

Научная статья

УДК 519-7

doi: 10.17223/19988605/61/9

Использование ансамбля сверточных нейронных сетей для распознавания дорожных знаков

Игорь Константинович Харченко¹, Игорь Георгиевич Боровской²,
Елена Александровна Шельмина³

Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

¹ igor.k.kharchenko@tusur.ru

² igor.g.borovskoi@tusur.ru

³ elena.a.shelmina@tusur.ru

Аннотация. Предлагается для распознавания дорожных знаков использовать ансамбль сверточных нейронных сетей, который является модификацией робастного метода распознавания на основе нейронных сетей глубокого обучения. Данный ансамбль повышает скорость работы робастного метода распознавания, а также позволяет увеличить быстродействие с сохранением высокой точности распознавания за счет удаления из набора данных значений, которые не представляют полезной нагрузки.

Ключевые слова: распознавание дорожных знаков; сверточные нейронные сети; нейросетевой ансамбль

Благодарности: Исследование выполнено при финансовой поддержке гранта на реализацию программы стратегического академического лидерства «Приоритет-2030» в рамках научного проекта «Пр2030-МОН наука БЧ/СП2/Б/196».

Для цитирования: Харченко И.К., Боровской И.Г., Шельмина Е.А. Использование ансамбля сверточных нейронных сетей для распознавания дорожных знаков // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 88–96. doi: 10.17223/19988605/61/9

Original article

doi: 10.17223/19988605/61/9

Usage of convolutional neural network ensemble for traffic sign recognition

Igor I. Kharchenko¹, Igor G. Borovskoy², Elena A. Shelmina³

Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

¹ igor.k.kharchenko@tusur.ru

² igor.g.borovskoi@tusur.ru

³ elena.a.shelmina@tusur.ru

Abstract. The article suggests using an ensemble of convolutional neural networks for the recognition of road signs, which is a modification of a robust recognition method based on deep learning neural networks.

This ensemble improves the speed of the robust recognition method, and also allows you to increase the speed while maintaining high recognition accuracy by removing values from the data set that do not represent a payload.

Keywords: recognition of road signs; convolutional neural networks; neural network ensemble

Acknowledgments: The study was carried out with the financial support by a grant for the implementation of the program strategic academic leadership “Priority-2030” within the framework of the research project “Пр2030-МОН science БЧ/СП2/Б/196”.

For citation: Kharchenko, I.K., Borovskoy, I.G., Shelmina, E.A. (2022) Usage of convolutional neural network ensemble for traffic sign recognition. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 88–96. doi: 10.17223/19988605/61/9

В современном мире идентификация различных дорожных знаков является частью системы помощи водителю (СПБ, ADAS – Advanced Driver-Assistance System) [1], а также систем беспилотного управления автомобилем. Многие передовые компании, такие как Google, Uber, Tesla Motors, Volkswagen, Hyundai, NVIDIA, Baidu, ведут активную деятельность по реализации и тестированию подобных систем. Например, в Германии, Китае, США беспилотные автомобили уже используются в качестве такси и для грузоперевозок на дальние расстояния. Поэтому можно говорить об актуальности данной задачи. Для использования подобных систем в реальных дорожных условиях они должны быть устойчивыми к влиянию на качество распознавания различных факторов среды, таких как дождь, снег, туман, размытие и загрязнение линзы устройства видеозахвата.

Проблематике распознавания и локализации дорожных знаков посвящено немало количество научных работ. Для этих целей используются различные алгоритмы, детальный обзор которых представлен в [1, 2]. Рассмотрим классические алгоритмы, которые основываются на цвете и форме изображения. Качество распознавания у методов цветовой сегментации может зависеть от выбранного цветового пространства: к примеру, в работе [3] обнаружено, что использование цветового пространства Hue-Saturation-Intensity (HSI; оттенок–насыщенность–интенсивность) улучшает качество распознавания по сравнению с результатами, полученными на цветовом пространстве RGB в работе [4]. Авторы исследований [5, 6] установили, что применение цветовых пространств YCbCr и YUV также дает приемлемое качество распознавания. Однако такие методы очень чувствительны к изменению освещения, что является важным фактором при их применении в реальных условиях: распознавание должно быть корректным для разного времени суток. Если обратиться к методам сегментации по форме, то достаточно распространенным методом является детектор границ Кенни [7–9], а также каскад Хаара [10], быстрые трансформации Фурье [11], гистограмма ориентированных градиентов [5]. Применение методов сегментации по форме изображения конкретно к распознаванию дорожных знаков достаточно затруднительно, поскольку размеры и масштабы знака зависят от его удаленности от устройства захвата видео. Ситуацию усложняет еще и то, что знак может быть перекрыт другим объектом. Следовательно, можно прийти к выводу, что не существует какого-то конкретного способа качественного определения наилучшего классификатора для задачи распознавания дорожных знаков.

