

АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ АВТОМОБИЛЬНЫХ НОМЕРОВ В ВИДЕОРЯДЕ: АДАПТАЦИЯ АЛГОРИТМОВ ДЛЯ СМАРТФОНА

© 2023 М. В. Абрамов^{1✉}, Д. А. Ельцов²

¹Санкт-Петербургский Федеральный исследовательский центр Российской академии наук
14-я линия В.О., 39, 199178 Санкт-Петербург, Российская Федерация

²Huawei Technologies Co. Ltd.

ул Тамбовская, 12, 192007 Санкт-Петербург, Российская Федерация

Аннотация. В данной области уже проведено немало исследований, и предложены различные модели, решающие задачу распознавания автомобильных номеров, есть промышленные образцы, используемые на камерах фиксации нарушений правил дорожного движения. Однако, такие модели обычно разработаны для клиент-серверных архитектур, поскольку архитектуры моделей, хорошо решающих данную задачу, зачастую имеют сотни миллионов параметров [5] и рассчитаны для того, чтобы их обучали и применяли на машинах с большим объемом памяти и производительными видеокартами. Предметом исследования данной статьи являются автомобильные государственные номерные знаки в видеоряде плохого качества, распознаваемые в условиях ограничения на вычислительные ресурсы. Для распознавания автомобильных номеров предлагается использовать следующие методы: простые сверточные нейронные сети для определения углов таблички автомобильного номера и ее последующего поворота; предобученная модель YOLOv5s для поиска в кадре табличек автомобильных номеров и предобученная сеть SCR-Net для распознавания на табличках номерных знаков. В статье предложены алгоритм и его реализация для автоматического распознавания автомобильного номера в условиях видео плохого качества и ограниченности вычислительных ресурсов. Теоретическая значимость результата заключается в разработке нового алгоритма, учитывающего указанные ограничения, способствующего развитию серии изысканий в области разработки методов и алгоритмов, функционирующих в условиях существенных ограничений на вычислительные мощности и качество видеоряда. Практическая значимость результата заключается в приложениях результата к решению прикладных задач, таких как фиксация неправильной парковки, поиск угоняемых автомобилей, иных правонарушений.

Ключевые слова: распознавание автомобильных номеров, алгоритмы распознавания, распознавание на смартфоне, распознавание номеров.

ВВЕДЕНИЕ

С развитием глубоких нейронных сетей и появлением видеокарт с большим объёмом памяти, область компьютерного зрения начала достигать результатов уровня человека в распознавании образов и обнаружении объектов

[1–4]. Это позволило значительно сократить человеческие трудозатраты, облегчить труд путём автоматизации монотонных процессов [5]. Одним из таких процессов стала автоматизация распознавания автомобильных номерных знаков. Решения по автоматическому распознаванию автомобильных номеров имеет множество востребованных приложений [6]. Результаты, полученные при решении этой задачи, используются в умных транспортных

✉ Абрамов Максим Викторович
e-mail: mva@dscs.pro



Контент доступен под лицензией Creative Commons Attribution 4.0 License.

The content is available under Creative Commons Attribution 4.0 License.

сетях, безопасном управлении парковками, наблюдении за дорожным движением и его контролем. Алгоритмы распознавания автомобильных номеров помогают полиции искать пропавшие транспортные средства.

На момент написания работы, в данной области уже проведено немало исследований, и предложены различные модели, решающие задачу распознавания автомобильных номеров [7]. Однако, такие модели обычно разработаны для клиент-серверных архитектур, поскольку архитектуры моделей, хорошо решающих данную задачу, зачастую имеют сотни миллионов параметров [8] и рассчитаны для того, чтобы их обучали и применяли на машинах с большим объемом памяти и производительными видеокартами. С другой стороны, иногда бывает необходимо решать данную задачу в отсутствии интернета и высоких вычислительных мощностей [9]. Алгоритмы, способные работать в таких ограничениях, найти не удалось. То есть с одной стороны существует потребность в алгоритмах распознавания автомобильных номеров, предъявляющих ограниченные требования к вычислительным ресурсам, работающим на мобильных устройствах, с другой стороны решений на текущий момент не выявлено. Таким образом цель настоящей статьи состоит в разрешении обозначенного противоречия — улучшение распознавания номерных знаков, за счет разработки и реализации соответствующего алгоритма, способного работать в условиях ограниченных вычислительных мощностей без доступа к интернету. Теоретическая значимость результата заключается в разработке нового алгоритма, учитывающего указанные ограничения, способствующего развитию серии изысканий в области разработки методов и алгоритмов, функционирующих в условиях существенных ограничений на вычислительные мощности. Практическая значимость результата заключается в приложениях результата к решению прикладных задач, таких как фиксация неправильной парковки, иных правонарушений.

