

УДК 004.942

ПРОЕКТИРОВАНИЕ ПРОГРАММНОГО КОМПЛЕКСА ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ДОРОЖНЫХ ЗНАКОВ В ПОТОКОВОМ ВИДЕО

Смолянинов В.А., Белов Ю.С.

*Московский государственный технический университет имени Н.Э. Баумана, филиал,
Калуга, e-mail: fn1-kf@mail.ru*

В данной статье описывается проектирование системы поиска и распознавания дорожных знаков в потоковом видео на основе библиотеки OpenCV и сверточной нейронной сети. Представленная модель использует последовательность видеокadres в качестве входных данных. Сначала выполняется перевод видеокadra в цветовое пространство HSV. Затем к HSV-изображению применяется фильтр размытия с ядром 5x5. Получившееся размытое изображение бинаризуется по порогу. Морфологические операции, проводимые над двоичным изображением, удаляют шумы. Следующим этапом проводится контурный анализ, в результате которого отбирается самый большой по площади. Затем из исходного изображения извлекается область с потенциальным знаком. На вход классификатору подаются видеокadres, полученные с помощью алгоритма поиска дорожных знаков. В результате классификации пользователь получает конечную информацию о детектированном дорожном знаке. Основное внимание уделяется тестированию производительности модели. Экспериментально было оценено качество работы алгоритмов поиска и распознавания дорожных знаков на примере видеопотока реального времени, сгенерированного при нормальном дневном освещении. Эксперимент показал, что точность детекции составляет 87,42%, точность распознавания – 97,64%. Общий процент работы системы составляет 80,52%.

Ключевые слова: поиск дорожных знаков, распознавание дорожных знаков, OpenCV, сверточная нейронная сеть

DESIGNING A SOFTWARE COMPLEX FOR TRAFFIC SIGN DETECTION AND RECOGNITION IN THE STREAMING VIDEO

Smolyaninov V.A., Belov Yu.S.

Bauman Moscow State Technical University, Kaluga branch, Kaluga, e-mail: fn1-kf@mail.ru

This article describes the design of a system for searching and recognizing traffic signs in streaming video based on the OpenCV library and a convolutional neural network. The presented model uses a sequence of video frames as input. First, the video frame is translated into the HSV color space. Then a 5x5 kernel blur filter is applied to the HSV image. The resulting blurred image is threshold binarized. Morphological operations performed on a binary image remove noise. The next step is contour analysis, as a result of which the largest area is selected. Then the area with a potential sign is extracted from the original image. The input to the classifier is video frames obtained using the traffic sign search algorithm. As a result of the classification, the user receives the final information about the detected road sign. The focus is on testing the performance of the model. The quality of the algorithms for the search and recognition of road signs was experimentally evaluated using the example of a real-time video stream generated under normal daylight illumination. The experiment showed that the detection accuracy is 87.42%, the recognition accuracy is 97.64%. The overall percentage of the system is 80.52%.

Keywords: traffic sign detection, traffic sign recognition, OpenCV, convolutional neural network

В настоящее время системы поиска и распознавания дорожных знаков привлекают к себе большое внимание. Цель таких систем – осведомить водителей о приближающихся знаках на дороге и предупредить их о возможных опасностях. Дорожные знаки включают полезную и важную визуальную информацию, такую как ограничение максимальной скорости, сужение полосы движения, направление движения, расстояние до места назначения, а также опасные или необычные условия, чтобы помочь водителю во время движения. Если водитель не замечает дорожный знак или не понимает содержащуюся в нем информацию, то разумно предположить, что условия движения становятся неоптимальными. Данная система может уменьшить эту возможность и значительно помочь водителю,

обнаруживая и распознавая попутные дорожные знаки.

Цель исследования: разработать комплекс обнаружения и распознавания дорожных знаков с помощью методов, минимизирующих выявленные недостатки существующих систем.

Описание модели

Изучив основные аспекты в области распознавания дорожных знаков, можно сделать вывод о том, что целесообразно объединять несколько алгоритмов для решения задачи распознавания [1]. Например, объединение алгоритмов цветовой сегментации с основным функционалом библиотеки компьютерного зрения OpenCV позволит максимизировать процент корректного распознавания и свести к минимуму недо-

статки алгоритмов. Работа системы обнаружения и распознавания дорожных знаков происходит в два этапа: детектирование дорожных знаков; распознавание дорожных знаков. Общая модель системы распознавания дорожных знаков состоит из ряда взаимосвязанных модулей.

