ГИАБ. Горный информационно-аналитический бюллетень /
MIAB. Mining Informational and Analytical Bulletin, 2023;(6):5-18
ОРИГИНАЛЬНАЯ СТАТЬЯ / ORIGINAL PAPER

УДК 622.6:004.93 DOI: 10.25018/0236_1493_2023_6_0_5

ЭФФЕКТИВНЫЙ АЛГОРИТМ ИДЕНТИФИКАЦИИ ТРАНСПОРТНЫХ СРЕДСТВ В СИСТЕМАХ ВИДЕОНАБЛЮДЕНИЯ

В.А. Епифанов[1,2], И.О. Темкин[1], С.Е. Краснояружский[2]

[1] НИТУ «МИСиС», Москва, Россия, e-mail: igortemkin@yandex.ru
[2] ООО «Ситилабс», Москва, Россия

Аннотация: Рассмотрена задача, связанная с повышением точности идентификации транспортных средств с использованием стационарных или мобильных видеокамер. Плохое качество изображений или их частей из-за различных внешних или внутренних факторов может приводить к потере значимой информации. К факторам, негативно влияющим на качество итоговых изображений, можно причислить избыточную яркость объектов кадра, высокую скорость перемещения объектов, недостаточные возможности камеры для съемки в темных или избыточно ярких условиях. Яркие источники света, например, автомобильные фары, подают избыточное количество света на матрицу камеры и делают некоторые части изображений неразличимыми, а высокая скорость объектов в кадре при недостаточном значении выдержки часто приводит к тому, что за объектом следует некоторый шлейф. Такие факторы приводят к тому, что интересующие объекты на изображении могут стать непригодными для анализа. Особенно важным это становится, когда необходимо точно определить регистрационный номер автомобиля. Высокая степень смазанности или засвеченности может привести к тому, что различные идентификационные метки, такие, как государственные регистрационные знаки (ГРЗ) автомобилей, могут быть нечитаемыми и непригодными для распознавания различными системами. Подобные проблемы характерны также для некоторых горнопромышленных объектов, где ведется стационарное видеонаблюдение за производственными процессами. Данная работа посвящена вопросам автоматического определения степени смазанности и засвеченности изображений государственных регистрационных знаков автомобилей. Предлагаются методы для быстрого определения смазанности и засвеченности изображений на примере ГРЗ, применимые для встраивания в различные системы на базе платформы микрокомпьютера Nvidia Jetson Nano, без существенных задержек системы. Данный подход, помимо классификации на читаемые и нечитаемые изображения, дает также численную оценку смазанности и засвеченности. Эти данные в свою очередь могут быть использованы для корректировки параметров камеры, таких, как значение выдержки и диафрагмы, что позволит повысить качество последующих кадров.

Ключевые слова: анализ изображений, идентификация транспортного средства, гистограммы, алгоритм, сверточные нейронные сети, Nvidia Jetson Nano.

Для цитирования: *Епифанов В. А., Темкин И. О., Краснояружский С. Е.* Эффективный алгоритм идентификации транспортных средств в системах видеонаблюдения // Горный информационно-аналитический бюллетень. – 2023. – № 6. – С. 5–18. DOI: 10.25018/0236_1493_2023_6_0_5.

© В.А. Епифанов, И.О. Темкин, С.Е. Краснояружский. 2023.

Efficient algorithm of transport identification in video surveillance systems

V.A. Epifanov[1,2], I.O. Temkin[1], S.E. Krasnoyaruzhskiy[2]

[1] National University of Science and Technology «MISiS», Moscow, Russia,
e-mail: igortemkin@yandex.ru
[2] «Citylabs» LLC, Moscow, Russia

Abstract: The objective is to enhance transport identification accuracy by stationary and movable video cameras. The poor quality of images or their parts due to some external or internal impacts can lead to losses of important information. The adverse effects to affect the picture quality include the excessive scene brightness, high velocity of transport and insufficient capabilities of cameras to shoot in the dark or in the excessive brightness. The bright sources of light, for instance, automobile headlights, feed the excessive light to the camera matrix and make some parts of images indistinguishable, while high velocity of vehicles produces a kind of tail behind a vehicle in the image when the exposure time is insufficient. As a result, the test subjects in the images become unsuitable for the analysis. For instance, it becomes sometimes impossible to read precisely the figures on the license plate of a car. High diffuseness or light-shorting can make the vehicle registration plates of Russia unreadable and undistinguishable for any identification system. This problem is similarly intrinsic to mining facilities with stationary video surveillance of production processes. This article focuses on automated estimation of the diffuseness or light-shorting rate in the images of the vehicle registration plates. The methods are proposed for the quick estimate of the diffuseness or light-shorting rate as a case-study of the vehicle registration plates. These methods are suitable for embedding in different systems on the basis of NVIDIA Jetson Nano Developer Kit platform without significant delays in operation. Alongside with distinguishing between the readable and unreadable images, this method enables quantification of diffuseness and light-shorting. The obtained results can assist in adjustment of such camera parameters as exposure and diaphragm, which can improve the quality of snaps.

