

РОССИЙСКАЯ БАЗА ИЗОБРАЖЕНИЙ АВТОДОРОЖНЫХ ЗНАКОВ

В.И. Шахуро¹, А.С. Конушин^{1,2}

¹ НИУ Высшая школа экономики, Москва, Россия,

² Московский государственный университет им. М.В. Ломоносова, Москва, Россия

Аннотация

Представлена новая открытая база изображений дорожных знаков. База предназначена для обучения и тестирования алгоритмов распознавания дорожных знаков. Описывается структура и правила работы с базой, проводится сравнение с другими базами дорожных знаков. На базе проведена экспериментальная оценка современных алгоритмов выделения и классификации дорожных знаков, которая показала, что существующие алгоритмы распознавания большого класса знаков не достигают требуемой для ряда практических приложений полноты и точности.

Ключевые слова: база изображений дорожных знаков, классификация и выделение дорожных знаков, каскад слабых классификаторов, свёрточная нейронная сеть.

Цитирование: Шахуро, В.И. Российская база изображений автодорожных знаков / В.И. Шахуро, А.С. Конушин // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 294-300. – DOI: 10.18287/2412-6179-2016-40-2-294-300.

Введение

Рассмотрим задачу распознавания автодорожных знаков на изображении. На вход алгоритму распознавания подаётся последовательность видеокadres. Выходом алгоритма являются прямоугольники, содержащие автодорожные знаки на кадрах и классы найденных знаков. Задача имеет несколько важных применений: 1) в системах помощи водителю (ADAS, Advanced Driver Assistance Systems); 2) при автоматическом составлении и пополнении навигационных карт; 3) в системах мониторинга знаков на дорогах для служб дорожного хозяйства.

В данной работе рассматривается случай, когда искомый алгоритм распознавания работает независимо для каждого кадра, т.е. не использует информацию, полученную с соседних кадров. Тогда работу алгоритма распознавания можно разделить на два этапа: выделение и классификация. На первом этапе выделяются все знаки на кадре, на втором этапе найденные знаки классифицируются.

Дорожный знак — стандартизированный по размеру и внешнему виду объект. Несмотря на это, распознавание знаков на изображении остаётся сложной задачей по нескольким причинам:

1. Классов дорожных знаков много (156 в базе, представленной в данной работе), они сильно отличаются формой и рисунком. Существующие системы помощи водителю находят ограниченный набор классов (ограничение скорости, стоп, уступи дорогу, пешеходный переход), что упрощает задачу.
2. Для таких задач, как мониторинг дорог и построение карт, решение должно работать с полнотой, близкой к 100% и 1 ложным обнаружением на минуту видеопотока. Такое количество ложных обнаружений достигается при точности более 90%.
3. Для задачи системы помощи водителю решение должно работать в реальном времени. Дорожные знаки на изображении могут быть небольшого размера (от 16×16 пикселей), и для их выделения обычно строится многомасштабная пирамида изображения, в которой поиск производится ок-

ном с небольшим шагом (например, 4 пикселя). В результате получается порядка миллиона возможных положений окна в пирамиде, и для работы в реальном времени алгоритм выделения знаков должен быстро отбрасывать ложные обнаружения.

4. Различные погодные условия (дождь, снег, яркое солнце) и перекрытия объектов из-за динамического окружения усложняют задачу.

Современные алгоритмы распознавания объектов на изображении используют машинное обучение. Характеристики решений, полученных с помощью машинного обучения, существенно зависят от размера и качества обучающей выборки. Основные результаты данной статьи – представление новой базы, собранной на российских дорогах; её сравнительный анализ с существующими публичными базами дорожных знаков и экспериментальная оценка избранных современных методов распознавания объектов. База будет полезна для исследования и улучшения работы алгоритмов распознавания знаков. База доступна по адресу <http://graphics.cs.msu.ru/en/research/projects/rtsd>

1. Обзор существующих работ

1.1. Базы дорожных знаков

Мы рассматриваем самые крупные открытые базы автодорожных знаков: немецкие (GTSDb и GTSDb), шведскую (STS), бельгийскую (BTSD) и американскую (LISA). На рис. 1а–г показаны примеры кадров из этих баз.