Веб-камера. В этом исследовании для сбора входных данных использовалась веб-камера Defender G-lens 2597 (рис. 1). Благодаря CMOS сенсору с разрешением 2 мегапикселя камера выдает детализированное широкоформатное изображение. Автоматическая подстройка светочувствительности позволяет получать высококачественное изображение даже при недостаточном освещении. Нет необходимости в постоянной настройке фокуса при перемещении в кадре: камера делает это автоматически.

Цветовая сегментация. Сегментация цвета является важным шагом, на котором с изображения (видеокадра) устраняются все фоновые объекты и неважная информация, и генерируется двоичное изображение, содержащее дорожные знаки и любые другие схожие объекты [2]. Результат цветовой сегментации показан на рис. 1.



Рис. 1. Результат цветовой сегментации

Детектирование дорожного знака реализуется посредством контурного анализа. На предобработанном двоичном изображении выполняется поиск контуров потенциального дорожного знака с помощью функции `cv2.findContours()`. Затем найденные контуры отображаются на изображении (`cv2.drawContours()`). Далее происходит поиск наибольшего по площади контура

и поиск области, в которой он располагается (`cv2.boundingRect()`) [3]. Завершается этап детекции вырезанием области с потенциальным дорожным знаком (рис. 2).



Рис. 2. Детектированный дорожный знак

Набор обучающих данных. Реализация эксперимента проходила с использованием набора данных, сгенерированного на основе российской базы дорожных знаков (RTSD). Для генерации базы данных RTSD был привлечен персонал из организации «Геоцентр-Консалтинг» [4]. Набор обучаемых данных для предлагаемой модели содержит 42689 изображений с 53 различными классами. Изображения неравномерно распределены между этими классами, поэтому модель может предсказывать одни классы с большей точностью, чем другие. Исходя из этого, рекомендуется сделать набор обучающих данных более согласованным, увеличив набор данных с помощью изменения существующих изображений. Такой подход называется искусственным увеличением данных. Изображения, содержащиеся в исходном наборе данных RTSD-3, имеют динамический диапазон измерений: от $16 \times 16 \times 3$ до $128 \times 128 \times 3$, поэтому его нельзя передать непосредственно в модель. Используя библиотеку компьютерного зрения OpenCV, была установлена размерность (40×40) всех изображений в наборе данных с помощью метода `cv2.resize (img, (40, 40))` (рис. 3).

id класса	Изображение	id класса	Изображение	id класса	Изображение
0	20200317-134001-319673 Файл "PNG" 3,38 KB	18	20200317-134645-035317 Файл "PNG" 3,88 KB	35	20200317-140254-953541 Файл "PNG" 3,16 KB
1	20200317-134115-033598 Файл "PNG" 3,25 KB	19	20200317-134657-206243 Файл "PNG" 3,35 KB	36	20200317-140306-206994 Файл "PNG" 3,62 KB
2	20200317-134128-311343 Файл "PNG" 3,81 KB	20	20200317-134712-881091 Файл "PNG" 3,55 KB	37	20200317-140327-379651 Файл "PNG" 3,43 KB
3	20200317-134140-460939 Файл "PNG" 3,69 KB	21	20200317-134724-231731 Файл "PNG" 3,41 KB	38	20200317-140340-597806 Файл "PNG" 3,42 KB
4	20200317-134158-850641 Файл "PNG" 4,18 KB	22	20200317-134736-772763 Файл "PNG" 3,61 KB	39	20200317-140352-065029 Файл "PNG" 3,53 KB
5	20200317-134216-723949 Файл "PNG" 4,27 KB	23	20200317-134750-627012 Файл "PNG" 3,86 KB	40	20200317-140404-445647 Файл "PNG" 3,20 KB

