор изображений символов. Пример сгенерированного изображения ТО представлен на рис. 4.

Для подтверждения инвариантности к масштабу ВМС (1) показано, что вероятность заполнения изображения символа информационным пикселем $p_{zs} = N_i / (h \cdot wh)$, где N_i – общее количество информационных пикселей в изображении символа, для каждой группы лежит в определенном диапазоне и не

зависит от размера шрифта:

- группа 1 $p_{zs} = [0.330...0.543]$; группа 2 $p_{zs} = [0.296...0.606]$;
- группа 3 $p_{zs} = [0.314...0.482]$; группа 4 $p_{zs} = [0.393...0.397]$.

Экспериментальные исследования

Предложенный метод локализации может работать в двух режимах – обучения и локализации. В первом режиме (в процессе обучения) с помощью модели ВМС и генератора ТО (рис. 2) создается обучающая выборка согласно приведенному выше алгоритму формирования символов. В ходе обучения изображения сформированных ТО последовательно подаются на вход вейвлет-преобразователя, где формируется набор вейвлет-пирамид, поступающих на НС-классификатор. Обучение классификатора происходит по методу «с учителем» и длится до получения желаемой точности классификации изображений обучающей выборки. Для обучения сети была создана обучающая выборка из 3730 изображений (36x64 пикселя), которые получены с помощью моделей символов текста. Также были добавлены изображения, содержащие только часть символов и без текста. Изображения имели сложный фон, т.е. содержали кроме ТО и другие объекты (фрагменты домов, деревьев и т.д.).

Для проверки качества обучения классификатора использовалась контрольная выборка, включающая как изображения ТО, сгенерированные с использованием модели ВМС, так и реальные изображения.

В режиме локализации текстовых областей на вход предварительно обученной системы локализации подаются реальные изображения (рис. 1) и НС-классификатор определяет ТО-кандидаты, которые затем верифицируются с использованием морфологических операций. В результате получают набор координат текстовых областей.

В ході комп'ютерного експеримента оцінювались точність локалізації фрагментів зображень із обучаючої і контрольної вибірок, які містили ТО. Для оцінки точності класифікації зображень, що містять ТО, використовувалась міра точності [14], яка показує частку областей, розпізнаних на зображенні як текстові, від дійсного числа ТО.

Після навчання мережі точність локалізації зображень для обучаючої вибірки склала 99,93 %, а для контрольної вибірки – 87,7 %. Приклад роботи запропонованого методу локалізації на реальному зображенні представлено на рис. 5.



Рис. 5. Приклад локалізації ТО на зображенні з використанням запропонованого методу: а – вихідне зображення; б – результат локалізації текстової області

Висновки

В роботі запропоновано метод локалізації тексту на зображенні з навчанням на основі ймовірнісної моделі символів.

Одною з особливостей запропонованого методу є створення навчаючої вибірки для навчання НС класифікатора, що усуває необхідність у отриманні вибірки реальних зображень з різними типами і розмірами шрифтів.

Для створення моделі ТО, здатної включати багато рядків тексту, і генерування на її основі ТО, виконано визначення статистичних характеристик зображень символів текстів з різними типами і розмірами шрифтів. Розроблена ймовірнісна модель символу дозволяє генерувати текстові області, що мають ті ж ймовірності розподілу інформативних пікселів, що і реальні символи різних розмірів і типів шрифтів.

З використанням сгенерованої моделі ТО зменшується залежність якості локалізації від характеристик ТО реальних зображень і підвищується гнучкість при створенні навчаючого набору, збільшується точність локалізації ТО після навчання сверточної нейронної мережі.

Проведено перевірку адекватності отриманої моделі, а також використано модель ТО для отримання навчаючого набору в двохетапній схемі локалізації ТО з допомогою сверточної нейронної мережі і багатомасштабного представлення зображення на основі дискретного вейвлет-перетворення. Після навчання мережі точність класифікації зображень навчаючої вибірки склала 99,93 %, а контрольної вибірки – 87,7 %.

В подальшому планується продовжити розробку моделі ТО і створити набір ТО для різних мов (українська, російська, в'єтнамська і т.д.).

Література

1. Андрианов, А. Локализация текста на изображениях сложных графических сцен /А.И. Андрианов // Современные проблемы науки и образования. – 2013. – № 3; URL: www.science-education.ru/109-9311 (дата звернення: 10.10.2014).
2. Coates, A. Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning [Text] / A. Coates, B. Carpenter, C. Case, S. Sathesh, B. Suresh, T. Wang, D. Wu, A. Ng. // The 11th International Conference on Document Analysis and Recognition (ICDAR).– 2011.– P. 440 – 445.
3. Epshtein, B. Detecting Text in Natural Scenes with Stroke Width Transform [Text] / B. Epshtein, E. Ofek, Y. Wexler // 23rd IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.V. *San Francisco*, 2010.
4. Kunishige, Y. Scenery Character Detection with Environmental Context [Text] /Y. Kunishige, F. Yaokai, S. Uchida // The 11th International Conference on Document Analysis and Recognition (ICDAR). – 2011.– P. 1049 – 1053.
5. Uchida, S. A Keypoint-Based Approach Toward Scenery Character Detection [Text] / S. Uchida, Y.

Shigeyoshi, Y. Kunishige, F. Yaokai. // The 11th International Conference on Document Analysis and Recognition (ICDAR).– 2011.– P. 819–823.

6. Du, Y. Dot Text Detection Based on FAST Points [Text] / Y. Du, H. Ai, S. Lao // The 11th International Conference on Document Analysis and Recognition (ICDAR).– 2011.– P. 435–439.