В табл. 1 представлены количественные характеристики баз знаков в сравнении с представленной в данной статье базой RTSD. Анализ этих характеристик показывает, что существующие базы имеют следующие особенности:

- небольшое количество кадров (GTSDb), что делает базу недостаточно репрезентативной для тестирования детектора знаков (например, некоторые классы знаков есть только в тренировочной выборке);
- ограниченное количество классов знаков (STS), что делает невозможной оценку обобщаемости алгоритма на большое количество классов знаков;

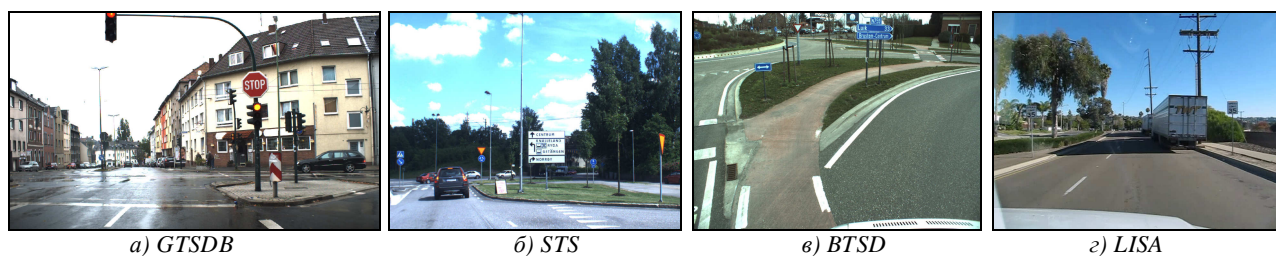


Рис. 1. Примеры кадров из баз дорожных знаков

Табл. 1. Статистика публичных баз дорожных знаков

	GTSRB [1]	GTSDDB [2]	STS [3]	BTSD [4]	LISA [5]	RTSD
Количество кадров	-	900	4000	25630	6610	179138
Количество знаков	43	43	7	108	47	156
Физических знаков	1728	1213	-	4565	-	15630
Изображений знаков	51839	1213	3488	13444	7855	104358

- количество изображений, приходящихся на один класс, небольшое (BTSD, LISA), что усложняет тестирование классификаторов, требующих больших обучающих выборок (например, сверточные нейронные сети).

Итого, существующие выборки нерепрезентативны либо для детектора, либо для классификатора, и поэтому не подходят для комплексного обучения и тестирования системы распознавания знаков (детектор + классификатор).

1.2. Методы распознавания знаков на изображениях

Алгоритмы распознавания объектов на изображении можно условно разделить на две группы: на основе эвристических алгоритмов и на основе машинного обучения.

Эвристические алгоритмы для поиска дорожных знаков используют то, что знаки имеют фиксированный цвет и форму. В [3] на входном изображении считается карта краёв, которая затем сопоставляется с шаблоном знака с помощью преобразования Фурье. В [6] рассматривается задача выделения знаков с красными рамками. Входное изображение преобразуется: переводится в пространство HSV, отсекается по порогу, фильтруется от шума. Финальные гипотезы получаются с помощью обобщённого преобразования Хафа. В [7] круглые знаки выделяются с помощью каналов, выделяющих области преобладания цвета, и преобразования Хафа.

К недостаткам эвристических алгоритмов можно отнести неустойчивость при размытых входных изображениях и сложность их построения в случае большого количества знаков разных цветов и формы. Далее рассмотрим основные подходы с использованием машинного обучения.

Подход на основе каскада слабых классификаторов начинается с работы Виолы и Джонса [8], в которой впервые была решена задача выделения лиц на изображении в реальном времени. Метод использует быстро вычисляемые интегральные признаки и неглубокие деревья решений (слабые классификаторы), которые объединяются с помощью бустинга в каскад (сильный

классификатор). Сильный классификатор последовательно применяет слабые классификаторы. После каждого слабого классификатора часть окон отбрасывается. Таким образом, целиком каскад проходят только окна с объектами и наиболее сложными примерами фона. Модифицированные варианты каскадного подхода показывают высокое качество и скорость на задачах выделения объектов с небольшой внутриклассовой изменчивостью: пешеходах [9–11], дорожных знаках [12, 13].