Рис. 3. Фрагмент набора изображений 40×40 для каждого из 53 классов

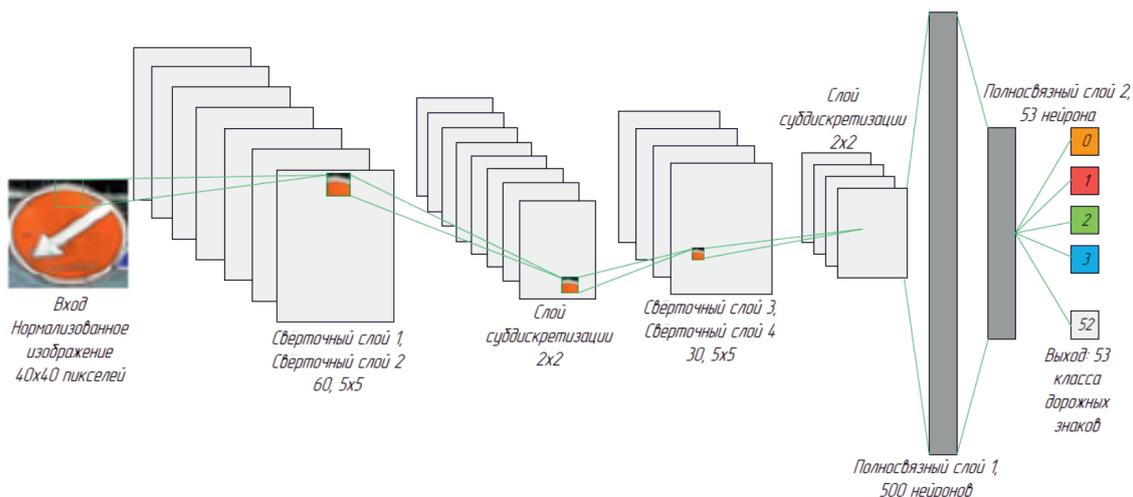


Рис. 4. Архитектура сверточной нейронной сети

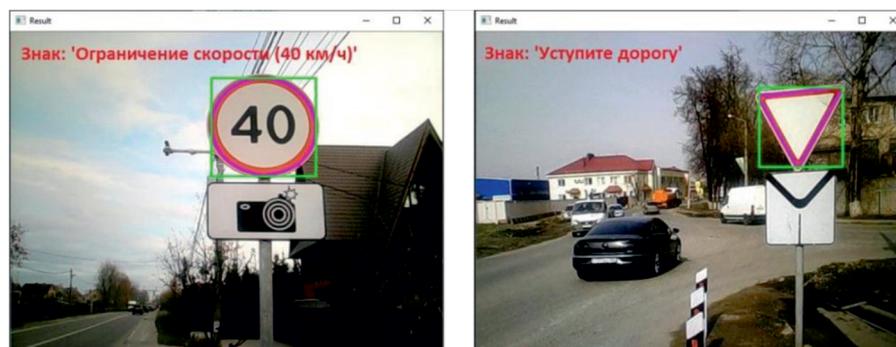


Рис. 5. Пример работы системы распознавания дорожных знаков

Сверточная нейронная сеть. Сверточная нейронная сеть, изображенная на рис. 4, спроектирована на основе модели распознавания рукописных цифр [5] и состоит из четырех сверточных слоев и двух полносвязных. Первые два слоя свертки содержат 60 фильтров, каждый из которых имеет размерность 5×5 . Последние два слоя содержат 30 фильтров с размером ядра 3×3 . Функция выпрямленных линейных единиц (ReLU) используется в качестве функции активации для всех сверточных слоев. После каждого из двух слоев свертки используются слои Maxpooling, размер пула которых составляет 2×2 . Чтобы уменьшить вероятность переобучения, после второго слоя maxpooling и первого слоя Dense добавляется 50% ный Dropout (отсев). Первый полносвязный слой содержит 500 нейронов, второй – 53 нейрона, что равняется количеству классов дорожных знаков. Последняя

функция активации – это функция softmax для классификации.

Классификация. Этап классификации (расознавания) выполняется с помощью сверточной нейронной сети. На вход классификатору подаются как тестовая выборка изображений дорожных знаков, так и изображения, полученные с помощью алгоритма поиска дорожных знаков. В результате классификации пользователь получает конечную информацию о детектированном дорожном знаке (рис. 5).

Анализ полученных результатов. За 30 эпох обучения нейронной сети точность на обучающих данных составила 0,9623, ошибка на обучающих данных уменьшилась до 0,1304; точность на проверочных данных достигла результата 0,9918, а ошибка на проверочных данных снизилась до 0,0319. С помощью библиотеки Matplotlib, для наглядности, были построены графики точности и ошибки обучения (рис. 6).