Key words: image analysis, transport identification, bar charts, algorithm, convolutional neutral networks, Nvidia Jetson Nano.

For citation: Epifanov V. A., Temkin I. O., Krasnoyaruzhskiy S. E. Efficient algorithm of transport identification in video surveillance systems. *MIAB. Mining Inf. Anal. Bull.* 2023;(6):5-18. [In Russ]. DOI: 10.25018/0236_1493_2023_6_0_5.

Введение

Одной из важных задач, возникающих при анализе дорожно-транспортных ситуаций, в том числе и в условиях технологических дорог, является задача идентификации конкретного транспортного средства (ТС). Идентификация ТС в большинстве случаев проводится по различным опознавательным знакам, таким, как государственный регистрационный знак (ГРЗ). Во время детектирования этих знаков на видео при нестабильных условиях съемки часто возникают случаи искажения изображений, что в свою очередь приводит к некорректной идентификации. К факторам, негативно влияющим на качество итоговых изображений, можно причислить избыточную яркость

объектов кадра, высокую скорость перемещения объектов, недостаточные возможности камеры для съемки в темных или избыточно ярких условиях. Яркие источники света — это, например, автомобильные фары, которые подают избыточное количество света на матрицу камеры и делают некоторые части изображений неразличимыми. Высокая скорость объектов в кадре при недостаточном значении выдержки часто приводит к тому, что за объектом следует некоторый шлейф. Также причиной низкого качества изображений могут стать специфические особенности места съемки, например, повышенная запыленность. Такие факторы приводят к тому, что интересующие объекты на изображении могут стать непригодными для анализа. Сегодня в мире наблюдается очевидный тренд в сторону ускорения разработки беспилотных транспортных средств и использования автономных мобильных устройств в социотехнических и индустриальных системах различного назначения [1]. Работоспособность и эффективность подобных систем зависит от решения комплекса функциональных задач, среди которых одними из важнейших являются оперативное определение местоположения и динамики перемещения транспортных средств, что позволяет интерпретировать и оценивать реальную ситуацию, складывающуюся в условиях городской или индустриальной инфраструктуры, а также уточнение геопространственных и качественных параметров этой инфраструктуры. В последние годы предложены различные платформенные решения этих задач [2—4], которые основаны на комплексном использовании различных технологий, таких как: навигационные GPS системы, системы дистанционного контроля (радары, лидары и т.д.), а также стационарные и мобильные системы технического зрения, в основе работы которых использование различного типа «умных» камер. Каждый из этих подходов к идентификации дорожно-транспортных или транспортно-технологических ситуаций имеет свои достоинства и недостатки. В частности, в промышленных приложениях чаще всего используется GPS навигация в сочетании с лидарными системами, а при анализе транспортных ситуаций в городских условиях в качестве основного инструмента выступают системы видеонаблюдения [5]. На горных предприятиях системы стационарного видеонаблюдения также находят достаточно широкое применение, однако идентификация самосвалов на основе анализа видеокадров может быть актуальна только для небольших карьеров, в которых осуществляется добыча нерудных строительных материалов (щебень, песок, гравий). В этом случае при контроле за въездом, выездом и перемещением карьерного автотранспорта (в том числе, по прилегающим дорогам общего назначения) возможно использование механизмов идентификации на основе распознавания регистрационных номеров.

При решении задачи идентификации транспортного средства на основе распознавания автомобильных номеров [6—8] традиционно предъявляются жесткие требования к качеству входного изображения номера, так как ошибка даже в одном символе может привести к некорректности итогового результата работы системы. По этой причине важно быть уверенным, что номер, подающийся в систему распознавания будет с высокой вероятностью распознан верно. Для повышения этой уверенности необходимы инструменты оценки качества входных изображений.

В связи с этим определение качества входных изображений в системах распознавания автомобильных номеров является одной из важнейших задач, решение которой позволяет отфильтровывать

их непосредственно перед работой системы распознавания, которая в случае подачи ей некорректного номера после затрат значительных вычислительных ресурсов выдаст заведомо ложные результаты.

В настоящее время для распознавания государственных регистрационных знаков в системах видеонаблюдения применяются различные инструментальные технологии, среди которых наиболее популярными являются подходы на основе сверточных нейронных сетей различных архитектур, таких, как GRCNN [9, 10], ViTSTR [11]. Данные нейронные сети способны с высокой точностью распознавать текст на изображениях. При этом они чувствительны к качеству изображения, и в случае некачественного изображения такие нейронные сети могут давать ошибки в распознавании [12, 13]. Например, путать различные символы, а в случае сильного искажения изображения порой и «додумывать» результаты.