Другой подход основан на гистограммах ориентированных градиентов (HOG) и машине опорных векторов (SVM). В [14] была показана эффективность метода для задачи выделения пешеходов на изображении. Описание изображения с помощью HOG оказалось эффективным и для задач многоклассовой классификации. В [13, 15] проводится сравнительный анализ различных гистограмм ориентированных градиентов, ядерных функций в SVM и других классификаторов на задаче классификации дорожных знаков базы GTSRB.

Последний подход, глубинное обучение, переживает бурный рост в последние годы. Подход стал активно использоваться для решения различных задач компьютерного зрения после работы [16], в которой сверточная нейронная сеть успешно используется для классификации изображений базы ImageNet на 1000 классов. В [17] комитет сверточных нейронных сетей используется для классификации базы GTSRB и превосходит человека по точности классификации. В [18] для обучения классификатора дорожных знаков используются синтетические изображения знаков. Сверточная нейронная сеть, обученная на таких данных, показывает качество, сравнимое с нейронной сетью, обученной на реальных данных. Это позволяет решить проблему нерепрезентативных выборок и редко встречающихся классов знаков. Однако эксперименты в [19] показывают, что детектор дорожных знаков, обученный на синтетических данных, показывает неудовлетворительное качество. В [20] каскад из трёх нейронных сетей используется для быстрого и качественного выделения лиц на изображении. Этот метод является перспективным и для задачи выделения дорожных знаков на изображении.

2. Описание российской базы знаков

Для составления базы RTSD использовались кадры, предоставленные компанией Геоцентр-Консалтинг (<http://geocenter-consulting.ru>). Кадры получены с широкоформатных видеорегистраторов, установленных за лобовым стеклом машины. Видеорегистраторы снимают со скоростью 5 кадров/с. Разрешение кадров – от 1280×720 до 1920×1080. Кадры сняты в различные времена года (весна, осень, зима), времена суток (утро, день, вечер) и при различных погодных условиях

(дождь, снег, яркое солнце). Примеры кадров представлены на рис. 2а-е.

Разметка знаков на кадрах проводилась в два этапа. На первом этапе выделялись треки физических знаков на последовательностях кадров. На втором этапе отбрасывались неразличимые изображения знаков, и каждому физическому знаку присваивался класс. На рис. 3а-б показаны интерфейсы использованных программ. Исходные коды программ для разметки треков объектов и классов дорожных знаков распространяются вместе с базой RTSD.

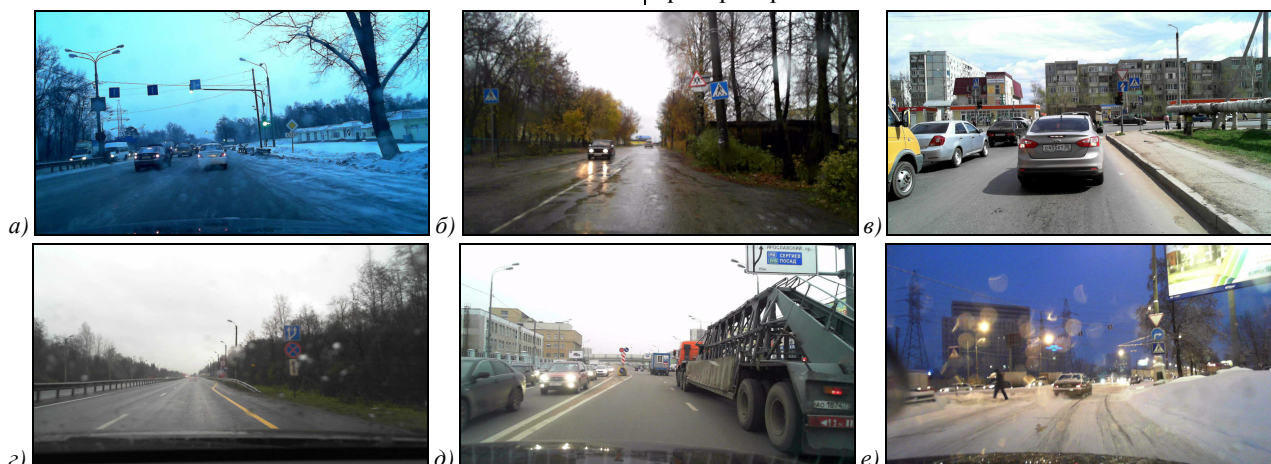


Рис. 2. Кадры из базы RTSD, демонстрирующие различные времена года, погоду и освещённость