7. Jung, C. Accurate text localization in images based on SVM output scores [Text] / C. Jung, Q.F. Liu, J. Kim // Image and Vision Computing. – 2009.–vol.27.–P. 1295–1301.

8. Адаптивна локалізація символічних написів на зображеннях методом вейвлет-аналізу / С.Г. Антошук, А.О. Ніколенко, О.Ю. Бабілунга, О.В. Ткаченко // Вісник Житомирського технол. ун-ту. – 2008. – № 4(47). – С. 124-130

9. Николенко, А. Обнаружение текстовых областей в видео-последовательностях [Текст] / А.А. Николенко, Тьен Т.К. Нгуен // Искусственный интеллект. — 2012. — № 4. — С. 227 — 234.

10. Delakis, M. Text detection with convolutional neural networks [Text] / M. Delakis, Cr. Garcia // International Conference on Computer Vision Theory and Applications. – 2008. – P. 290-294.

11. Ніколенко, А. Метод виявлення текстових областей в відеопослідовності з використанням вейвлет-перетворення [Текст] / А.О. Ніколенко, О. Ю. Бабілунга, Тьен Т.К. Нгуен – XI Всеукраїнська міжнародна конференція “Оброблення сигналів і зображень та розпізнавання образів” (УкрОБРАЗ’2012). – Збірник наукових праць. – Київ, МННЦІТС, 2012.– С. 99 – 102.

12. Николенко, А. Локализация текстовых областей на изображениях с использованием сверточной нейронной сети [Текст] / А.А. Николенко, О. Ю. Бабилунга, Тьен Т.К. Нгуен // Вісник Національного технічного університету “Харківський політехнічний інститут”. Збірник наукових праць. Серія: Інформатика та моделювання. – Харків: НТУ “ХПІ”. – 2013. – № 19 (992). – С. 121–127.

13. Нгуен, Т. Обнаружение и распознавание текстов на изображениях сложных графических сцен с помощью сверточной нейронной сети [Текст] / Тьен Т.К. Нгуен // Електротехн. та комп’ют. системи. – 2014. – № 13 (89). – С. 125–130.

14. Вежневцев, В. Оценка качества работы классификаторов [Текст] / В. Вежневцев // Компьютерная графика и мультимедиа. — Выпуск №4. — 2006. — С. 106.

References

1. Andrianov, A. I. (2013). Lokalizatsiya teksta na izobrazheniyah slozhnykh graficheskikh stsen. Sovremennyye problemy nauki i obrazovaniya, URL: www.science-education.ru/109-9311.

2. Coates, A., Carpenter, B., Case, C., Satheesh, S., Suresh, B., Wang, T., Wu, D., Ng, A. (2011). Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning. The 11th International Conference on Document Analysis and Recognition (ICDAR), P. 440–445.

3. Epshtein, B., Ofek, E., Wexler, Y. (2010). Detecting Text in Natural Scenes with Stroke Width Transform. 23rd IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.V.

4. Kunishige, Y., Yaokai, F., Uchida, S. (2011). Scenery Character Detection with Environmental Context. The 11th International Conference on Document Analysis and Recognition (ICDAR), P. 1049–1053.

5. Uchida, S., Shigeyoshi, Y., Kunishige, Y., Yaokai, F. (2011). A Keypoint-Based Approach Toward Scenery Character Detection. The 11th International Conference on Document Analysis and Recognition (ICDAR), P. 819–823.

6. Du, Y., Ai, H., Lao, S. (2011) Dot Text Detection Based on FAST Points. The 11th International Conference on Document Analysis and Recognition (ICDAR), P. 435–439.

7. Jung, C., Liu, Q.F., Kim, J. (2009). Accurate text localization in images based on SVM output scores. Image and Vision Computing, vol. 27, P. 1295–1301.

8. Antoschuk, S.G., Nikolenko, A.O., Babluniga, O.Yu., Tkachenko, O.V. (2008). Adaptivna lokalizatsiya simvolnih napisiv na zobrazheniyah metodom veyvlet-analizu. Visnik Zhitomirskogo tehnol. un-tu, 4(47), S. 124-130

9. Nikolenko, A.A., Nguyen, Tien T.K. (2012). Obnaruzhenie tekstovyykh oblastey v video-posledovatel'nostyakh. Iskusstvennyy intellekt, 4, S. 227 — 234.

10. Delakis, M., Garcia, Cr. (2008). Text detection with convolutional neural networks. International Conference on Computer Vision Theory and Applications, P. 290-294.

11. Nikolenko, A.A., Nguyen, Tien T.K. (2012). Obnaruzhenie tekstovyykh oblastey v video-posledovatel'nostyakh. Iskusstvennyy intellekt, 4, S. 227 — 234.

12. Nikolenko, A. A., Babiluniga, O. Yu., Nguyen, Tien T.K. (2013). Lokalizatsiya tekstovyykh oblastey na izobrazheniyah s ispolzovaniem svertochnoy neyronnoy seti. Visnik Natsionalnogo tehnichnogo universitetu “Harkivskiy politehnichnyy institut”. Zbirnik naukovih prats. Seriya: Informatika ta modelyuvannya, Harkiv: NTU “HPI”, 19 (992), S. 121–127.

13. Nguyen, Tien T.K. Obnaruzhenie i raspoznavanie tekstov na izobrazheniyah slozhnykh graficheskikh stsen s pomoschyu svertochnoy neyronnoy seti (2014). Elektrotehn. ta komp'yut. sistemi, 13 (89), S. 125–130.

14. Vezhnevets, V. Otsenka kachestva raboty klassifikatorov (2006). Kompyuternaya grafika i multimedia, 4, S. 106.

Рецензія/Peer review : 3.10.2014 р.

Надрукована/Printed :5.11.2014 р.