В настоящей работе основное внимание уделено оценке качества изображения по двум характеристикам: степени засвеченности изображения и степени смазанности изображения. Основной причиной засвеченности изображения, как правило, является попадание избыточного количество света на матрицу камеры [14 — 16]. Это может происходить по таким, например, причинам: высокая чувствительность матрицы, свет на матрицу поступал дольше, чем нужно, в кадре присутствует объект с избыточной яркостью. Смазанность изображения в свою очередь происходит, как правило, из-за высокой скорости объектов в кадре или высокой запыленности среды наблюдения. И, как правило, для того чтобы минимизировать этот эффект, необходимо корректировать значение выдержки. При этом важно соблюдать баланс между «пересвеченностью» и «затемненностью».

Помимо экономии вычислительных ресурсов, информация о степени засвеченности и смазанности может быть полезна для использования в автоматической подстройке параметров камеры (выдержка и диафрагма), что в результате положительно повлияет на общее качество работы системы распознавания. Для реализации данного функционала недостаточно классификации изображения на 2 класса, необходимо численное значение степени засвеченности и смазанности. Для повышения стабильности распознавания можно использовать различные техники улучшения изображений с применением различных фильтров [17] и нейросетевых технологий [18]. Такие подходы часто дают повышение качества итогового изображения, но зачастую требуют существенных вычислительных ресурсов и применяются, как правило, ко всем изображениям видеопотока. Корректировка же параметров камеры требуется нечасто, в основном в моменты изменения внешних условий съемки.

В работе рассматривается решение задачи распознавания номерных знаков, которая производится непосредственно на микрокомпьютере Nvidia Jetson Nano. Ресурсы данного устройства весьма ограничены, при этом на практике требуется возможность работы системы в реальном времени. В системах подобного рода основная вычислительная нагрузка создается модулями детектирования объектов [19] и распознавания текста, которые потребляют большую часть вычислительных ресурсов устройства, и на какие-либо дополнительные инструменты вычислительных мощностей остается крайне мало.

По этой причине при реализации решения задачи определения качества входного изображения должно быть задействовано минимальное количество вычислительных ресурсов.

Определение степени засвеченности

Для определения степени засвеченности автомобильного номера нами предлагается использовать анализ гистограммы яркостей. В большинстве случаев при визуальном анализе регистрационных автомобильных номеров общая характеристика изображения может сильно отличаться от конкретного участка, содержащего сам номер. По этой причине в случае работы системы необходимо полагаться лишь на части изображений, содержащие номера. Для этого детектирование положения номера производится поэтапно (рис. 1).

Разработанный нами алгоритм можно представить, как многоэтапную процедуру:

1. Сначала на кадре видео производится поиск транспортных средств. После обнаружения в заданной области на кадре транспортного средства определяются его точные координаты на изображении.

2. Далее на уже вырезанном участке, содержащем автомобиль, ищется его регистрационный знак. Такой двухэтапный алгоритм поиска позволяет повысить уверенность в том, что обнаруженная область действительно является номером.

Это позволяет отсечь случаи, когда якобы автомобильный номер находится в области без автомобиля. Для детектирования как транспортных средств, так и автомобильных номеров используется хорошо известная нейронная сеть yolov5. Для детектирования автомобилей и номеров при обучении нейронных сетей было сформировано и размечено 2 датасета. Данные датасеты формировались с учетом времени суток, сезонности и погоды. В результате обучения нейронных сетей удалось добиться стабильных результатов работы.

При этом полученные нейросетевые модели инвариантны относительно времени суток, сезонов и особых погодных условий.

Рис. 1. Этапы поиска номера
Fig. 1. Stages of the number search

Рис. 2. Определение информативной области номера
Fig. 2. Determination of informative area of number

После определения координат номерного знака на изображении для определения степени засвеченности необходимо выделить центральную его часть. Для этого отсекается по 20% с каждого края изображения номера (рис. 2).

Это необходимо для того, чтобы в исследуемую область не попадали фрагменты, не принадлежащие знаку. Такая ситуация возможна, когда детектируемые знаки захватываются под некоторым углом, а вхождение частей, не принадлежащих номеру, в область анализа отрицательно влияет на итоговый результат. После определения области, представляющей интерес, выбранный участок из трехмерного цветового пространства RGB преобразуется к одномерному «серому» по формуле:

$$Gray = 0{,}299 * R + 0{,}587 * G + 0{,}114 * B. \quad (1)$$

Результат преобразования цветного изображения номера к серому (рис. 3).