Релевантные работы. Традиционным образом задача распознавания автомобильных номеров решается путем разбиения ис-

ходного видео на кадры, где частота кадров зависит от конкретной задачи и влияет на скорость обработки видео алгоритмом [10]. После проведения раскадровки задача поиска государственного номерного знака разбивается на следующие две подзадачи. Во-первых, необходимо определить присутствует ли на данном кадре автомобиль с государственным номерным знаком и какие у него координаты — это задача выделения объектов на изображении. Во-вторых, необходимо провести предобработку таблички и её сегментацию на отдельные символы с целью последующей их трансляции в цифровой формат — это задача оптического распознавания символов. Решать эти задачи можно совершенно независимо и одновременно. Лучшие на данный момент алгоритмы совмещают оба подхода, решая задачи по отдельности, но делясь информацией между собой [6].

Существует пласт коммерческих сервисов для решения задачи распознавания автомобильного номерного знака в видеопотоке. Рассмотрим основные из них более подробно.

1. OpenALPR¹ — сервис имеет как Open-Source, так и коммерческое решение. Open-Source версия показывает низкий процент точности в распознавании и использует специфические библиотеки, а, следовательно, не может быть использована в нашем проекте. С другой стороны, коммерческая версия проекта показывает хорошие результаты за приемлемую цену и могла бы использоваться в случае наличия возможности обратиться к серверу приложения.

2. Neuro-Recognitor² — сервис с хорошим качеством распознавания номеров, показывает высокую точность даже в плохую погоду и при слабом освещении. Однако, он хорошо работает только с российскими номерами и не может быть расширен на номерные знаки других стран. Кроме того, для создателей проекта, сервис является скорее хобби, чем работой, а значит может оказаться нестабильным решением.

¹ OpenALPR. – URL: <https://www.openalpr.com/software/blue>

² Neuro-Recognitor. – URL: http://cvml.ru/?page_id=89

3. НомерОК³ — стабильный сервис, широко используемый в российских системах. Распространяется, как платное программное обеспечение, работающее без доступа к интернету, но поддерживает только платформу Windows и не дает возможности доработки для удовлетворения поставленных условий задачи.

Таким образом, на текущий момент нет комплексных сервисов, позволяющих решать задачу распознавания автомобильных номеров в условиях ограниченности вычислительных ресурсов. Напомним, что задача распознавания автомобильных номеров обычно декомпозируется на две: поиск рамки автомобильного номера и распознавание его символов. Для каждой из этих задач также существуют решения, рассмотрим их. Для обнаружения рамки автомобильного номера на снимке наиболее удачными решениями являются следующие.

1. Faster R-CNN [11]. Один из первых алгоритмов по предложению областей с вероятностями объекта на них. Является полностью сверточной нейронной сетью, использующей алгоритмы внимания, введенные в статье, посвященной архитектуре Transformer [9]. Имеет высокую скорость обнаружения объектов и обеспечивает среднюю точность в популярных наборах данных. Алгоритм полезен для изучения, но не может конкурировать с аналогами.

2. Single Shot Multibox Detector [13]. Подход данной модели основан на простой сверточной сети прямого распространения, которая создает стандартный набор ограничивающих рамок фиксированного размера и оценивает наличие экземпляров классов в этих полях. После этого, модель предлагает улучшение для рамки, чтобы больше подходить под объект на изображении. Данная модель работает быстрее и показывает большую точность, чем стандартный современный алгоритм для подобного рода задач — YOLOv1 [14] и показывает значительно большую точность, но с потерей во времени работы, в сравнении с алгоритмом Faster R-CNN (Region Based Convolutional Neural Network) [11]. Алго-

³ НомерОК. – URL: <https://avtonomerok.su/products>

ритм может быть включен в рассмотрение, но напрямую использовать его не получится, в силу повышенной требовательности к вычислительным ресурсам.

3. You only look once [14]. YOLO — современный подход к задаче обнаружения объектов на изображениях. Рассматривает обнаружение объектов как задачу регрессии к пространственно разделенным ограничивающим рамкам и связанным с ними вероятностями классов. Нейронная сеть предсказывает ограничительные рамки и вероятности классов в этих рамках с полного изображения за один проход, вследствие чего алгоритм и имеет такое название. Базовая модель работает довольно быстро, но также имеет и упрощенную версию Fast YOLO, работающую в 3 раза быстрее, но теряющую в точности. Алгоритм постоянно улучшается и сейчас уже существует YOLOv5 [15], основанный на архитектуре Transformer [12]. Прекрасно подходит для решения общей задачи и может быть легко использована в нашем проекте, хотя и имеет альтернативы, имеющие большую точность вследствие их адаптированности к задаче.

Мы рассмотрели существующие решения для поиска на изображении рамки государственного номерного знака, далее перейдем к решениям задачи оптического распознавания символов в рамке.

1. Google Tesseract [16]. Бесплатный движок для оптического распознавания символов, поддерживаемый Google. Работает с более чем ста языками сразу и может быть легко натренирован для распознавания, как других языков, так и уникальных шрифтов, необходимых в задаче. Хорошо справляется с общей задачей распознавания отдельных символов, но требует нетривиальной предобработки автомобильных номеров, поэтому для задачи распознавания символов автомобильных номеров имеет лучшие аналоги, хотя может быть использован в качестве первого решения.