Недавние исследования в области машинного обучения привели к созданию инновационных подходов к решению данной задачи. Авторы работы [12] представили архитектуру сети, позволяющей одновременно вычислять и локализовать участок, на котором расположен знак. В работе [13] авторы пошли дальше и разделили локализацию и распознавание знака на две отдельные независимые нейросети: в качестве локализатора выступает гибридная архитектура SegU-Net, созданная на основе двух архитектур – SegNet [14] и U-Net [15], а для распознавания классов дорожных знаков используется архитектура VGG16 [16]. Далее авторы в рамках работы [17] решили модифицировать свой подход, добавив улучшение изображения в качестве отдельных этапов: классификатор на VGG16 обучен на распознавание погодных условий, если погодная ситуация требует нормализации (снег, дождь, туман), то изображение проходит через генеративно-сопоставительную сеть, которая приводит изображение к виду без плохих погодных условий, и после этого сеть SegU-Net локализует области дорожных знаков и распознает их.

Для локализации и распознавания объектов на изображениях используются два вида архитектур: одноэтапные и двухэтапные. В двухэтапных архитектурах применяются две сети: одна для локализации области расположения объекта, а вторая для распознавания его класса. В одноэтапных архитектурах оба действия производятся одной сетью. В работе [18] приводится сравнение нескольких архитектур нейросетевых детекторов: Faster R-CNN, R-FCN, Mask R-CNN, YOLO, SSD; из сравнения можно сделать вывод, что архитектура Faster R-CNN обладает высокой точностью, но низкой скоро-

стью работы, при этом алгоритмы семейства YOLO обладают самой высокой скоростью, при этом точность работы чуть ниже альтернатив.

Одним из существенных недостатков в данных работах является необходимость обучения сети на большом объеме обучающей выборки. К тому же скорость распознавания у YOLO гораздо выше предложенной авторами архитектуры SegU-Net.

Цель же данной работы – реализация нейросетевого ансамбля, подобного реализованному в [17], с заменой локализатора с SegU-Net на YOLO v4 [19]. Также для уменьшения скорости обучения предложено удалять области изображения, не содержащие никаких знаков. Получены следующие результаты: точность работы YOLO v4 оказалась в среднем на 5% ниже по каждому типу погодных условий, что обосновано неполной обучающей выборкой в данном исследовании в отличие от исследования [17]. При этом скорость увеличилась до 13 кадров в секунду по сравнению с 3 кадрами в секунду в работе [17]. Также за счет отсека областей без дорожных знаков на изображении удалось добиться уменьшения объема обучающей выборки в два раза по сравнению с исследованием [13].

1. Нейросетевой ансамбль

1.1. Описание подхода

Предложенный подход распознавания дорожных знаков состоит из трех архитектур нейронных сетей, каждая из которых выполняет отдельный набор действий. Первый модуль – классификатор погодных условий, отвечающий за распознавание типа погодных условий на изображении. Второй модуль отвечает за удаление плохих погодных условий на изображении, т.е. за приведение его к виду без плохих условий. Наконец, третий модуль – локализация и распознавание дорожных знаков на изображении.

Алгоритм данного нейросетевого ансамбля:

1. Исходное изображение передается классификатору погодных условий.
2. Если классификатором определяется одно из ухудшающих погодных условий, то изображение пропускается через модуль удаления плохих погодных условий.
3. Результирующее изображение подается на вход локализатора дорожных знаков, и в случае их наличия локализатор возвращает их ограничивающие прямоугольники (bounding box) с точностью распознавания для каждого такого прямоугольника.