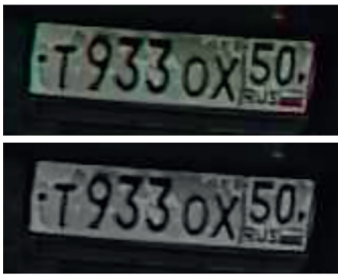

Рис. 3. Преобразование изображения из цветного в серое
Fig. 3. Converting an image from color to gray

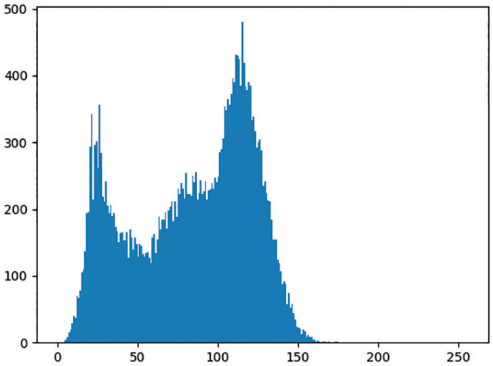

Рис. 4. Гистограмма яркости
Fig. 4. Histogram of brightness

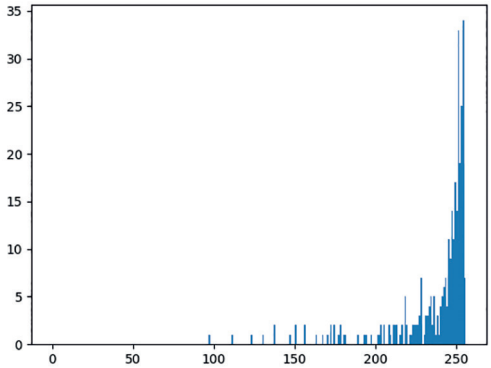

Рис. 5. Гистограмма «пересвеченного» номерного знака
Fig. 5. Histogram of the overexposed license plate

«Серое» изображение далее используется для расчета гистограммы яркостей (рис. 4). Данная гистограмма строится на основе всех пикселей полученной области интереса. Каждый пиксель «серого» изображения имеет значение от 0 (абсолютно черный) до 255 (абсолютно белый). Для каждого бина гистограммы считается сумма всех пикселей области, значение которых принадлежит данному бину. Так, например, значение 5-го бина равно количеству всех пикселей со значением 5, а значение 31-го бина равно количеству всех пикселей со значением 31.

После подсчета гистограммы яркостей серого изображения выделяется та ее часть, которая будет отвечать за «пересвеченность». На основе проведенных вычислительных экспериментов было выявлено, что оптимально использовать пятую часть гистограммы, то есть бины от 204 до 255. Это связано с тем, что в случае «пересвеченности» изображения более 90% гистограммы приходится как раз на этот участок. Пример гистограммы «пересвеченного» изображения номера автомобиля представлен на рис. 5.

Далее для оценки засвеченности изображения необходимо посчитать отношение количества пикселей из бинов, отвечающих за засвеченность, то есть бины от 204 до 255, ко всем пикселям участка изображения:

$$O = \frac{\sum_{i=204}^{255} H_i}{I_w * I_h}, \qquad (2)$$

где H — гистограмма изображения; I_w — ширина изображения; I_h — высота изображения.

Значение засвеченности, исходя из (2), может принимать значение от 0 до 100, а для того чтобы определить порог, при пересечении которого номер считается пересвеченным, нами был собран датасет, состоящий из номеров с разной степенью засветки (рис. 6).

Для каждого номера из данного датасета расчитывалось значение функции засвета. Далее все номера разделялись на

Рис. 6. Автомобильные номера с разной степенью засвеченности
Fig. 6. License plates with different degrees of overexposure

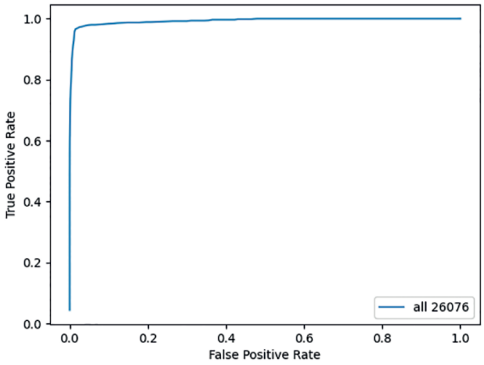

Рис. 7. ROC кривая по засвеченным номерам
Fig. 7. ROC curve by overexposed numbers

2 класса (читаемые, не читаемые). После чего была построена ROC [20] кривая (рис. 7). Данный график показывает отношение TP к FP (один из простых, но эффективных критериев качества классификации). Здесь TP (True Positive) — величина, показывающая отношение количества объектов, верно классифицированных как засвеченные, ко всем засвеченным объектам выборки. А FP (False Positive) — величина, показывающая отношение количества объектов, неверно классифицированных как засвеченные, ко всем незасвеченным номерам в выборке.