2. Convolutional Recurrent Neural Network [17]. Модель использует сверточную нейронную сеть, чтобы извлечь признаки из исходного изображения, а на извлеченных признаках применить двустороннюю рекуррентную нейронную сеть, которая позволяет учиты-

вать контекст при распознавании символов и повышает качество распознавания. Хотя модель показывает высокие результаты, она сильно проигрывает по производительности моделям, не использующим архитектуру рекуррентной сети.

3. Semantic Segmentation Network [18]. Данный алгоритм состоит из двух частей. Семантическая сегментация, создающая семантическую карту исходного изображения, где каждый символ заменяется на его класс внутри данной автомобильной таблички. Затем применяется уточнение подсчета, так как некоторые символы могли слиться в один. По сравнению с традиционными методами, такой подход может использовать больше информации в дополнении к высокой эффективности, поскольку всё изображение обрабатывается за раз. В силу своей эффективности является привлекательным вариантом для использования в проекте, но проигрывает следующей архитектуре по всем параметрам.

4. Squeeze Character Recognition Network [6]. Модель проводит вырезание автомобильного номера и его нормализацию, состоящую в повороте и растягивании изображения. Затем модель применяет сверточные нейронные слои с батч нормализацией и динамической регуляризацией [19]. После этого создаются признаки изображения с применением блоков с остаточными связями [8] и на полученных признаках применяются классификаторы, ответственные за символы автомобильного номера. Данная модель значительно превосходит в своей производительности и требовательности к ресурсам предыдущие подходы [6], вследствие чего является самым привлекательным вариантом для использования в решении.

1. ПОСТРОЕНИЕ АЛГОРИТМА

Мы рассмотрели комплексные варианты решения задачи распознавания номерных знаков и варианты решения отдельных задач, на которые указанная задача может декомпозироваться. Далее сконфигурируем свой алгоритм для решения задачи распознавания

автомобильного номера в условиях ограниченности вычислительных ресурсов.

Первой задачей, которую необходимо решить, является задача идентификации рамки номерного знака автомобиля. Для решения этой задачи был выбран алгоритм YOLOv5. Так как данный алгоритм постоянно развивается и используется в различных областях компьютерного зрения для задачи распознавания объектов на изображении, он имеет широкую аудиторию людей его поддерживающих и отвечающих на возникающие во время его использования вопросы. Обученные модели экспортируются в Tensorflow Lite формат, который может быть использован на мобильном устройстве. Кроме того, официальная реализация алгоритма поддерживает сразу несколько размеров модели, начиная от малой модели для мобильных устройств и заканчивая большой моделью для серверов. На рис. 1 представлен график, сравнивающий эти модели по их производительности и эффективности, измеренной на стандартном наборе данных [20]. В табл. 1 представлена производительность некоторых архитектур модели, вычисленная на разных поколениях iPhone.

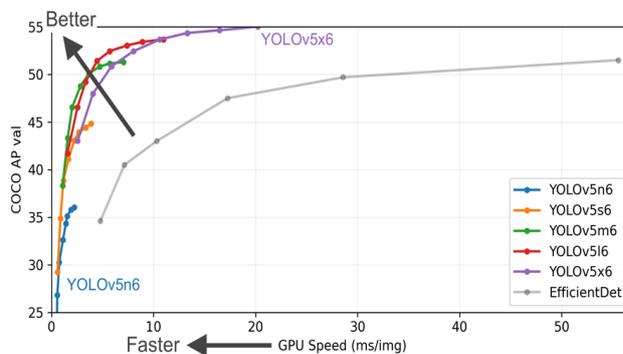


Рис. 1. Сравнение производительности моделей [20]

[Fig. 1. Model performance comparison [20]]

Так как в нашей задаче важна производительность модели, было решено начать с самой производительной из представленных — YOLOv5s. Также, в целях повышения производительности и учет низкого разрешения камер смартфонов было решено попробовать тренировать изображения с понижен-

ным разрешением: 256×256 пикселей. Следуя рекомендациям разработчиков алгоритма, модель начинала тренироваться не с нуля, а загружала предобученные веса и затем тренировалась на найденных размеченных изображениях номерных знаков стран СНГ, которые были предварительно сжаты до нужного разрешения.

Таблица 1. Сравнение производительности моделей на разных устройствах [20]
[Table 1. Comparison of model performance on different devices [20]]

Иphone	Год	Процес-сор	YOLOv5s (мс)	YOLOv5m (мс)	YOLOv5l (мс)	YOLOv5x (мс)
6	2014	A8-20mm	90	180	350	500
6s	2015	A9-16mm	148	216	304	289.8
7	2016	A10-16mm	94.7	140.7	216.7	289.8
8/X	2017	A11-10mm	-	-	-	-
XR/XS	2018	A12-7mm	22.3	25.8	43.2	57.7
11	2019	A13-7mm	17.4	21.3	27.8	41.2
12	2020	A14-5mm	14.3	16.5	21.0	28.8

Для распознавания символов на табличке автомобильного номера в силу высокой заявленной эффективности и производительности было решено использовать архитектуру нейронной сети SCR-Net. Детали архитектуры представлены в табл. 2. Классификаторы представляют собой простые полносвязные слои и призваны заменить с сохранением эффективности, но улучшением производительности рекуррентную нейронную сеть, используемую для предсказания символов в решениях с высокой точностью [11].