Для проверки качества обучения нейронных сетей в области распознавания дорожных знаков существуют специальные наборы данных, такие как GTSDb [20], BelgiumTS [21]. Однако в данных наборах нет достаточного количества погодных условий, на которых можно было бы обучить нейросеть без применения методов аугментации обучающей выборки [22] (которые в свою очередь тоже не гарантируют полноту данных), для обучения. Чтобы преодолеть это ограничение, в работе [23] авторы реализовали набор данных, состоящий из видеорядов с добавленными погодными условиями, влияющими на качество распознавания изображения. Набор состоит из двух поднаборов из 49 видеорядов в каждом: один набор – это реальная съемка из автомобиля, второй – съемка в виртуальной смоделированной среде. В рамках данной работы используется только первая часть данного набора, снятая из реального автомобиля.

1.2. Модуль классификации плохих погодных условий

Классификатором погодных условий выступает сеть VGG19 [16], предобученная на наборе данных ImageNet и обученная на наборе данных CURE-TSD путем метода «передачи обучения» (transfer learning). Идея «передачи обучения» заключается в использовании свойств сверточных сетей: такие сети на каждом последующем слое распознают все более сложные фрагменты, например на неглубоких слоях сеть распознает углы, края, простые формы, на более глубоких слоях – простые фигуры и т.д. Таким образом, вся информация о признаках изображений хранится в 16 сверточных

слоях, а обобщения – в трех полносвязных выходных слоях. Следовательно, обучению подлежат только последние три слоя; остальные обучению не подлежат.

Экспериментальным путем установлено, что разрешения входного изображения в 224×224 пикселя достаточно для точности распознавания погодных условий более 99%. Поэтому на вход сети подаются изображения, уменьшенные до размера 224×224 пикселя; каждому изображению соответствует метка с погодными условиями: 0 – хорошие условия, от 1 до 6 – ухудшающие условия. Модель обучается на 15 эпохах обучения. Если ошибка обучения не уменьшается в течение трех эпох, шаг оптимизатора уменьшается наполовину (начальный шаг оптимизатора равен 0,0001).

1.3. Модуль нормализации плохих погодных условий

Для целей нормализации могут подойти такие архитектуры, как сверточный автоэнкодер и генеративно-сопоставительная сеть. Однако сверточный автоэнкодер неприменим для данной работы, поскольку его основное назначение – не улучшение изображения, а извлечение признаков из последнего. Поэтому для нормализации изображения применяется генеративно-сопоставительная сеть.

Архитектура состоит из двух сетей: генератора и дискриминатора. Генератор – это сеть, генерирующая изображения с шумом, которые подаются на вход дискриминатора, который, в свою очередь, учится «отделять» зашумленные изображения от незашумленных и тем самым удалять шум из изображения. Топология сети генератора и дискриминатора соответствует используемой в работе [17].

При обучении данной архитектуры используется результат суммы двух функций ошибки, приведенных авторами в работе [17].

Первая – это функция потерь с использованием средней абсолютной ошибки

$$L_{Wasserstein} = \frac{1}{H * W * C} \sum_{k=0}^C \sum_{j=0}^W \sum_{i=0}^H |R_{ijk} - T_{ijk}|, \quad (1)$$

где H и W – высота и ширина изображения, C – количество каналов изображения, R и T – реконструированное сетью и оригинальное изображение соответственно. Смысл использования именно такой функции (1) вместо наиболее часто используемой функции среднеквадратической ошибки в том, что, как обнаружено в работе [24], последняя накладывает эффект размытия на изображение, из-за чего нейросеть не может определить на нем мелкие детали. Применение функции средней абсолютной ошибки [25] дает лучшие результаты при реконструкции мелких деталей изображения.

Вторая – это так называемая перцептивная функция потерь

$$L_{perceptive} = \frac{1}{H * W * C} \sum_{k=0}^C \sum_{j=0}^W \sum_{i=0}^H |\varphi(R)_{ijk} - \varphi(T)_{ijk}|, \quad (2)$$

где $\varphi(R)$ и $\varphi(T)$ – признаки изображения, извлеченные из седьмого слоя сети VGG16, предобученной на наборе данных ImageNet, для реконструированного и оригинального изображения. Основная идея такого подхода с использованием функции (2) заключается в том, что необходимо максимально сохранить информацию о мелких частях дорожных знаков, а внутренние слои предобученной сети VGG16 уже содержат необходимую нам информацию о таких признаках, как углы, края и т.п.

