

Колосова Владислава Вадимовна,
обучающаяся кафедры искусственного
интеллекта и системного анализа
Донецкого национального технического
университета, г. Донецк
E-mail: vladi.kolosova@gmail.com

Научный руководитель
Бычкова Елена Викторовна,
ст. преподаватель
Донецкого национального технического
университета, г. Донецк
E-mail: elena.bichkova@gmail.com

ВЫБОР МЕТОДОВ РАСПОЗНАВАНИЯ АЛФАВИТНО–ЦИФРОВОЙ ИНФОРМАЦИИ НА ФОТОИЗОБРАЖЕНИИ

ИНФОРМАТИЗАЦИЯ И ЖИЗНЬ

УДК 004.93

Суть нашего научного исследования состоит в разработке алгоритмов и всевозможных программных средств, которые позволят увеличить эффективность детектирования символов на фоне, подверженном шумовым, аффинным и проекционным искажениям. В данной статье предлагается математический подход к решению задачи распознавания символов, включающий такие основные этапы, как детектирование области расположения символов на изображении, стандартизация отображения изображения и выделение отдельных символов, а также распознавание их.

The work deals with the logic design and software environment which enable to increase the characters recognition efficiency on a complex background with noise, affine and projection distortions. The paper proposes mathematical approach to solve the character recognition problem, which includes three main stages: the region character locations detection in the image; image display standardization and individual characters selection; character recognition.

Ключевые слова

Обработка изображений, алгоритм Виолы-Джонса, преобразование Хафа, искусственные нейронные сети, распознавание символов.

Keyword

Image processing, Viola-Jones algorithm, Hough transformation, artificial neural networks, character recognition.

Задача обработки и анализа изображений является относительно новой и получила свое развитие во второй половине XX века, параллельно с развитием компьютерных технологий [1, 2]. Среди особо важных и интересных задач в данной предметной области можно выделить распознавание объектов на изображениях со сложным (неоднородным по цвету и фактуре) фоном, что имеет большую зна-

чимость в таких областях, как государственная оборона, обеспечение правопорядка, предотвращение несанкционированного доступа к секретным объектам и объектам военного назначения [2].

За последние годы был представлен ряд методов и алгоритмов, применяемых на различных стадиях процесса распознавания символов на изображениях. Однако их применение недостаточно эффективно в реальных условиях, которые характеризуются наличием сложного фона на изображениях, различной степенью освещенности, наличием шумов, аффинными и проекционными искажениями объектов, возникающими из-за изменения углов регистрации. Эффективность работы большинства алгоритмов становится ниже технологически приемлемого уровня при наличии искажений подобного рода. Все это требует применения различных алгоритмов предварительной обработки, что в свою очередь усложняет процесс распознавания, делает его более громоздким, увеличивает объем и время вычислительного процесса. Таким образом, данная задача до сих пор не решена в полном объеме и является актуальной на сегодняшний день.

Одним из наиболее актуальных и интенсивно развивающихся направлений данной области является распознавание автомобильных номерных знаков (государственных регистрационных знаков транспортных средств). Поэтому было принято решение осуществлять реализацию и тестирование алгоритмов на примере распознавания символов на автомобильных номерных знаках при наличии сложного фона на изображениях.

Процесс распознавания символов на изображениях со сложным фоном можно разделить на три этапа [3]:

1. Детектирование области расположения символов на изображении.
2. Нормализация изображения и выделение отдельных символов.
3. Распознавание символов.

Автор статьи предлагает разработку алгоритмов и программных средств, позволяющих повысить эффективность распознавания символов на сложном фоне.

АЛГОРИТМ ДЕТЕКТИРОВАНИЯ ОБЛАСТИ РАСПОЛОЖЕНИЯ СИМВОЛОВ НА ИЗОБРАЖЕНИИ

Общий смысл процесса детектирования объектов на изображениях заключается в отнесении входных изображений к одному из двух классов: имеющие и не имеющие в обрабатываемой области объекты, соответствующие заданным критериям. Сегодня существует множество алгоритмов, применяемых для решения этой задачи. Среди наиболее популярных и эффективных алгоритмов можно выделить гистограммный анализ цветовых компонент изображения [4], алгоритмы с использованием искусственных нейронных сетей [5], метод главных компонент [6], дескрипторы локальных особенностей изображения (SURF [7], SIFT [8], HOG [9], FAST [10]), алгоритм Виолы-Джонса [11].