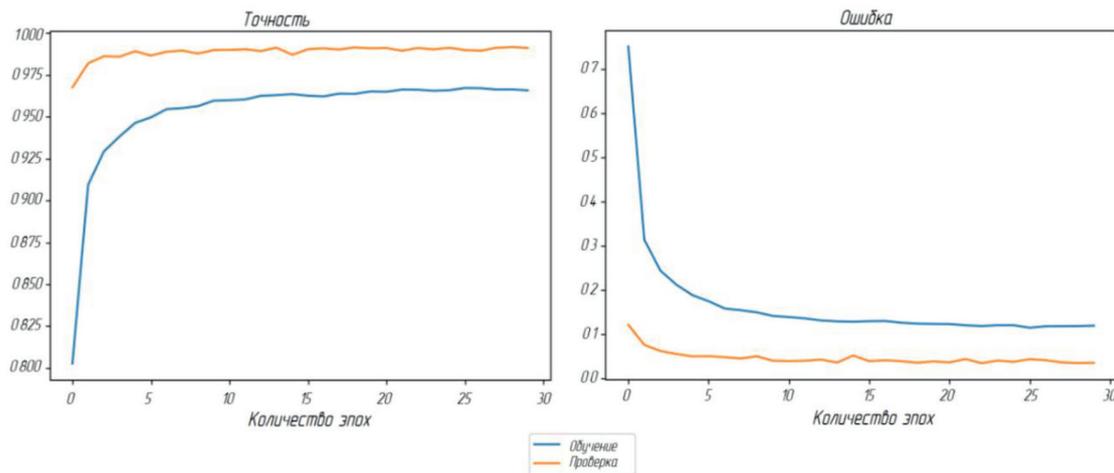


Рис. 6. Показатели точности и ошибки на протяжении 30 эпох обучения

Таблица 1

Результаты работы алгоритма обнаружения в потоковом видео

Форма контура знака	Общее количество знаков на видео	Количество корректно обнаруженных знаков	Верное обнаружение (%)
Треугольник	276	220	79,71
Четырехугольник	180	148	82,22
Окружность	160	140	87,5
Всего	616	508	82,47

Реализация модели обеспечивает точность в 99,22%, тем самым превышая результат из исследования [6], в котором сверточная нейронная сеть, обученная на основе базы немецких дорожных знаков (GTSRB), достигает точности 95%.

Тестирование модели. Вследствие своих недостатков предлагаемый алгоритм поиска дорожных знаков неустойчив к некоторому роду шумов: частичное перекрытие знаков, низкое качество видео ввиду малого разрешения камеры, а также плохой ее стабилизации и недостаточной цветопередачи. В режиме реального времени алгоритм корректно детектирует 82,47% дорожных знаков.

Согласно табл. 2 средний показатель качества распознавания знаков на изображениях, полученных с видеопотока, составляет 97,64%, что является доверительным показателем. Результаты распознавания знаков ограничения скорости были объединены ввиду отсутствия всех четырех наименований этого класса в видеопотоке.

Стоит обратить внимание на то, что процент классификации изображений, извлеченных из видеопотока, немного ниже, чем процент классификации изображений, сгенерированных при тестировании.

Это расхождение может быть оправдано тем, что изображения, содержащиеся в обучающем наборе, имеют то же происхождение, что и тестовые синтезированные изображения, в то время как реальные изображения, в свою очередь, имеют другое отношение к изображениям из обучающего набора данных [7]. Согласно результатам оценки качества работы разработанной системы поиска и распознавания объектов дорожной инфраструктуры, общий процент работы предлагаемой модели составил 80,52%. Исходя из проведенного эксперимента, можно сделать вывод о том, что детектор является слабым местом в модели и нуждается в доработке.

В табл. 3 показано время выполнения обнаружения и распознавания. Процессорная обработка алгоритмов модели занимает 1299,46 мс для всех вариантов разрешения изображений.

Стоит отметить, что разрешение входного кадра положительно сказывается на времени работы алгоритмов: чем ниже разрешение изображения, тем быстрее модель его обрабатывает. Однако с уменьшением разрешения входного видеокadra снижается качество работы модели.