При обучении модуля выбираются два равных фрагмента размером 1024×1024 пикселя с двух верхних углов каждого изображения. Обоснование такого подхода приведено в разд. 2. Обучение производится на 30 эпохах; в случае если в течение трех эпох нет качественного улучшения ошибки сети, то шаг обучения оптимизатора сокращается вдвое (начальный шаг равен 0,0001).

1.4. Модуль локализации и распознавания дорожных знаков

Принцип обучения модуля локализации и распознавания дорожных знаков на архитектуре YOLO v4 [19] схож с принципом обучения модуля нормализации: на вход поступают два равных фрагмента 1024×1024 пикселя по двум верхним углам изображения. Обучение производится на

4 000 эпохах: такое большое количество обосновано самой архитектурой сети, поскольку каждое изображение проходит через несколько этапов обработки.

2. Анализ набора данных, используемого для обучения модулей

В качестве набора данных выступает набор данных CURE-TSD, представляющий собой набор из 49 видеорядов по 300 кадров в каждом. Для каждого видеоряда есть по 70 фрагментов с нанесенными эффектами, чтобы смоделировать плохие погодные условия: 14 погодных условий с пятью уровнями ухудшения каждого условия. В рамках данного исследования отобраны только такие, которые встречаются в реальных условиях: дождь, снег, туман, загрязнение и размытие линзы, и снятые в реальном автомобиле; видеоряды, снятые в смоделированной среде, в данной работе не рассматриваются. Каждое из погодных условий также подразделяется на пять уровней: на первом хорошо различимы почти все знаки (погодные условия почти не влияют на качество распознавания), на последнем же знаки, находящиеся вдали, практически полностью невидны. Таким образом, для распознавания необходимо использовать по 26 видеорядов для каждого фрагмента снятого видео, в общей сложности 1 274 фрагмента.

Необходимо также принимать во внимание угол наклона камеры относительно дороги: от этого зависит, какая часть изображения действительно несет в себе полезную нагрузку, а сколько пикселей никогда не будет задействовано при распознавании. Соответственно, большая часть знаков расположена на правой части, по оси OY области пикселей знаков имеют нормальное распределение и расположены в основном в области от нуля до 646 пикселей. Следовательно, при таком угле наклона камеры можно отсечь почти 52% от всей области изображения без потери точности распознавания.

Также выявлено, что на одном кадре может располагаться от одного до пяти знаков с расстоянием между ними, превышающим 1 024 пикселя. Это говорит о том, что метод, который используют авторы в статье [17] (выборка фрагмента изображения размером $1\,024 \times 1\,024$ пикселя случайным образом), может быть применен только в ситуации, когда на изображении есть не более одного-двух знаков и они гарантированно будут находиться в одном фрагменте. Следовательно, это вполне могло оказать негативное влияние на точность распознавания знаков. С другой стороны, авторы статьи [13] для обучения сети SegU-Net выбирают четыре фрагмента размером $1\,024 \times 1\,024$ пикселя путем обрезки изображения по четырем углам (верхним и нижним).

Минимальная ширина знака, на котором обучается локализатор дорожных знаков, равна восьми, максимальная же составляет 246. Следовательно, минимальный размер фрагмента, на котором можно обучить второй и третий модули, может быть 512 или 1 024 (256 не берется из-за возможной потери информации о краях знака). Размер фрагмента должен быть кратен двум, поскольку в ином случае возможна потеря информации о контексте изображения при операции макс-пулинга во внутренних слоях сверточной сети [13]. Уменьшение размера фрагмента до 512 пикселей сократит размер сети, но не сократит время обучения, поэтому размер фрагмента обучения в рамках данной работы выбран равным $1\,024 \times 1\,024$ пикселя.

3. Результаты тестирования

Обучение производилось с использованием следующих вычислительных ресурсов: процессор Intel® Xeon® Gold 6226R, видеокарта Nvidia RTX 3090, объем оперативной памяти 64 ГБ.