Таблица 2. Детальная архитектура SCR-NET [6]
[Table 2. Detailed architecture SCR-NET SCR-NET [6]]

Наименование	Размер	Ch	Описание
Input	64×256	3	Input image
S-Conv-1	32×128	16	Conv-4-BN-ReLU
S-Stage-1	32×128	93	{BN-Conv3-BN-ReLU-Conv3-BN-Dy}×4
S-Stage-2	16×64	176	{BN-Conv3-BN-ReLU-Conv3-BN-Dy}×4
S-Stage-3	8×32	256	{BN-Conv3-BN-ReLU-Conv3-BN-Dy}×4
S-Conv-2	1×32	256	Conv3-BN-ReLU-Conv8×1-BN-ReLU
S-Stage-4	1×32	256	{{Res-block}×2-Conv1×3-BN-ReLU}×2
Classifiers	1×1	N	N выходной размер классификатора.

- Ch — количество каналов в итоговом изображении;
- Conv4 — стандартный свёрточный блок с размером фильтра 4×4 и шагом 2;
- Conv8×1 — стандартный свёрточный блок с фильтром размера 8×1;
- Conv1×3 — стандартный свёрточный блок с размером фильтра 1×3;
- BN-батч нормализация [21];
- ReLU — Rectified Linear Unit [22];
- Dy — динамическая регуляризация параметров [19];
- {...}×n — означает применение структуры n раз.

Для предотвращения переобучения модели и подготовки ее к плохим снимкам автомобильных номеров также был применен SpatialDropout2D [23] и различные виды аугментации изображений, включая растяжения, повороты, зум и изменение яркости.

Объединение моделей. Для построения полноценного решения стоит заметить, что

необходимо только из предсказанной моделью YOLOv5pb прямоугольной области выделить четырёхугольник с номерным знаком и преобразовать его в подходящий прямоугольник для модели SCR-Net. Для этого можно использовать известное в компьютерном зрении преобразование Хаффа для выделения прямых линий на изображении и последующего его поворота, но он работает эффективно только для знаков хорошего качества с небольшим поворотом. Для решения данной проблемы был предложен следующий подход. Вырезанный из изображения прямоугольник обычно имел размер в десятки пикселей, так как исходное изображение было сжато, поэтому он вырезался из исходного изображения и сжимался до прямоугольника размера 64×256 для однородности данных. После этого предлагалось повысить яркость и контрастность изображений для упрощения выделения контуров. Так как обычно внутри прямоугольника из объектов была только автомобильная табличка, то предлагалось использовать простую сверточную нейросеть для определения углов таблички и ее последующего поворота.

Полное решение задачи следующее:

1. На вход нашему алгоритму подается видеоряд, который разбивается на кадры с подобранной частотой, удовлетворяющей запросам пользователя библиотеки;

2. Для каждого кадра видео применялся следующий алгоритм действий:

а. Если предобученная модель YOLOv5s не нашла на кадре ни одного автомобильного номера, тогда кадр пропускался и алгоритм возвращал пустой список;

б. Если модель нашла какие-либо автомобильные номера, то для каждого из них происходила предобработка описанной выше нейросетью с целью его нормализации и подача на вход предобученной сети SCR-Net, которая распознает на нём автомобильный номер и возвращает его.

3. В результате работы алгоритма, для каждого кадра видео с указанной частотой он возвращает список автомобильных номеров и их координаты на конкретном кадре.

Данные. Было найдено несколько наборов данных, которые могут быть использованы для тестирования полученного алгоритма:

1. CLPD (Chinese License Plate Dataset)⁴ — набор данных, состоящий из фотографий машин с китайскими автомобильными номерами в разную погоду с разным освещением.

2. AOLP (Application-Oriented License Plate Recognition)⁵ — набор данных, содержащих выборку из машин преимущественно с европейскими номерами.

3. CCPD (Chinese City Parking Dataset)⁶ — большой набор данных, представляющий собой размеченные фотографии машин на парковках Китая.

4. Nomeroff Dataset⁷ — размеченный набор данных, содержащий фотографии машин как из стран СНГ, так и из стран Европы. Кроме того, содержит отдельно вырезанные автомобильные номера машин с разметкой и символы номеров с разметкой.

Так как модель должна быть обучена для её применения прежде всего в странах СНГ, то для решения задачи распознавания автомобильных номеров подходит только последний набор данных, ибо только он содержит соответствующие размеченные данные. Тем не менее, для задачи определения положения автомобильных номеров подходят все вышеперечисленные наборы данных.