Одним из наиболее эффективных подходов является алгоритм Виолы-Джонса, предложенный в 2001 г. ис-

следователями P. Viola и M. Jones [11]. Для классификации объектов на изображении в данном алгоритме используются признаки Хаара, которые представляют собой прямоугольные области, состоящие из нескольких смежных частей (рис. 1).

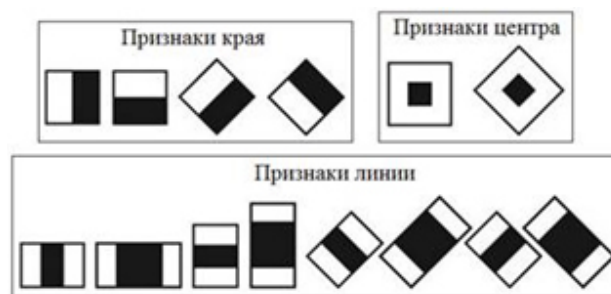


Рис. 1. Признаки Хаара

Данный алгоритм имеет каскадную структуру, состоящую из классификаторов. Каждый уровень этого каскада представляет собой классификатор, называемый ансамблем, который состоит из нескольких признаков Хаара. Выполняется последовательное сканирование изображения окном заданного размера, и анализируемая область изображения проверяется каждым из классификаторов на соответствие искомому образцу. Если на каком-либо уровне определяется, что исследуемая область не содержит признаки искомого объекта, то процедура заканчивается и данная область помечается как ложная. Таким образом, каскадный классификатор отбрасывает из рассмотрения ложные области изображения. В итоге остаются только области, имеющие наибольшую вероятность содержания искомого объекта [12].

Ансамбли первых уровней состоят из малого количества классификаторов. С увеличением порядкового номера уровня увеличивается количество классификаторов, содержащихся в нем. Ансамбли последних уровней могут содержать десятки и сотни классификаторов для более тщательного анализа областей изображения.

Значение признака для исследуемой области изображения вычисляется путем суммирования интенсивностей пикселей в черных и белых областях признаков Хаара, затем вычисляется разность между суммами и сравнивается с заданным порогом:

где L – сумма значений пикселей, соответствующих белым областям; D – сумма значений пикселей, соответствующих черным областям.

Для сокращения числа операций при вычислении признаков в данном алгоритме используется интегральное представление изображения [13]. Данный подход предполагает представление исходного изображения в виде матрицы, в которой значение каждой ячейки является суммой интенсивностей всех пикселей, находящихся выше и левее данной ячейки:

$$I[x,y] = \sum_{i=0}^x \sum_{j=0}^y lm[i, j]$$

где $lm[i, j]$ – яркость пикселя изображения, x, y – порядковые номера пикселя.

Таким образом, матрица $I[x,y]$ состоит из элементов, каждый из которых включает сумму значений пикселей прямоугольной области от $(0, 0)$ до (x, y) . Расчет интегральной матрицы выполняется за один проход по всем пикселям изображения:

$$I(x,y) = lm(x, y) - I(y - 1, x - 1) + I(y - 1, x) + I(y, x - 1)$$

Обучение алгоритма Виолы-Джонса осуществляется на основе бустинга (англ. adaptive boosting – адаптивное улучшение), где классификаторы, представляющие собой всевозможные признаки Хаара, комбинируются таким образом, чтобы получить более сложный классификатор. Из всего множества признаков выбираются те, которые наиболее оптимально подходят для обнаружения искомого объекта [13].

Для обучения алгоритма была сформирована база данных из 5 000 изображений сегментированных автомобильных номерных знаков (размер 28×52 пикселей), соответствующих ГОСТ Р 50577 – 93, группе 1, типу 1 (рис. 2 (а)). Основная часть изображений была получена из источников [14], [15]. Остальная часть изображений была сформирована самостоятельно.