По техническим ограничениям в рамках данной работы обучающая выборка ограничена только первыми 20 видеорядами. Для предотвращения переобучения обучающая, тестовая и валидационная выборки состоят из 16, 2 и 2 фрагментов соответственно.

Точность обучения классификатора погодных условий на валидационной выборке составила 99,54%. Ошибки распознавания возникают при распознавании дождя и размытия линзы на первом уровне, что в целом является приемлемым результатом. Авторы архитектуры [17] получили результат в 99,89%; такое весомое расхождение можно обосновать тем, что авторы использовали всю обучающую выборку, состоящую из 49 видеофрагментов, для обучения ансамбля.

Результаты работы модуля нормализации и локализации для каждого типа погодных условий на валидационной выборке (в столбце «Полученная») по сравнению с полученными в работе [17] (в столбце «Эталонная») приведены в таблице.

Точность модуля нормализации и локализации дорожных знаков

Условие	Уровень	Полученная	Эталонная
Без плохих условий	–	95,91	99,02
Дождь	1	97,28	98,15
	2	97,41	97,72
	3	97,41	97,38
	4	96,17	96,88
	5	96,05	96,03
Снег	1	96,67	98,95
	2	96,42	98,24
	3	96,54	98,57
	4	96,42	96,04
	5	93,29	94,94
Загрязнение линзы	1	95,99	98,55
	2	94,81	98,56
	3	94,19	98,33
	4	92,58	97,94
	5	92,11	97,33
Размытие линзы	1	96,30	98,62
	2	95,19	98,51
	3	94,45	98,22
	4	93,99	97,57
	5	93,12	97,31
Туман	1	95,68	98,31
	2	94,08	98,76
	3	93,71	98,85
	4	93,02	98,38
	5	92,83	97,74

Из результатов видно, что точность обучения на погодных условиях, которые перекрывают часть знака (дождь, снег), сопоставима с эталонной, однако на остальных условиях точность меньше в среднем на 5%. Это можно объяснить тем, что YOLO v4 немного хуже справляется с более мелкими элементами на изображении и не всегда может с высокой точностью распознать мелкие детали на большом удалении от знака.

В работе [17] скорость работы всего ансамбля в среднем равна 2,79 кадров в секунду. При использовании модифицированного ансамбля скорость работы увеличена до 12,76 кадров в секунду.

Заключение

В данной статье представлен нейросетевой ансамбль, являющийся модификацией работы [17], улучшающий скорость работы робастного метода распознавания, а также сокращающий скорость обучения с сохранением высокой точности распознавания за счет удаления из набора данных части изображений, которые не представляют полезной нагрузки. Классификатор погодных условий определяет, нужно ли улучшать изображение: в случае необходимости в работу подключается нейросеть для улучшения изображения. Нейросеть YOLO v4 локализует область знака и распознает его класс: использование данной нейросети вместо связки SegU-Net и VGG16 дает выигрыш в скорости распознавания классификатора при потере точности распознавания в 5% при таких условиях, как туман, размытие и загрязнение линзы. Обученные нейросети могут использоваться и на российских дорогах, поскольку видеоряды в обучающей выборке были сняты в Бельгии, где движение правостороннее, а знаки схожи с российскими.

При этом данный подход к распознаванию применим и к смежным задачам: например, к распознаванию госномеров автомобилей или определению на проезжей части людей, пересекающих проезжую часть в неположенном месте, для предотвращения аварийных ситуаций.