Построение ROC кривой позволило нам определить оптимальное значение порога засвеченности. Исходя из табл. 1, построенной по нашей выборке, оптимальным пороговым значением являет-

Таблица 1

Данные ROC кривой по засвеченным номерам
ROC curve data from overexposed numbers

Пороговое значение	% отсеченных засвеченных	% отсеченных незасвеченных
65	0,973536488	0,029924685
66	0,972734563	0,025856861
67	0,969526864	0,021466833
68	0,967923015	0,018164243
69	0,965517241	0,014700552
70	0,957497995	0,012445125
71	0,928628709	0,010632728
72	0,913392141	0,008941158
73	0,891740176	0,007289863
74	0,864474739	0,005678843
75	0,830793905	0,004833058
76	0,792301524	0,003584518
77	0,765838011	0,00249708
78	0,733761026	0,001731846
79	0,701684042	0,001288815
80	0,675220529	0,000926336
81	0,641539695	0,000644408
82	0,598235766	0,000281928
83	0,570970329	0,000161102
84	0,546110666	0,000120826
85	0,519647153	0,000120826
86	0,485966319	0
87	0,462710505	0

ся 70. При данном значении порога 95,7% номеров верно классифицируются как засвеченные, при этом только в 1,2% случаев незасвеченные номера классифицируются как засвеченные.

Определение степени смазанности

Для определения степени смазанности была построена нейронная сеть с уникальной архитектурой, которая должна была обеспечить точность классификации не ниже 90% при минимальном времени обработки одного номера на микрокомпьютерах типа Nvidia Jetson Nano, в пределах 20 мс. Требование к скорости работы выставлено согласно с тем, что данная нейронная сеть должна быть пригодна к использованию в системе реального времени совместно с основными элементами. В результате экспериментов для достижения указанных требований была получена нейронная сеть со следующей архитектурой (рис. 8).

Финальная архитектура нейронной сети, которая была сформирована в результате подбора оптимальной комбинации сверточных слоев и слоев пулинга, состоит из трех сверточных слоев, трех слоев пуллинга и двух полносвязных слоев. Последний полносвязный слой содержит 1 нейрон с сигмоидой в качестве функции активации для решения задачи бинарной классификации. Первый сверточный слой содержит 32 сверточных ядра размером 3 с шагом 1. Второй сверточный слой содержит 64 сверточных ядра размером 3 с шагом 1. Третий сверточный слой содержит 16 сверточных ядер размером 3 с шагом 1. Все сверточные слои и первый полносвязный слой имеют функцию активации ReLU. Пуллинг слои имеют размер ядра 2 и работают по принципу нахождения максимума. Также для минимизации вероятности переобучения нейронной сети к ее слоям добавлен слой «дропаут». Входное изображение, подающееся в нейронную сеть, имеет разрешение 224×224 пикселя с тремя цветовыми каналами.

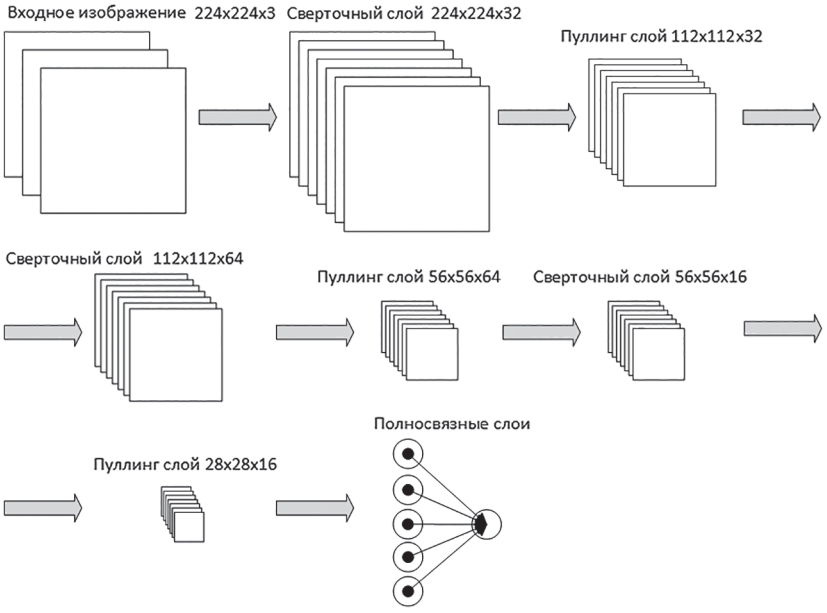

Рис. 8. Архитектура нейронной сети определения смазанности
Fig. 8. Architecture of the neural network for detecting blurring

Эксперименты с модификацией нейронной сети показали, что добавление дополнительных слоев не дает сильного прироста в качестве классификации, при этом время обработки одного изображения увеличивается.