2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Модель YOLOv5s обучалась на найденном датасете в течение 100 эпох с размером батча в 128 изображений. Обучение модели происходило на GPU на протяжении двух часов. Изменения функции ошибки и метрик на протяжении эпох представлены на рис. 2. На рисунке используются следующие обозначения:

⁴ China License Plate Dataset. – URL: https://github.com/wangpengnorman/CLPD_dataset

⁵ Application-Oriented License Plate Recognition. – URL: <https://github.com/AvLab-CV/AOLP>

⁶ Chinese City Parking Dataset. – URL: <https://github.com/detectRecog/CCPD>

⁷ Nomeroff Datasets. – URL: <https://nomeroff.net.ua/datasets/>

- `box_loss` — среднеквадратичная ошибка от предсказаний настоящих ширины, высоты, x-координаты левого угла, y-координаты левого угла;
- `obj_loss` — перекрестная энтропия для предсказанных объектов и настоящих объектов (функция ошибки для предсказания категорий объектов);
- `loss_cls` — ошибка предсказаний прямоугольника с объектом;
- `precision` — стандартная определенность, обозначающая долю автомобильных табличек среди всех объектов, классифицированных как таблички;
- `recall` — чувствительность модели, обозначает какую долю от всех автомобильных табличек распознает модель;
- `mAP_0.5` — точность при пороге уверенности 0.5;
- `mAP_0.5:0.9` — средняя точность при порогах уверенности от 0.5 до 0.9.

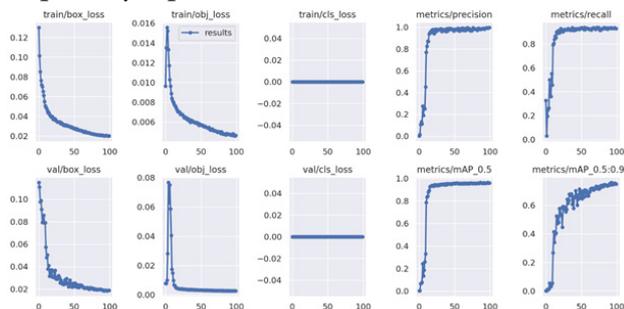


Рис. 2. Результаты обучения модели YOLOv5s [Fig. 2. Model training results YOLOv5s]

Для модели с лучшими весами среди всех эпох были построены графики для метрик Precision на рис. 3, Recall на рис. 4.

В результате оценки модели с лучшими весами была получена точность — 0.95, чувствительность — 0.93, определенность — 0.99. То есть модель распознала 93 % всех номерных знаков на фотографиях машин и, если модель что-то распознала, то это с точностью 99 % будет действительно номерным знаком.

Модель SCR-Net обучалась на протяжении 100 эпох и размером батча из 128 изображений. Обучение модели происходило на GPU на протяжении пяти часов. В результате, лучшая получившаяся модель имеет точность

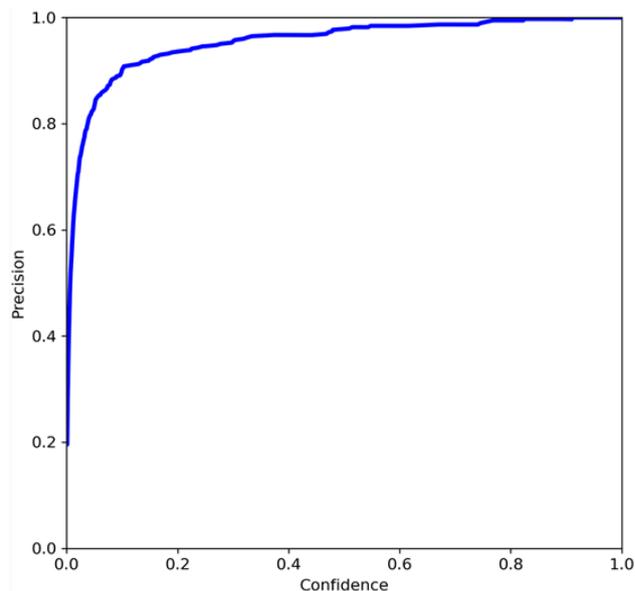


Рис. 3. YOLOv5s Precision [Fig. 3. YOLOv5s Precision]

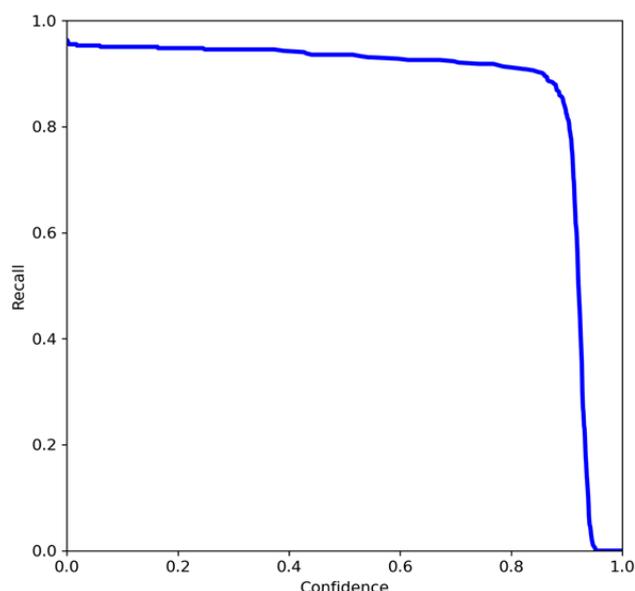


Рис. 4. YOLOv5s Recall [Fig. 4. YOLOv5s Recall]

0.7, если считать точность через количество правильно классифицированных символов и точность 0.15, если считать точность только тогда, когда модель полностью распознала всю автомобильную табличку.