Список источников

1. Mogelmose A., Trivedi M.M., Moeslund T.B. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey // *IEEE Trans. Intell. Transp. Syst.* 2012. V. 13 (4). P. 1484–1497.
2. Gudigar A., Shreesha C., Raghavendra U. A review on automatic detection and recognition of traffic sign // *Multimedia Tools and Applications*. 2011. V. 75 (1). P. 333–364.
3. Chen Y., Xie Y., Wang Y. Detection and recognition of traffic signs based on HSV vision model and shape features // *Journal of Computers*. 2013. V. 8 (5). P. 1366–1370.
4. Zaklouta F., Stanculescu B. Real-time traffic sign recognition in three stages // *Robotics and Autonomous Systems*. 2014. V. 62 (1). P. 16–24.
5. Creusen I.M., Wijnhoven R.G.J., Herbschleb E., de With P.H.N. Color exploitation in HOG-based traffic sign detection. URL: <https://ieeexplore.ieee.org/document/5651637> (accessed: 11.08.2022).
6. Prieto M.S., Allen A.R. Using self-organising maps in the detection and recognition of road signs // *Image and Vision Computing*. 2009. V. 27 (6). P. 673–683.
7. Garcia-Garrido M.A., Ocaña M., Llorca D.F., Sotelo M.A., Arroyo E., Llamazares A. Robust traffic signs detection by means of vision and v2i communications. URL: <https://ieeexplore.ieee.org/document/6082844> (accessed: 11.08.2022).
8. Gonzalez A., Garrido M.A.G., Fernandez-Llorca D., Gavilan M. Automatic traffic signs and panels inspection system using computer vision // *IEEE Transactions on Intelligent Transportation Systems*. 2011. V. 12 (2). P. 485–499.
9. Greenhalgh J., Mirmehdi M. Recognizing text-based traffic signs. // *IEEE Transactions on Intelligent Transportation Systems*. 2015. V. 16 (3). P. 1360–1369.
10. Baro X., Escalera S., Vitria J., Pujol O., Radeva P. Traffic sign recognition using evolutionary AdaBoost detection and Forest-ECOC classification // *IEEE Transactions on Intelligent Transportation Systems*. 2009. V. 10 (1). P. 113–126.
11. Jimenez P.G., Bascon S.M., Moreno H.G., Arroyo S.L., Ferreras F.L. Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies // *Signal Processing*. 2008. V. 88 (12). P. 2943–2955.
12. Lee H.S., Kim K. Simultaneous traffic sign detection and boundary estimation using convolutional neural network // *IEEE Transactions on Intelligent Transportation Systems*. 2018. V. 19 (5). P. 1652–1663.
13. Kamal U., Tonmoy T.I., Das S., Hasan M.K. Automatic traffic sign detection and recognition using SegU-Net and a modified Tversky loss function with L1-constraint // *IEEE Transactions on Intelligent Transportation Systems*. 2020. V. 21 (4). P. 1467–1479.
14. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017. V. 39 (12). P. 2481–2495.
15. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // *Medical Image Computing and Computer Assisted Intervention : Proc. of the 21st International Conference*. 2015. V. 9351. P. 234–241.
16. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. URL: <https://arxiv.org/abs/1409.1556v6> (accessed: 11.08.2022).
17. Ahmed S., Kamal U., Hasan M.K. DFR-TSD: A Deep Learning Based Framework for Robust Traffic Sign Detection Under Challenging Weather Conditions // *IEEE Transactions on Intelligent Transportation Systems*. 2021. V. 23 (6). P. 5150–5162.
18. Sanchez S.A., Romero H.J., Morales A.D. A review: Comparison of performance metrics of pretrained models for object detection using the TensorFlow framework. URL: https://www.researchgate.net/publication/342570032_A_review_Comparison_of_performance_metrics_of_pretrained_models_for_object_detection_using_the_TensorFlow_framework (accessed: 11.08.2022).
19. Bochkovskiy A., Wang C.-Ya., Mark L.H.-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection. URL: <https://arxiv.org/abs/2004.10934> (accessed: 11.08.2022).
20. Houben S., Stallkamp J., Salmen J., Schlipsing M., Igel C. Detection of traffic signs in real-world images: The German traffic sign detection benchmark. URL: <https://ieeexplore.ieee.org/document/6706807> (accessed: 11.08.2022).
21. Timofte R., Zimmermann K., Van Gool L. Multi-view traffic sign detection, recognition, and 3D localization // *Machine Vision and Applications*. 2014. V. 25 (3). P. 633–647.
22. Емельянов С.О., Иванова А.А., Швец Е.А., Николаев Д.П. Методы аугментации обучающих выборок в задачах классификации изображений // *Сенсорные системы*. 2018. Т. 32, № 3. С. 236–245.
23. Temel D., Alshawi T., Chen M.-H., AlRegib G. CURE-TSD: Challenging unreal and real environments for traffic sign detection // *IEEE Dataport*. URL: <https://ieee-dataport.org/open-access/cure-tds-challenging-unreal-and-real-environment-traffic-sign-detection> (accessed: 11.08.2022).
24. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error. URL: <https://arxiv.org/abs/1511.05440> (accessed: 11.08.2022).
25. Zhao H., Gallo O., Frosio I., Kautz J. Loss functions for neural networks for image processing. URL: <http://arxiv.org/abs/1511.08861> (accessed: 11.08.2022).