Отдельной задачей в ходе работы над нейронной сетью для определения смазанности было создание датасета для обучения. Условия, при которых изображения получаются смазанными, являются специфичными, и чтобы отобрать из огромного количества данных те, которые были пригодны для класса смазанных номеров, ушло огромное количество времени. Для увеличения количества сма-

Таблица 2

Результат работы алгоритма определения засвеченности на ПК
The result of the algorithm for determining the overexposure on a PC on a PC

№	Корректный статус	Статус алгоритма	Значение засвеченности	Время работы алгоритма, с
1	0	0	0	$98{,}903*10^{-6}$
2	1	1	100	$7{,}826*10^{-6}$
3	0	0	17	$6{,}883*10^{-6}$
4	0	0	23	$6{,}772*10^{-6}$
5	0	0	30	$7{,}578*10^{-6}$
6	0	0	41	$6{,}572*10^{-6}$
7	0	0	50	$6{,}353*10^{-6}$
8	0	0	57	$6{,}373*10^{-6}$
9	0	0	61	$6{,}329*10^{-6}$
10	0	0	63	$6{,}517*10^{-6}$
11	0	0	65	$6{,}481*10^{-6}$
12	0	1	68	$6{,}591*10^{-6}$
13	0	0	70	$6{,}093*10^{-6}$
14	1	1	74	$6{,}149*10^{-6}$
15	1	1	78	$6{,}356*10^{-6}$
16	1	1	82	$6{,}262*10^{-6}$
17	1	1	87	$5{,}769*10^{-6}$
18	1	1	89	$6{,}612*10^{-6}$
19	1	1	91	$6{,}068*10^{-6}$
20	1	1	92	$5{,}945*10^{-6}$
21	1	1	95	$6{,}236*10^{-6}$
22	1	1	96	$5{,}62*10^{-6}$
23	1	1	98	$5{,}99*10^{-6}$
24	0	0	9	$6{,}089*10^{-6}$

Таблица 3

Результат работы алгоритма определения засвеченности на Nvidia Jetson Nano
The result of the algorithm for determining the overexposure on Nvidia Jetson Nano

№	Корректный статус	Статус алгоритма	Значение засвеченности	Время работы алгоритма, с
1	0	0	0	$28{,}026*10^{-6}$
2	1	1	100	$37{,}917*10^{-6}$
3	0	0	17	$41{,}146*10^{-6}$
4	0	0	23	$36{,}875*10^{-6}$
5	0	0	30	$40{,}418*10^{-6}$
6	0	0	41	$40{,}834*10^{-6}$
7	0	0	50	$34{,}323*10^{-6}$
8	0	0	57	$36{,}459*10^{-6}$
9	0	0	61	$45{,}105*10^{-6}$
10	0	0	63	$42{,}761*10^{-6}$
11	0	0	65	$39{,}584*10^{-6}$
12	0	0	68	$40{,}157*10^{-6}$
13	0	0	70	$32{,}709*10^{-6}$
14	1	1	74	$37{,}24*10^{-6}$
15	1	1	78	$45{,}782*10^{-6}$
16	1	1	82	$41{,}303*10^{-6}$
17	1	1	87	$39{,}22*10^{-6}$
18	1	1	89	$35{,}313*10^{-6}$
19	1	1	91	$32{,}553*10^{-6}$
20	1	1	92	$34{,}271*10^{-6}$
21	1	1	95	$35{,}73*10^{-6}$
22	1	1	96	$43{,}803*10^{-6}$
23	1	1	98	$37{,}396*10^{-6}$
24	0	0	9	$36{,}303*10^{-6}$

занных номеров были записаны видео с измененным значением выдержки, так, чтобы процент смазанных номеров на них был выше. Это позволило эффективнее собирать датасет для класса смазанных номеров.

Также для расширения датасета применялась аугментация.

В результате работы удалось получить нейронную сеть, которая выдает 94 — 95% точности в задаче бинарной классификации на классы смазанных и несмазанных ГРЗ. Средняя скорость работы сети при этом на ПК составляет 0,073 мс на одно изображение номера, на микрокомпьютере Nvidia Jetson Nano скорость об-

Таблица 4

Результат работы нейросети для детектирования смазанности на ПК
The result of a neural network for blur detection on PC

№	Корректный статус	Статус алгоритма	Значение смазанности	Время работы алгоритма, с
1	1	1	0,071	0,00301
2	1	1	0,989	0,00066
3	1	1	0,796	0,0005
4	1	1	0,996	0,0005
5	1	1	0,998	0,00049
6	1	1	0,993	0,0005
7	0	0	0,396	0,00049
8	1	1	0,995	0,00049
9	0	0	0,068	0,00049
10	0	0	0,21	0,00048
11	0	0	0,148	0,0007
12	0	0	0,05	0,00076
13	1	1	0,986	0,0006
14	0	0	0,009	0,0006
15	1	1	0,999	0,00058
16	0	0	0,148	0,00058
17	1	1	0,997	0,00057
18	0	0	0,108	0,00055
19	0	0	0,17	0,00056
20	0	0	0,062	0,00055
21	0	0	0,397	0,00055
22	1	1	0,851	0,00054
23	1	1	0,983	0,00056
24	1	1	1,0	0,00053