Таблица 2

Результат работы алгоритма обнаружения дорожных знаков в видеопотоке

Порядковый номер класса	Наименование класса	Распознавание (%)
27–31	Знаки ограничения скорости	96,96
22	Въезд запрещен	98,09
26	Обгон запрещен	97,47
36	Движение прямо	97,83
38	Движение налево	98,1
37	Движение направо	97,06
39	Движение прямо или направо	98,12
40	Движение прямо или налево	97,82
41	Объезд препятствия справа	97,99
42	Объезд препятствия слева	97,07
47	Пешеходный переход	97,94
8	Уступить дорогу	97,64
14	Дети	97,63
19	Примыкание второстепенной дороги справа	97,29

Таблица 3

Производительность предлагаемой модели

Разрешение кадра	Обнаружение (мс)	Распознавание (мс)	Всего (мс)	Точность детекции (%)	Количество кадров / с
1920 x 1080	352,26	135,45	487,71	82,47	2,05
1632 x 918	252,03	95,69	347,72	82,12	2,88
1344 x 756	167,99	64,97	232,96	80,71	4,29
1056 x 594	97,33	40,16	137,49	77,73	7,27
768 x 432	49,02	20,92	69,94	70,7	14,30
480 x 270	15,47	8,17	23,64	42,7	42,30
Всего	934,1	365,36	1299,46		

Заключение

Для решения поставленной задачи поиска и распознавания объектов дорожной инфраструктуры целесообразно объединить несколько алгоритмов. Например, комбинирование алгоритмов цветовой сегментации с анализом контуров позволит повысить точность распознавания и минимизировать недостатки каждого из алгоритмов.

Была предложена модель системы поиска и распознавания дорожных знаков в потоковом видео на основе библиотеки OpenCV и сверточной нейронной сети. Детектор основан на алгоритмах компьютерного зрения. На вход классификатору подаются видеокadres, полученные с помощью алгоритма поиска дорожных знаков. В результате классификации пользователь получает конечную информацию о детектированном дорожном знаке.

В ходе экспериментов был продемонстрирован весь функционал, определен-

ный на шаге конструирования модели. Эксперименты показали, что детектор дорожных знаков, основанный на алгоритмах компьютерного зрения, неустойчив к некоторому роду шумов: частичное перекрытие знаков, низкое качество видео ввиду малого разрешения камеры, а также плохой ее стабилизации и недостаточной цветопередачи.

Таким образом, можно сделать вывод о том, что детектор недостаточно эффективен для решения поставленной задачи в реальном времени. В дальнейшем планируется улучшить алгоритмы детектора, что позволит минимизировать количество ложных срабатываний.

Список литературы

1. Смольянинов В.А., Белов Ю.С. Сравнительный анализ алгоритмов поиска дорожных знаков // Высокие технологии и инновации в науке: сборник избранных статей Международной научной конференции. 2020. С. 186–190.

2. Смольянинов В.А., Гришунов С.С., Белов Ю.С. Подходы к проектированию и разработке программного детектора дорожных знаков // Сборник избранных статей по материалам научных конференций ГНИИ «Нацразвитие»: материалы Международных научных конференций. 2020. С. 145–149.
3. Bradski G., Kaehler A. Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol: O'Reilly Media, 2008. 571 p. [Electronic resource]. URL: <http://bookre.org/reader?file=1077338> (date of access: 23.04.2021).
4. Шахуро В., Конушин А. Российская база изображений автодорожных знаков // Компьютерная оптика. 2016. № 2. С. 294–300.
5. Shawon A., Mahmud F., Md. Jamil-Ur Rahman, Arefin Zaman M.M. Bangla Handwritten Digit Recognition Using Deep CNN for Large and Unbiased Dataset. International Conference on Bangla Speech and Language Processing (ICBSLP). IEEE, 2018.
6. Class Data Science Project 2020 Traffic Signs Recognition [Electronic resource]. URL: <https://medium.com/dataflair/class-data-science-project-for-2020-traffic-signs-recognition-12b09c131742> (date of access: 23.04.2021).
7. Yasmina D., Karima R., Ouahiba A. Traffic signs recognition with deep learning. 2018 International Conference on Applied Smart Systems (ICASS), Medea, Algeria, 2018. P. 1–5. DOI: 10.1109/ICASS.2018.8652024.