Предложенные модели демонстрируют хорошие результаты для поиска табличек номерных знаков в видеоряде, но дают недостаточные по качеству результаты для распознавания номерного знака. Вместе с тем алгоритм полностью адаптирован для работы

на смартфоне. Необходимая точность может быть достигнута в ходе дообучения моделей, реконфигурации или усложнения используемых методов.

ЗАКЛЮЧЕНИЕ

Таким образом, в статье предложен алгоритм автоматического распознавания автомобильного номерного знака в видеоряде низкого качества, адаптированный для работы в условиях ограничения вычислительных ресурсов. Теоретическая значимость результата заключается в разработке нового алгоритма, учитывающего указанные ограничения, способствующего развитию серии изысканий в области разработки методов и алгоритмов, функционирующих в условиях существенных ограничений на вычислительные мощности. Практическая значимость результата заключается в приложениях результата к решению прикладных задач, таких как фиксация неправильной парковки, иных правонарушений. Дальнейшие направления исследований будут связаны с повышением точности распознавания на найденной табличке государственного номерного знака.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Geirhos R., Janssen D. H. J., Schütt H. H., Rauber J., Bethge M., Wichmann F. A. Comparing deep neural networks against humans: object recognition when the signal gets weaker // arXiv preprint arXiv:1706.06969. – 2017. – URL: <https://arxiv.org/pdf/1706.06969.pdf>.

2. Олисеенко В. Д., Абрамов М. В., Тулупьев А. Л. Идентификация аккаунтов пользователей при помощи сравнения изображений: подход на основе рHash // Научно-технический вестник информационных технологий, механики и оптики. – 2021. – Т. 21, № 4. –

С. 562–570. doi: 10.17586/2226-1494-2021-21-4-562-570

3. Корепанова А. А., Абрамов М. В., Тулупьев А. Л. Идентификация аккаунтов пользователей социальных сетей при помощи сравнения графического контента // Научно-технический вестник информационных технологий, механики и оптики. – 2021. – Т. 21, № 6. – С. 942–950. doi: 10.17586/2226-1494-2021-21-6-942-950

4. Oliseenko V. D., Abramov M. V. identification of user profiles in online social networks: a combined approach with face recognition // Journal of Physics: Conference Series. – 2021. – Vol. 1864. 012119. doi: 10.1088/1742-6596/1864/1/012119

5. Shahdadpuri N. Real Image of Computer Vision Application and its Impact: Future and Challenges. -- URL: https://www.researchgate.net/publication/349278277_Real_Image_of_Computer_Vision_Application_and_its_Impact_Future_and_Challenges

6. Wang Y., Bian Z.-P., Zhou Y., Chau L.-P. Rethinking and designing a high-performing automatic license plate recognition approach // IEEE Transactions on Intelligent Transportation Systems. – 2021. – URL: <https://arxiv.org/pdf/2011.14936.pdf>.

7. Du S., Ibrahim M., Shehata M. Automatic license plate recognition (ALPR): A state-of-the-art review // IEEE Transactions on circuits and systems for video technology. – 2012. – Т. 23, № 2. – С. 311–325. -- URL: <https://ieeexplore.ieee.org/document/6213519>.

8. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 770–778. – URL: <https://arxiv.org/abs/1512.03385>.

9. Saha S., Basu S., Nasipuri M., Basu D. K. Localization of License Plates from Indian Vehicle Images Using Iterative Edge Map Generation. – 2011. -- URL: <https://arxiv.org/abs/1003.1072>.

10. Laroca R., Severo E., Zanlorensi L. A., Oliveira L. S., Goncalves G. R., Schwartz W. R., Menotti D. A robust real-time automatic license plate recognition based on the YOLO detector // 2018 international joint conference on neural