Таблица 5

Результат работы нейросети для детектирования смазанности на Nvidia Jetson Nano
The result of a neural network for blur detection on Nvidia Jetson Nano

№	Корректный статус	Статус алгоритма	Значение смазанности	Время работы алгоритма, с
1	1	1	0,989	0,01855
2	1	1	0,986	0,01879
3	0	0	0,05	0,01854
4	0	0	0,399	0,01854
5	0	0	0,148	0,01849
6	0	0	0,062	0,01854
7	0	0	0,068	0,01865
8	1	1	0,796	0,01848
9	1	1	1	0,01841
10	1	1	0,983	0,01859
11	1	1	0,995	0,01835
12	0	0	0,17	0,01835
13	0	0	0,148	0,01845
14	0	0	0,395	0,01841
15	1	1	0,993	0,01854
16	0	0	0,211	0,01938
17	1	1	0,998	0,0183
18	1	1	0,999	0,0186
19	0	0	0,071	0,01851
20	1	1	0,996	0,01877
21	1	1	0,851	0,01866
22	0	0	0,108	0,01844
23	0	0	0,01	0,01859
24	1	1	0,997	0,01692

работки одного изображения составляет 18,4 мс.

Для того, чтобы результатом работы нейронной сети было определение степени смазанности, на ее выходе добавлена сигмоидальная функция. В результате на выходе нейронной сети мы получили значение в интервале от 0 до 1. С помощью ROC кривой было определено, что пороговое значение для смазанных номеров нашего датасета равно 0,65. При данном значении порога отфильтровывается 95,2% смазанных номеров, при этом 2,2% несмазанных номеров ошибочно отфильтровываются как смазанные.

Результаты работы программного модуля на различных устройствах

Для апробации подхода мы использовали два типа устройств: персональный компьютер и микрокомпьютер Nvidia Jetson Nano. ПК обладает следующими техническими характеристиками: процессор Amd Ryzen5 3400G, оперативная память 16 Гб DDR4-3200, видеокарта nvidia RTX3080. Nvidia Jetson Nano имеет следующие характеристики: четырехядерный процессор Arm A57 с тактовой частотой 1,43 ГГц, оперативная память — 4 ГБ LPDDR4, графический процессор NVIDIA Maxwell 128 ядер.

Тестирование алгоритма определения засвеченности показало следующие результаты (см. табл. 2, 3).

Результаты тестирования нейросети для детектирования смазанности представлены в табл. 4, 5.

Заключение

Разработанные алгоритмы для определения смазанности и засвеченности государственных регистрационных знаков автомобилей позволяют достаточно точно оценить степень пригодности изображений для их дальнейшего распознавания. Очевидным преимуществом подхода с применением предварительной классификации качества информации, зарегистрированной камерой перед распознаванием, является существенная экономия вычислительных ресурсов и повышение точности работы системы, что особенно важно для сравнительно небольших горнопромышленных объектов. Своевременное отсеивание заведомо некорректных изображений ГРЗ позволяет не задействовать впустую вычислительные ресурсы для распознавания номеров, а также снижает вероятность некорректного распознавания.

В ходе экспериментов предложенные подходы показали свою эффективность на различных устройствах, таких как ПК и микрокомпьютер Nvidia Jetson Nano. Это позволяет говорить о том, что функционал определения смазанности и засвеченности с применением предложенных в статье методов применим к использованию как в серверных решениях, так и в мобильных «коробочных» решениях, где камера и вычислительное устройство представляют собой единое устройство.

Тестирование модуля определения засвеченности показало, что на ПК (процессор Amd Ryzen 53400G) средняя скорость обработки одного изображения колеблется от $5 \cdot 10^{-6}$ до $8 \cdot 10^{-6}$ с, а на микрокомпьютере Nvidia Jetson Nano составляет $36{,}8 \cdot 10^{-6}$ с.

Тестирование модуля определения смазанности на ПК показало скорость работы в интервале от $0{,}5 \cdot 10^{-3}$ до $3 \cdot 10^{-3}$ с, а на микрокомпьютере Nvidia Jetson Nano — в интервале от 0,016 до 0,019 с.

При этом качество работы модуля засвеченности следующее: 95,7% засвеченных номеров верно классифицировано как засвеченные при 1,2% случаев, когда незасвеченные номера классифицируются как засвеченные.