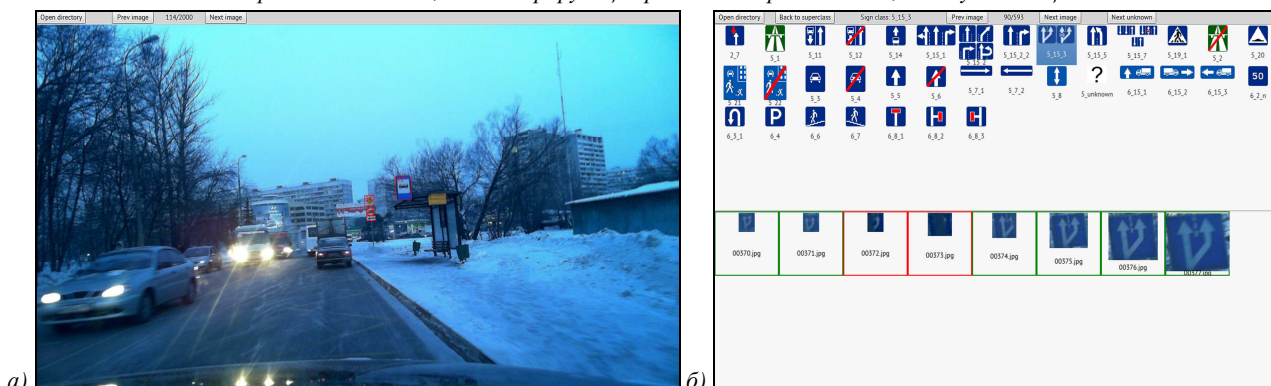


Рис. 3. Интерфейс программ, использовавшихся для разметки базы RTSD: разметка треков объектов (а), разметка классов знаков (б)

Для тестирования алгоритмов распознавания знаков было создано несколько выборок. Выборки содержат группы классов «предписания» (синие круги), «запреты» (красные треугольники), «ограничения» (круги с красной рамкой), «главная дорога» (жёлтый ромб), «сервис» (прямоугольники с синей рамкой), «особые предписания» (синие прямоугольники). Классы знаков, не принадлежащие данным группам, при составлении выборок не использовались. Кроме того, исключались редкие классы знаков (имеющие меньше 3 физических знаков или меньше 20 изображений).

При формировании выборок для детектора с каждого физического знака бралось по одному изображению (т.е. кадры прореживались так, чтобы каждому физическому знаку соответствовало одно изображение). Выборка RTSD-D1 по классам знаков аналогична базе GTSD. Размеры выборок для де-

тектора RTSD-D1, RTSD-D2, RTSD-D3 представлены в табл. 2.

Для классификатора формировались две выборки RTSD-R1 и RTSD-R3. Они совпадают по классам знаков с выборками RTSD-D1 и RTSD-D3 и содержат вырезанные изображения всех физических знаков. Количественные характеристики выборок для классификации представлены в табл. 3.

Как и в GTSD, выборки из RTSD для различных групп классов используют общие кадры. Это позволяет смоделировать полноценную систему распознавания дорожных знаков, в которой знаки сначала выделяются несколькими детекторами различных классов знаков, а затем в совокупности распознаются классификатором. Тренировочные и тестовые части выборок получались путём деления выборки в пропорции 3:1.

Табл. 2. Статистики выборок из RTSD и результаты тестирования детектора на них








						
RTSD-D1						
обучение (3821 кадров)	1054	1594	1842			
тестирование (1274 кадра)	396	578	605			
Качество детектора (AUC)	0,79	0,90	0,83			
RTSD-D2						
обучение (4786 кадров)	1033	1617	1848	1268		
тестирование (1596 кадров)	455	591	626	329		
Качество детектора (AUC)	0,82	0,89	0,80	0,92		
RTSD-D3						
обучение (9065 кадров)	1164	1800	2099	1678	1235	6843
тестирование (3022 кадра)	501	651	684	431	474	2085
Качество детектора (AUC)	0,80	0,86	0,72	0,90	0,83	0,76

Табл. 3. Результаты эксперимента с неплотной пирамидой и игнорированием знаков других классов при подсчёте качества выделения знаков на выборке RTSD-D2