References

1. Mogelmose, A., Trivedi, M.M. & Moeslund, T.B. (2012) Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. *IEEE Transactions on Intelligent Transportation Systems*. 13(4). pp. 1484–1497. DOI: 10.1109/TITS.2012.2209421
2. Gudigar, A., Shreesha, C. & Raghavendra, U. (2016) A review on automatic detection and recognition of traffic sign. *Multimedia Tools and Applications*. 75(1). pp. 333–364. DOI: 10.1007/s11042-014-2293-7
3. Chen, Y., Xie, Y. & Wang, Y. (2013) Detection and recognition of traffic signs based on HSV vision model and shape features. *Journal of Computers*. 8(5). pp. 1366–1370. DOI: 10.4304/jcp.8.5.1366-1370
4. Zaklouta, F. & Stanculescu, B. (2014) Real-time traffic sign recognition in three stages. *Robotics and Autonomous Systems*. 62(1). pp. 16–24. DOI: 10.1016/j.robot.2012.07.019
5. Creusen, I.M., Wijnhoven, R.G.J., Herbschleb, E. & de With, P.H.N. (2010) *Color exploitation in HOG-based traffic sign detection*. [Online] Available from: <https://ieeexplore.ieee.org/document/5651637>. (Accessed: 11th August 2022).
6. Prieto, M.S. & Allen, A.R. (2009) Using self-organising maps in the detection and recognition of road signs. *Image and Vision Computing*. 27(6). pp. 673–683. DOI: 10.1016/j.imavis.2008.07.006
7. Garcia-Garrido, M.A., Ocaña, M., Llorca, D.F., Sotelo, M.A., Arroyo, E. & Llamazares, A. (2011) *Robust traffic signs detection by means of vision and v2i communications*. [Online] Available from: <https://ieeexplore.ieee.org/document/6082844>. (Accessed: 11th August 2022).
8. Gonzalez, A., Garrido, M.A.G., Fernandez-Llorca, D. & Gavilan, M. (2011) Automatic traffic signs and panels inspection system using computer vision. *IEEE Transactions on Intelligent Transportation Systems*. 12(2). pp. 485–499. DOI: 10.1109/TITS.2010.2098029
9. Greenhalgh, J. & Mirmehdi, M. (2015) Recognizing text-based traffic signs. *IEEE Transactions on Intelligent Transportation Systems*. 16(3). pp. 1360–1369. DOI: 10.1109/TITS.2014.2363167
10. Baro, X., Escalera, S., Vitria, J., Pujol, O. & Radeva, P. (2009) Traffic sign recognition using evolutionary AdaBoost detection and Forest-ECOC classification. *IEEE Transactions on Intelligent Transportation Systems*. 10(1). pp. 113–126. DOI: 10.1109/TITS.2008.2011702
11. Jimenez, P.G., Bascon, S.M., Moreno, H.G., Arroyo, S.L. & Ferreras, F.L. (2008) Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies. *Signal Processing*. 88(12). pp. 2943–2955. DOI: 10.1016/j.sigpro.2008.06.019
12. Lee, H.S. & Kim, K. (2018) Simultaneous traffic sign detection and boundary estimation using convolutional neural network. *IEEE Transactions on Intelligent Transportation Systems*. 19(5). pp. 1652–1663. DOI: 10.48550/arXiv.1802.10019
13. Kamal, U., Tonmoy, T.I., Das, S. & Hasan, M.K. (2020) Automatic traffic sign detection and recognition using SegU-Net and a modified Tversky loss function with L1-constraint. *IEEE Transactions on Intelligent Transportation Systems*. 21(4). pp. 1467–1479.
14. Badrinarayanan, V., Kendall, A. & Cipolla, R. (2017) SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 39(12). pp. 2481–2495. DOI: 10.48550/arXiv.1511.00561
15. Ronneberger, O., Fischer, P. & Brox, T. (2015) U-net: Convolutional networks for biomedical image segmentation. *Proc. of the 21st Medical Image Computing and Computer Assisted Intervention International Conference*. 9351. pp. 234–241.
16. Simonyan, K. & Zisserman, A. (n.d.) *Very Deep Convolutional Networks for Large-Scale Image Recognition*. [Online] Available from: <https://arxiv.org/abs/1409.1556v6>. (Accessed: 11th August 2022).