- networks (ijcnn). – IEEE, 2018. – С. 1–10. – URL: <https://arxiv.org/pdf/1802.09567.pdf>
11. Ren S., He K., Girshick R., Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks // Advances in neural information processing systems. – 2015. – Т. 28. – URL: <https://arxiv.org/abs/1506.01497>.
 12. Vaswani A., Shazeer A., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. Attention is all you need // Advances in neural information processing systems. – 2017. – Т. 30. – URL: <https://arxiv.org/abs/1706.03762>.
 13. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A. C. Ssd: Single shot multi-box detector // European conference on computer vision. – Springer, Cham, 2016. – С. 21–37. – URL: <https://arxiv.org/abs/1512.02325>.
 14. Redmon J., Divvala S., Girshick R., Farhadi A. You only look once: Unified, real-time object detection // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 779–788. – URL: <https://arxiv.org/abs/1506.02640>.
 15. Zhu X., Lyu S., Wang X., Zhao Q. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios // Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2021. – С. 2778–2788. – URL: <https://arxiv.org/abs/2108.11539>.
 16. Smith R. An overview of the Tesseract OCR engine // Ninth international conference on document analysis and recognition (ICDAR 2007). – IEEE, 2007. – Т. 2. – С. 629–633 – URL: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/33418.pdf>.
 17. Wang W., Yang J., Chen M., Wang P. A light CNN for end-to-end car license plates detection and recognition // IEEE Access. – 2019. – Т. 7. – С. 173875–173883. – URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8915848>.
 18. Zhuang J., Hou S., Wang Z., Zha Z.-J. Towards human-level license plate recognition // Proceedings of the European Conference on Computer Vision (ECCV). – 2018. – С. 306–321. – URL: http://home.ustc.edu.cn/~saihui/papers/eccv2018_towards.pdf.
 19. Wang Y., Bian Z.-P., Hou J., Chau L.-P. Convolutional neural networks with dynamic regularization // IEEE Transactions on Neural Networks and Learning Systems. – 2020. – Т. 32, № 5. – С. 2299–2304. – URL: <https://arxiv.org/pdf/1909.11862.pdf>.
 20. YOLOv5 Github. -- URL: <https://github.com/ultralytics/yolov5>.
 21. Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift // International conference on machine learning. – PMLR, 2015. – С. 448–456. -- URL: <https://arxiv.org/pdf/1502.03167.pdf>
 22. Agarap A. F. Deep learning using rectified linear units (relu) // arXiv preprint arXiv:1803.08375. – 2018. – URL: <https://arxiv.org/pdf/1803.08375.pdf>.
 23. Tompson J., Goroshin R., Jain A., LeCun Y., Bregler C. Efficient object localization using convolutional networks // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2015. – P. 648–656. — URL: https://openaccess.thecvf.com/content_cvpr_2015/html/Tompson_Efficient_Object_Localization_2015_CVPR_paper.html.
17. Wang W., Yang J., Chen M., Wang P. A light CNN for end-to-end car license plates de-

Абрамов Максим Викторович — канд. техн. наук, старший научный сотрудник, руководитель лаборатории теоретических и междисциплинарных проблем информатики, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук; доцент кафедры информатики, Санкт-Петербургский государственный университет.
E-mail: mva@dscs.pro
ORCID iD: <https://orcid.org/0000-0002-5476-3025>

Ельцов Даниил Андреевич — инженер по исследованиям и разработкам в команде облачных игр, Huawei Technologies Co. Ltd. E-mail: sthfaceless@mail.ru
ORCID iD: <https://orcid.org/0000-0002-9048-8585>

NEURAL NETWORKS WITH LSTM AND GRU IN APPLICATION TO THE TASK OF MULTICLASS CLASSIFICATION OF TEXT POSTS OF SOCIAL NETWORK USERS

© 2023 M. V. Abramov^{1✉}, D. A. Yeltsov²

¹*St. Petersburg Federal Research Center of the Russian Academy of Sciences
39, 14th Line V.O., 199178 St. Petersburg, Russian Federation*

²*Huawei Technologies Co. Ltd.
12, Tambovskaya Street, 192007 St. Petersburg, Russian Federation*

Annotation. A lot of research has already been done in this area, and various models have been proposed that solve the problem of license plate recognition, there are industrial designs used on cameras for recording traffic violations. However, such models are usually developed for client-server architectures, since model architectures that solve this problem well often have hundreds of millions of parameters [5] and are designed to be trained and applied on machines with large memory and productive video cards. The subject of the study of this article is the automobile state license plates in the video sequence of poor quality, recognizable under conditions of limited computing resources. To recognize license plates, it is proposed to use the following methods: simple convolutional neural networks to determine the angles of a license plate and its subsequent rotation; a pretrained YOLOv5s model for searching license plates in a frame and a pretrained SCR-Net network for recognizing license plates on plates. The article proposes an algorithm and its implementation for automatic license plate recognition under conditions of poor video quality and limited computing resources. The theoretical significance of the result lies in the development of a new algorithm that takes into account these limitations, which contributes to the development of a series of studies in the field of developing methods and algorithms that operate under conditions of significant limitations on computing power and video quality. The practical significance of the result lies in the application of the result to solving applied problems, such as fixing incorrect parking, searching for stolen cars, and other offenses.

Keywords: License Plate Recognition, Detection Algorithms, Smartphone Recognition, Number Plate Recognition.

CONFLICT OF INTEREST

The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

REFERENCES

1. Geirhos R., Janssen D. H. J., Schütt H. H., Rauber J., Bethge M., Wichmann F. A. (2017) Comparing deep neural networks against humans: object recognition when the signal gets weaker // arXiv preprint arXiv:1706.06969. URL: <https://arxiv.org/pdf/1706.06969.pdf>.

2. Олисеенко В. Д., Абрамов М. В., Тулупьев А. Л. (2021) Идентификация аккаунтов пользователей при помощи сравнения изображений: подход на основе рHash. *Научно-технический вестник информационных технологий, механики и оптики*. Т. 21, No 4. P. 562–570. doi: 10.17586/2226-1494-2021-21-4-562-570

3. Корепанова А. А., Абрамов М. В., Тулупьев А. Л. (2021) Идентификация аккаунтов пользователей социальных сетей при помощи сравнения графического контента. *Научно-технический вестник информационных технологий, механики и оптики*. Т. 21, No 6. P. 942–950. doi: 10.17586/2226-1494-2021-21-6-942-950

4. Oliseenko V. D., Abramov M. V. (2021) Identification of user profiles in online social networks: a combined approach with face recognition.