				
Неплотная пирамида	0,82	0,89	0,80	0,92
+ игнорирование знаков других классов	0,86	0,90	0,82	0,94

3. Экспериментальная оценка существующих алгоритмов

3.1. Детектор

Для экспериментов по выделению знаков использовалась реализация каскадного метода на интегральных признаках из библиотеки Петра Доллара [21]. Детектор обучался с параметрами, схожими с [13]: 10 каналов для подсчёта признаков (LUV, величина градиента, шесть направлений градиента), каскад из 400 деревьев решений глубины 2, который обучается в четыре этапа с бутстрэппингом (по 2000 негативных примеров на каждом этапе), на этапах обучаются {50, 100, 200, 400} деревьев решений соответственно. Для многомасштабного поиска дорожных знаков (от 16×16 до 128×128 пикселей) строится пирамида из 50 масштабов изображений. Для каждой группы классов знаков обучается модель размером 56×56 пикселей. Для повышения точности выделения на GTSDDB каждая обученная модель масштабируется и тестируется пять раз (с соотношениями ширина/высота модели {0,8, 0,9, 1,0, 1,1, 1,2}), итоговые обнаружения сливаются.

Для подсчёта точности и полноты использовалась мера PASCAL пересечения двух прямоугольников. Эта мера равна отношению площади пересечения к площади объединения прямоугольников.

Обнаружение считается верным, если оно пересекается с прямоугольником из разметки более чем на 0,5. В качестве метрики качества использовалась площадь под ROC-кривой, AUC (Area Under Curve). Эта метрика является стандартной в задаче выделения объектов на изображении и используется при оценке детекторов на базе GTSDDB [2,13].

Результаты тестирования детектора на выборках RTSD-D1, RTSD-D2, RTSD-D3 представлены в табл. 2.

Наши эксперименты показывают, что растяжение модели в 5 различных пропорциях, как в [13], незначительно улучшает качество (около 0,005 AUC). Также заметим, что пирамида в детекторе [13], как в наших первых экспериментах, строится очень плотная – 50 слоёв. Эксперимент с 25 слоями пирамиды и 1 пропорцией модели показывает, что качество меняется незначительно (в пределах 0,01 AUC).

Был проведён эксперимент с игнорированием знаков других групп классов. Фиксируется группа искомым классов, обучается и тестируется детектор для данной группы классов. Если детектор нашёл знаки других групп классов, то эти обнаружения не включаются в подсчёт оценки качества выделения. Результаты эксперимента показывают, что синие круги со стрелками («предписание») часто путаются с кругами с красной рамкой («запрет»). Это связано с тем, что у некоторых классов знаков с красной рамкой синий фон (запрет, ограничение парковки).

Результаты детектора на выборке RTSD-D2 в экспериментах с неплотной пирамидой и игнорированием классов представлены в табл. 3.

Сравним полученные результаты с заявленными во введении требованиями практических приложений (близко к 100 % полноты и порядка 90 % точности). Только детекторы, обученные на поиск узких классов знаков (главная дорога и красные треугольники), достигают требуемого качества. Решение общей задачи выделения, получаемое как совокупность детекторов, обученных для всех групп классов знаков, показывает неудовлетворительное качество.

3.2. Классификатор

Для классификации знаков использовалась свёрточная нейронная сеть. Модель взята из [17] и состоит из 8 слоёв. Описание конфигурации сети приведено в табл. 4.

Для реализации нейросети использовалась открытая библиотека Caffe [22]. На базе GTSRB реализация достигает точности 98%, что сопоставимо с результатом из [17], где нейронная сеть достигает точности 98,5%.

На базе RTSD нейросеть тестировалась на выборках RTSD-R1, RTSD-R3, RTSD-R1, RTSD-R3. Последние две выборки – вырезанные изображения знаков из выборок для детектора. Результаты тестирования нейросети на

RTSD представлены в табл. 5. Можно заметить, что увеличение обучающих данных при переходе от RTSD-D1 к RTSD-R1 и от RTSD-D3 к RTSD-R3 положительно сказывается на качестве классификации. Меньшая по сравнению с GTSDB точность классификации может объясняться более плохим качеством изображений знаков в RTSD (многие изображения сильно смазаны и сняты при сложных условиях освещения).