Изображения автомобильных номерных знаков представлены при различных углах отклонения относительно регистрирующего устройства. Для обеспечения высокой степени инвариантности алгоритма к различным искажениям, к некоторым изображениям из обучающего набора применялось изменение освещения, контраста и размытие.



Рис. 2. Изображения из обучающей выборки:
а) изображения сегментированных автомобильных номерных знаков;
б) фоновые изображения

Также для обучения была сформирована база из 6 000 фоновых изображений, не содержащих образы автомобильных номерных знаков (рис. 2 (б)). Для обеспечения эффективного процесса обучения данная база содержит множество ложных образов, имеющих визуальные признаки, схожие с автомобильными номерными знаками.

Обучение классификатора осуществлялось с помощью модуля обучения алгоритма Виолы-Джонса в библиотеке OpenCV (opencv_traincascade.exe). Время обучения алгоритма составило 32,4 часа.

Тестирование классификатора производилось на 2 000 изображений транспортных средств со сложной фоновой структурой. Точность классификации при тестировании составила 98,21 %. После того как искомым объект на изображении найден, данная область изображения сегментируется для последующей обработки. Одна-

ко сложность заключается в том, что автомобильный номерный знак может быть расположен под неизвестным углом на изображении, что может внести определенные трудности в процесс распознавания символов на данном номерном знаке (рис. 3). В связи с этим требуется применять алгоритмы нормализации изображения.



Рис. 3. Выделение автомобильного номерного знака

АЛГОРИТМ НОРМАЛИЗАЦИИ ИЗОБРАЖЕНИЯ И ВЫДЕЛЕНИЕ ОТДЕЛЬНЫХ СИМВОЛОВ

С целью минимизации влияния на результат распознавания шумов и фона предлагается применять выделение контуров на изображении (англ. Edge detection) оператором Собеля [16] (рис. 4 (б)).

При анализе различных алгоритмов было принято решение определять угол отклонения автомобильного номерного знака на основе горизонтальных граней пластины номерного знака. Данный подход предполагает, что на полученном сегментированном изображении горизонтальные грани пластины номерного знака представляют собой линии максимальной длины. Поэтому все найденные линии, длина которых превышает заданный порог, принимаются за горизонтальные грани пластины номерного знака. Поиск прямых линий на изображении выполняется с помощью преобразования Хафа (англ. Hough transform) [16, 17] (рис. 4 (в)).

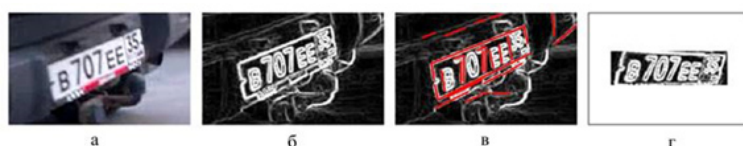


Рис. 4. Нормализация изображения:
а) исходное сегментированное изображение;
б) результат выделения контуров; в) результат преобразования Хафа; г) выравнивание изображения

Из всех найденных линий выбираются линии с максимальной длиной относительно ширины изображения. Данные линии предположительно будут являться горизонтальными гранями пластины номерного знака. После нахождения горизонтальных граней пластины номерного знака изображение выравнивается путем поворота в горизонтальную плоскость и выполняется обрезка лишних областей (рис. 4 (г)).

На следующем этапе необходимо выделить области

расположения символов на изображении для их последующего распознавания. Для этого выполняется поиск замкнутых контуров. Найденные замкнутые контуры на изображении выделяются прямоугольными областями для последующей сегментации [16] (рис. 5).

Однако на изображении могут присутствовать ложные объекты, представляющие собой замкнутые контуры. Для этого выполняется проверка обнаруженных областей на соответствие следующему условию: $h > h_{max}$, где h – высота области расположения символа, h_{max} – высота изображения. При выполнении данного условия рассматриваемая область считается ложной.