✉ Abramov Maxim V.
e-mail: mva@dscs.pro

Journal of Physics: Conference Series. Vol. 1864. 012119. doi: 10.1088/1742-6596/1864/1/012119

5. Shahdadpuri N. Real Image of Computer Vision Application and its Impact: Future and Challenges. URL: https://www.researchgate.net/publication/349278277_Real_Image_of_Computer_Vision_Application_and_its_Impact_Future_and_Challenges

6. Wang Y., Bian Z.-P., Zhou Y., Chau L.-P. (2021) Rethinking and designing a high-performing automatic license plate recognition approach. *IEEE Transactions on Intelligent Transportation Systems*. URL: <https://arxiv.org/pdf/2011.14936.pdf>.

7. Du S., Ibrahim M., Shehata M. (2012) Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology*. T. 23, No 2. P. 311–325. URL: <https://ieeexplore.ieee.org/document/6213519>.

8. He K., Zhang X., Ren S., Sun J. (2016) Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. P. 770–778. URL: <https://arxiv.org/abs/1512.03385>.

9. Saha S., Basu S., Nasipuri M., Basu D. K. (2011) Localization of License Plates from Indian Vehicle Images Using Iterative Edge Map Generation. URL: <https://arxiv.org/abs/1003.1072>.

10. Laroca R., Severo E., Zanlorensi L. A., Oliveira L. S., Goncalves G. R., Schwartz W. R., Menotti D. (2018) A robust real-time automatic license plate recognition based on the YOLO detector. *2018 international joint conference on neural networks (ijcnn)*. *IEEE*. P. 1–10. URL: <https://arxiv.org/pdf/1802.09567.pdf>

11. Ren S., He K., Girshick R., Sun J. (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*. T. 28. URL: <https://arxiv.org/abs/1506.01497>.

12. Vaswani A., Shazeer A., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. (2017) Attention is all you need. *Advances in neural information processing systems*. T. 30. URL: <https://arxiv.org/abs/1706.03762>.

13. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A. C. (2016) Ssd: Single

shot multibox detector. *European conference on computer vision*. Springer, Cham. P. 21–37. URL: <https://arxiv.org/abs/1512.02325>.

14. Redmon J., Divvala S., Girshick R., Farhadi A. (2016) You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*. P. 779–788. URL: <https://arxiv.org/abs/1506.02640>.

15. Zhu X., Lyu S., Wang X., Zhao Q. (2021) TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. P. 2778–2788. URL: <https://arxiv.org/abs/2108.11539>.

16. Smith R. (2007) An overview of the Tesseract OCR engine. *Ninth international conference on document analysis and recognition (ICDAR 2007)*. *IEEE*. T. 2. P. 629–633. URL: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/33418.pdf>.

17. Wang W., Yang J., Chen M., Wang P. (2019) A light CNN for end-to-end car license plates detection and recognition. *IEEE Access*. T. 7. P. 173875–173883. – URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8915848>.

18. Zhuang J., Hou S., Wang Z., Zha Z.-J. (2018) Towards human-level license plate recognition. *Proceedings of the European Conference on Computer Vision (ECCV)*. P. 306–321. URL: http://home.ustc.edu.cn/~saihui/papers/eccv2018_towards.pdf.

19. Wang Y., Bian Z.-P., Hou J., Chau L.-P. (2020) Convolutional neural networks with dynamic regularization. *IEEE Transactions on Neural Networks and Learning Systems*. T. 32, No 5. P. 2299–2304. URL: <https://arxiv.org/pdf/1909.11862.pdf>.

20. YOLOv5 Github. URL: <https://github.com/ultralytics/yolov5>.

21. Ioffe S., Szegedy C. (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning*. PMLR, P. 448–456. URL: <https://arxiv.org/pdf/1502.03167.pdf>

22. Agarap A. F. (2018) Deep learning using rectified linear units (relu) // arXiv pre-

print arXiv:1803.08375. URL: <https://arxiv.org/pdf/1803.08375.pdf>.

23. *Tompson J., Goroshin R., Jain A., LeCun Y., Bregler C.* (2015) Efficient object localization using convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. P. 648–656. URL: https://openaccess.thecvf.com/content_cvpr_2015/html/Tompson_Efficient_Object_Localization_2015_CVPR_paper.html.

Abramov Maxim V. — PhD, senior researcher, is a Head of Theoretical and Interdisciplinary Computer Science Lab St. Petersburg Federal Research Center of the Russian Academy of Sciences. Associate professor at the Informatics Department, St. Petersburg State University.

E-mail: mva@dscs.pro

ORCID iD: <https://orcid.org/0000-0002-5476-3025>

Yeltsov Daniil A. — R&D Engineer в Cloud Gaming Team, Huawei Technologies Co. Ltd.

E-mail: sthfaceless@mail.ru

ORCID iD: <https://orcid.org/0000-0002-9048-8585>