Табл. 4. Архитектура свёрточной нейронной сети, использовавшейся для классификации знаков

Слой	Тип	Количество карт в слое и нейронов	Ядро
0	Входной	3 карты по 48×48 нейронов	
1	Свёрточный	100 карт по 100×100 нейронов	7×7
2	Max pooling	100 карт по 21×21 нейронов	2×2
3	Свёрточный	150 карт по 18×18 нейронов	4×4
4	Max pooling	150 карт по 9×9 нейронов	2×2
5	Свёрточный	250 карт по 6×6 нейронов	4×4
6	Max pooling	250 карт по 3×3 нейронов	2×2
7	Полносвязный	300 нейронов	1×1
8	Полносвязный	43 нейрона (число классов)	1×1

Табл. 5. Результаты классификации выборок из RTSD с помощью свёрточной нейронной сети

Выборка	Точность классификации (%)
RTSD-D1 обучение – 4490 знаков тестирование – 1579 знаков	85,18
RTSD-R1 (66 классов) обучение – 25432 знака тестирование – 7551 знак	90,78
RTSD-D3 обучение – 14819 знаков тестирование – 4826 знаков	90,08
RTSD-R3 (106 классов) обучение – 70687 знаков тестирование – 22967 знаков	92,90

Заключение

В данной статье была представлена база российских автодорожных знаков RTSD. Эта база превосходит другие публичные базы дорожных знаков по количеству кадров, классов знаков, физических знаков и их изображений. Кроме того, база содержит кадры с различными погодными условиями, освещением и временами года. На базе проведён анализ работы алгоритма выделения объектов на основе градиентных признаков и мягкого каскада и алгоритма классификации объектов на основе свёрточных нейронных сетей. Анализ показывает, что текущие решения по детектированию и классификации знаков работают недостаточно качественно для практических приложений.

Благодарность

Работа выполнена при поддержке Сколковского института науки и технологии, договор №081-R, приложение А1.

Литература

1. **Stallkamp, J.** Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition / J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel // Proceedings of IEEE International Joint Conference on Neural Networks. – 2012. – Vol. 32. – P. 323-332.

2. **Houben, S.** Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark / S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, C. Igel // Proceedings of IEEE International Joint Conference on Neural Networks. – 2013. – P. 1-8.
3. **Larsson, F.** Using Fourier descriptors and spatial models for traffic sign recognition / F. Larsson, M. Felsberg // Image Analysis. – Berlin, Heidelberg: Springer, 2011. – P. 238-249.
4. **Timofte, R.** Multi-view traffic sign detection, recognition, and 3D localization / R. Timofte, K. Zimmermann, L. Van Gool // Machine Vision and Applications. – 2014. – Vol. 25, Issue 3. – P. 633-647.
5. **Mogelmoose, A.** Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey / A. Mogelmoose, M.M. Trivedi, T.B. Moeslund // IEEE Transactions on Intelligent Transportation Systems. – 2012. – Vol. 13(4). – P. 1484-1497.
6. **Якимов, П.Ю.** Предварительная обработка цифровых изображений в системах локализации и распознавания дорожных знаков // Компьютерная оптика. – 2013. – Т. 37, № 3. – С. 401-405.
7. **Ruta, A.** A New Approach for In-Vehicle Camera Traffic Sign Detection and Recognition / A. Ruta, Y. Li, F. Porikli, S. Watanabe, H. Kage, K. Sumi // Machine Vision and Applications. – 2009. – P. 509-513.
8. **Viola, P.** Rapid object detection using a boosted cascade of simple features / P. Viola, M. Jones // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2001. – Vol. 1. – P. 511-518.