Рис. 5. Выделение областей расположения символов:

а) выделение замкнутых контуров на изображении; б) исключение ложных областей

Таким образом находятся области прямоугольной формы, которые соответствуют границам символов, что делает возможным дальнейшее их распознавание. Полученные области изображения масштабируются к размеру 29x29 пикселей и подаются на вход алгоритму распознавания. Данный размер масштабирования соответствует размеру изображений символов из обучающей выборки, а также размеру входного слоя сверточной нейронной сети.

АЛГОРИТМ РАСПОЗНАВАНИЯ СИМВОЛОВ

После этапа выделения отдельных символов следует последний – распознавание символов. Основные сложности здесь заключаются в том, что символы на изображении могут быть подвержены шумовым, аффинным и проекционным искажениям.

При анализе различных алгоритмов, применяемых для решения данной задачи, было принято решение использовать сверточные нейронные сети (СНС) [18, 19], так как они обладают следующими преимуществами:

1. Структура СНС хорошо подходит для обработки двумерных данных.
2. Наслаивающиеся друг на друга локальные рецептивные поля нейронов (области входных сигналов нейронов) обеспечивают взаимосвязь пространственно-зависимых областей изображения.
3. Повышенная устойчивость к аффинным и проекционным искажениям входных данных, шумам, изменению масштаба [19].

Исходя из перечисленных преимуществ для распознавания символов на автомобильных номерных знаках была разработана СНС, представленная на рис. 6.

Разработанная СНС состоит из двух типов слоев – сверточные и подвыборочные. Каждый слой состоит из набора карт признаков, которые, в свою очередь, состоят из нейронов.

При выполнении работы получены следующие

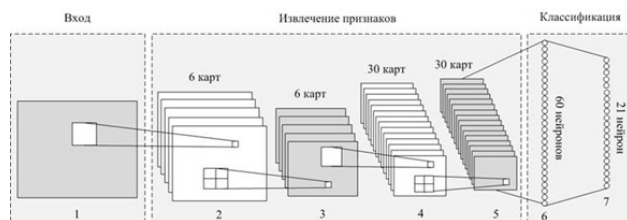


Рис. 6. Структура СНС.

$C1, C2$ – сверточные слои; $P1, P2$ – подвыборочные слои; $N1, N2$ – слои из обычных нейронов

научные и практические результаты:

1. Предложена реализация алгоритма Виолы-Джонса для решения задачи детектирования области расположения символов на изображениях со сложным фоном. Алгоритм обеспечивает точность классификации на уровне 98,21 %, скорость работы на уровне 34 мс и высокую устойчивость к шумовым, аффинным и проекционным искажениям входных данных.

2. Разработан алгоритм нормализации изображений символов, основанный на преобразовании Хафа и выделении замкнутых контуров на изображении. Алгоритм обеспечивает обнаружение точных границ пластины номерного знака, расположенного под различными углами наклона.

3. Разработан алгоритм распознавания символов на изображении на основе сверточной нейронной сети. Алгоритм обеспечивает точность распознавания символов на изображениях на уровне 98 %, скорость работы – на уровне 33 мс и высокую устойчивость к шумовым, аффинным и проекционным искажениям входных данных.

4. На основе разработанных алгоритмов была создана программная система, которая обеспечивает точность распознавания автомобильных номерных знаков, соответствующих ГОСТ Р 50577 – 93, группе 1, типу 1 на уровне 97,8 %. Программная реализация алгоритмов является инициативой авторов, выполнялась с целью написания научных трудов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений. М.: Высшая школа. – 1983. – С. 295.
2. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. – М.: Наука. – 1974. – С. 416.
3. Елизаров А.И., Афонасенко А.В. Методика построения систем распознавания автомобильного номера. Известия томского политехнического университета. – 2006. – Т. 309. – № 8. – С. 118 – 122.
4. Друки А.А., Спицын В.Г. Применение сверточных нейронных сетей для выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном. Известия Томского политехнического университета. – 2014. – Т. 324. – № 5. – С. 85 – 91.
5. Massa F., Aubry M., Marlet R. Convolutional Neural Networks for joint object detection and pose estimation: A comparative study. Computer Vision and Pattern Recognition (CVPR). – 2015. – Vol. 4. – Pp. 231 – 239.