17. Ahmed, S., Kamal, U. & Hasan, M.K. (2021) DFR-TSD: A Deep Learning Based Framework for Robust Traffic Sign Detection Under Challenging Weather Conditions. *IEEE Transactions on Intelligent Transportation Systems*. 23(6). pp. 5150–5162. DOI: 10.13140/RG.2.2.18341.86249
18. Sanchez, S.A., Romero, H.J. & Morales, A.D. (2020) A review: Comparison of performance metrics of pretrained models for object detection using the TensorFlow framework. *IOP Conference Series Materials Science and Engineering*. 844:012024. DOI: 10.1088/1757-899X/844/1/012024
19. Bochkovskiy, A., Wang, C.-Ya. & Mark, L.H.-Y. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. [Online] Available from: <https://arxiv.org/abs/2004.10934>. (Accessed: 11th August 2022).
20. Houben, S., Stallkamp, J., Salmen, J., Schlipsing, M. & Igel, C. *Detection of traffic signs in real-world images: The German traffic sign detection benchmark*. [Online] Available from: <https://ieeexplore.ieee.org/document/6706807> (Accessed: 11th August 2022).
21. Timofte, R., Zimmermann, K. & Van Gool, L. (2014) Multi-view traffic sign detection, recognition, and 3D localization. *Machine Vision and Applications*. 25(3). pp. 633–647. DOI: 10.1007/s00138-011-0391-3
22. Emelyanov, S.O., Ivanova, A.A., Shvets, E.A. & Nikolaev, D.P. (2018) Methods of training data augmentation in the task of image classification. *Sensornye sistemy – Sensory Systems*. 32(3). pp. 236–245. DOI: 10.1134/S0235009218030058
23. Temel, D., Alshawi, T., Chen, M.-H. & AlRegib, G. (2021) CURE-TSD: Challenging unreal and real environments for traffic sign detection. *IEEE Dataport*. [Online] Available from: <https://iee-dataport.org/open-access/cure-tds-challenging-unreal-and-real-environment-traffic-sign-detection>. (Accessed: 11th August 2022).
24. Mathieu, M., Couprie, C. & LeCun, Y. (2016) *Deep multi-scale video prediction beyond mean square error*. [Online] Available from: <https://arxiv.org/abs/1511.05440>. (Accessed: 11th August 2022).
25. Zhao, H., Gallo, O., Frosio, I. & Kautz, J. (n.d.) *Loss functions for neural networks for image processing*. [Online] Available from: <http://arxiv.org/abs/1511.08861>. (Accessed: 11th August 2022).

Информация об авторах:

Харченко Игорь Константинович – аспирант кафедры автоматизированных систем управления Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: igor.k.kharchenko@tusur.ru

Боровской Игорь Георгиевич – профессор, доктор физико-математических наук, заведующий кафедрой экономической математики, информатики и статистики Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: igor.g.borovskoi@tusur.ru

Шельмина Елена Александровна – кандидат физико-математических наук, доцент кафедры экономической математики, информатики и статистики Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: elena.a.shelmina@tusur.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Kharchenko Igor Konstantinovich (Post-graduate student, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: igor.k.kharchenko@tusur.ru

Borovskoy Igor Georgievich (Doctor of Physics and Mathematics, Professor, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: igor.g.borovskoi@tusur.ru

Shelmina Elena Aleksandrovna (Candidate of Physics and Mathematics, Associate Professor, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: elena.a.shelmina@tusur.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Received 12.08.2022; accepted for publication 29.11.2022

Поступила в редакцию 12.08.2022; принята к публикации 29.11.2022