9. **Dollár, P.** Crosstalk cascades for frame-rate pedestrian detection / P. Dollár, R. Appel, W. Kienzle // *Computer Vision—ECCV 2012*. – Berlin, Heidelberg: Springer, 2012. – P. 645-659.
10. **Dollár, P.** Fast feature pyramids for object detection / P. Dollár, R. Appel, S. Belongie, P. Perona // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2014. – Vol. 36, Issue 8. – P. 1532-1545.
11. **Benenson, R.** Pedestrian detection at 100 frames per second / R. Benenson, M. Mathias, R. Timofte, L. Van Gool // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. – 2012. – P. 2903-2910.
12. **Overett, G.** Creating robust high-throughput traffic sign detectors using centre-surround HOG statistics / G. Overett, L. Tychsen-Smith, L. Petersson, N. Pettersson, L. Andersson // *Machine Vision and Applications*. – 2014. – Vol. 25, Issue 3. – P. 713-726.
13. **Mathias, M.** Traffic sign recognition – how far are we from the solution? / M. Mathias, R. Timofte, R. Benenson, L. Van Gool // *Proceedings of IEEE International Joint Conference on Neural Networks*. – 2013. – P. 1-8.
14. **Dalal, N.** Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. – 2005. – Vol. 1. – P. 886-893.
15. **Лисицын, С.О.** Распознавание дорожных знаков с помощью метода опорных векторов и гистограмм ориентированных градиентов / С.О. Лисицын, О.А. Байда // *Компьютерная оптика*. – 2012. – Т. 36, № 2 – С. 289-295.
16. **Krizhevsky, A.** Imagenet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // *Advances in Neural Information Processing Systems*. – 2012. – P. 1097-1105.
17. **Cireşan, D.** Multi-column deep neural network for traffic sign classification / D. Cireşan, U. Meier, J. Masci, J. Schmidhuber // *Proceedings of IEEE International Joint Conference on Neural Networks*. – 2012. – Vol. 32. – P. 333-338.
18. **Moiseev, B.** Evaluation of Traffic Sign Recognition Methods Trained on Synthetically Generated Data / B. Moiseyev, A. Konev, A. Chigorin, A. Konushin // *Advanced Concepts for Intelligent Vision Systems*. – 2013. – P. 576-583.
19. **Chigorin, A.** A system for large-scale automatic traffic sign recognition and mapping / A. Chigorin, A. Konushin // *CMRT13 – City Models, Roads and Traffic 2013 (ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences)*. – 2013. – Vol. 3. – P. 13-17.
20. **Li, H.** A convolutional neural network cascade for face detection / H. Li, Z. Lin, X. Shen, J. Brandt, G. Hua // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. – 2015. – P. 5325-5334.
21. **Dollár, P.** Piotr's Computer Vision Matlab Toolbox (PMT) [Electronical Resource]. – URL: <http://vision.ucsd.edu/~pdollar/toolbox/doc/index.html> (request date 01.04.2015).
22. **Jia, Y.** Convolutional architecture for fast feature embedding / Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell // *Proceedings of the ACM International Conference on Multimedia*. – 2014. – P. 675-678.

Сведения об авторах

Шахуро Владислав Игоревич, 1993 года рождения, в 2015 году окончил МГУ имени М.В. Ломоносова. Аспирант НИУ Высшая школа экономики. Научные интересы: обработка изображений, компьютерное зрение, машинное обучение, программирование. E-mail: vshahuro@hse.ru.

Конушин Антон Сергеевич, 1980 года рождения, в 2002 году окончил МГУ имени М.В. Ломоносова. В 2005 году защитил кандидатскую диссертацию в ИПМ имени М.В. Келдыша РАН. Работает доцентом на ВМК МГУ имени М.В. Ломоносова. Научные интересы: компьютерное зрение, машинное обучение. E-mail: ktosh@graphics.cs.msu.ru.

Поступила в редакцию 10 июля 2015 г. Окончательный вариант – 25 февраля 2016 г.

RUSSIAN TRAFFIC SIGN IMAGES DATASET

V.I. Shakhuro¹, A.S. Konushin^{1,2}

¹*NRU Higher School of Economics, Moscow, Russia,*

²*Lomonosov Moscow State University, Moscow, Russia*

Abstract

A new public dataset of traffic sign images is presented. The dataset is intended for training and testing the algorithms of traffic sign recognition. We describe the dataset structure and guidelines for working with the dataset, comparing it with the previously published traffic sign datasets. The evaluation of modern detection and classification algorithms conducted using the proposed dataset has shown that existing methods of recognition of a wide class of traffic signs do not achieve the accuracy and completeness required for a number of applications.

Keywords: traffic sign dataset, traffic sign classification and detection, cascade of weak classifiers, convolutional neural network.

Citation: Shakhuro VI, Konushin AS. Russian traffic sign images dataset. *Computer Optics* 2016; 40(2): 294-300. DOI: 10.18287/2412-6179-2016-40-2-294-300.

Acknowledgements: This work was supported by the Skolkovo Institute of Science and Technology, the contract №081-R, Annex A1.

References

- [1] Stallkamp J, Schlipsing M, Salmen J, Igel C. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. *Neural networks* 2012; 32: 323-332.
- [2] Houben S, Stallkamp J, Salmen J, Schlipsing M, Igel C. Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark. *The International Joint Conference Neural Networks*; 2013: 1-8.
- [3] Larsson F, Felsberg M. Using Fourier descriptors and spatial models for traffic sign recognition. *Image Analysis* 2011; 238-249.
- [4] Timofte R, Zimmermann K, Van Gool L. Multi-view traffic sign detection, recognition, and 3d localisation. *Machine Vision and Applications* 2014; 25(3): 633-647.
- [5] Mogelmoose A, Trivedi MM, Moeslund TB. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. *IEEE Transactions on Intelligent Transportation Systems* 2012; 13(4): 1484-1497.
- [6] Yakimov PYu. Preprocessing of digital images in systems of location and recognition of road signs. *Computer Optics* 2013; 37(3): 401-405.
- [7] Ruta A, Li Y, Porikli F, Watanabe S, Kage H, Sumi K. A New Approach for In-Vehicle Camera Traffic Sign Detection and Recognition. *Machine Vision and Applications* 2009; 509-513.
- [8] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2001: 511-518.
- [9] Dollár P, Appel R, Kienzle W. Crosstalk cascades for frame-rate pedestrian detection. *Computer Vision (ECCV)*; 2012: 645-659.
- [10] Dollár P, Appel R, Belongie S, Perona P. Fast feature pyramids for object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2014; 36(8):1532-1545.
- [11] Benenson R, Mathias M, Timofte R, Van Gool L. Pedestrian detection at 100 frames per second. *IEEE Conference on Computer Vision and Pattern Recognition*; 2012: 2903-2910.
- [12] Overett G, Tychsen-Smith L, Petersson L, Pettersson N, Andersson L. Creating robust high-throughput traffic sign detectors using centre-surround HOG statistics. *Machine Vision and Applications* 2014; 25(3): 713-726.
- [13] Mathias M, Timofte R, Benenson R, Van Gool L. Traffic sign recognition – How far are we from the solution? *International Joint Conference on Neural Networks*; 2013: 1-8.
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2005: 886-893.
- [15] Lisitsyn SO, Bayda OA. Road sign recognition using support vector machines and histogram of oriented gradients. *Computer Optics* 2012; 36(2): 289-295.
- [16] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*; 2012: 1097-1105.
- [17] Cireşan D, Meier U, Masci J, Schmidhuber J. Multi-column deep neural network for traffic sign classification. *Neural Networks* 2012; 32: 333-338.
- [18] Moiseev B, Konev A, Chigorin A, Konushin A. Evaluation of Traffic Sign Recognition Methods Trained on Synthetically Generated Data. *Advanced Concepts for Intelligent Vision Systems*; 2013: 576-583.
- [19] Chigorin A, Konushin A. A system for large-scale automatic traffic sign recognition and mapping. *CMRT13 – City Models, Roads and Traffic 2013 (ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences)* 2013; 3: 13-17.
- [20] Li H, Lin Z, Shen X, Brandt J, Hua G. A convolutional neural network cascade for face detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2015: 5325-5334.
- [21] Dollár P. Piotr's image and video Matlab Toolbox (PMT). Source: <http://vision.ucsd.edu/~pdollar/toolbox/doc/index.html>.
- [22] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the ACM International Conference on Multimedia*; 2014: 675-678.

Authors' information

Vladislav Igorevich Shakhuro, (b. 1993), graduated from Lomonosov Moscow State University in 2015. Currently graduate student at NRU Higher School of Economics. Research interests are image processing, computer vision, machine learning, and programming. E-mail: vshahuro@hse.ru.

Anton Sergeevich Konushin, (b. 1980), graduated from Lomonosov Moscow State University in 2002. In 2005 he successfully defended his PhD thesis in M.V. Keldysh Institute for Applied Mathematics RAS. He is currently associate professor at Lomonosov Moscow State University. Research interests are computer vision and machine learning. E-mail: ktosh@graphics.cs.msu.ru.

Received July 10, 2015. The final version – February 25, 2016.