Качество работы модуля смазанности: 96,4% смазанных номеров верно

классифицированы как смазанные при 1,2% случаев, когда незасвеченные номера классифицируются как засвеченные. Разработанные модули являются кросс-платформенными, и собраны в виде отдельного Docker контейнера, что упрощает развертывание на различных устройствах.

СПИСОК ЛИТЕРАТУРЫ

1. *Arena F., Pau G., Severino A.* An overview on the current status and future perspectives of smart cars // Infrastructures. 2020, vol. 5, no. 7, article 53. DOI: 10.3390/infrastructures 5070053.

2. *Temkin I., Myaskov A., Deryabin S., Konov I., Ivannikov A.* Design of a digital 3D model of transport — technological environment of open-pit mines based on the common use of telemetric and geospatial information // Sensors. 2021, vol. 21, no. 18, article 6277. DOI: 10.3390/s21186277.

3. *Темкин И. О., Клебанов Д. А., Дерябин С. А., Конов И. С.* Построение интеллектуальной геоинформационной системы горного предприятия с использованием методов прогнозной аналитики // Горный информационно-аналитический бюллетень. — 2020. — № 3. — С. 114 — 125. DOI: 10.25018/0236-1493-2020-3-0-114-125.

4. *Zhang H., Tao P., Meng X., Liu M., Liu X.* An optimum deployment algorithm of camera networks for open-pit mine slope monitoring // Sensors. 2021, vol. 21, no. 4, article 1148. DOI: 10.3390/s21041148.

5. *Brooks C. N., Dobson R. J., Dean D. B., Banach D., Oommen T., Havens T. C., Ahlborn T., Cook S. J., Clover A.* Evaluating the use of unmanned aerial vehicles for transportation purposes // Michigan Department of Transportation, Michigan, USA, 2015. Report No. RC-1616, 201 p.

6. *Болотова Ю. А., Спицын В. Г., Рудомёткина М. Н.* Распознавание автомобильных номеров на основе метода связных компонент и иерархической временной сети // Компьютерная оптика. — 2015. — Т. 39. — № 2. — С. 275 — 280. DOI: 10.18287/0134-2452-2015-39-2-275-280.

7. *Полтавский А. В., Юрушкина Т. Г., Юрушкин М. В.* Автоматическое распознавание автомобильных номерных знаков // Вестник Донского государственного технического университета. 2020. — Т. 20. — № 1. — С. 93 — 99. DOI: 10.23947/1992-5980-2020-20-1-93-99.

8. *Yang X., Wang X.* Recognizing license plates in real-time // arXiv: 2019. 1906.04376. DOI: 10.48550/arXiv.1906.04376.

9. *Wang J., Hu X.* Convolutional neural networks with gated recurrent connections // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022, vol. 44, pp. 3421 — 3435. DOI: 10.1109/TPAMI.2021.3054614.

10. *Wang J., Hu Y.* Gated recurrent convolution neural network for OCR / NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017, pp. 334 — 343.

11. *Atienza R.* Vision transformer for fast and efficient scene text recognition // Document Analysis and Recognition — ICDAR. 2021, pp. 319 — 334. DOI: 10.1007/978-3-030-86549-8_21.

12. *Kissos I., Dershowitz N.* OCR error correction using character correction and feature-based word classification / 12th IAPR Workshop on Document Analysis Systems (DAS). 2016. DOI: 10.1109/DAS.2016.44.

13. *Mokhtar K., Bukhari S., Dengel A.* OCR Error Correction: State-of-the-art vs An NMT Based Approach / 13th IAPR International Workshop on Document Analysis Systems (DAS). 2018. DOI: 10.1109/DAS.2018.63.

14. *Jatzkowski I., Wilke D., Maurer M.* A deep-learning approach for the detection of overexposure in automotive camera images / 21st International Conference on Intelligent Transportation Systems (ITSC). 2018. DOI: 10.1109/ITSC.2018.8569692.

15. *Singh M., Tiwari R., Swami K., Vijayvargiya A.* Detection of glare in night photography / 23rd International Conference on Pattern Recognition (ICPR). 2016. DOI: 10.1109/ICPR.2016.7899744.

16. *Chen Y., Liu F., Pei K.* Self-supervised sun glare detection CNN for self-aware autonomous driving / Machine Learning for Autonomous Driving Workshop at the 35th Conference on Neural Information Processing System. 2021.

17. *Зотин А. Г., Пахирка А. И., Дамов М. В., Савчина Е. И.* Улучшение визуального качества изображений, полученных в сложных условиях освещенности на основе инфракрасных данных // Программные продукты и системы. — 2016. — № 3. — С. 109—120. DOI: 10.15827/0236-235X.115.109-120.

18. *Xin Yang, Ke Xu, Yibing Song, Qiang Zhang, Xiaopeng Wei, Rynson Lau* Image correction via deep reciprocating HDR transformation / Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018. DOI: 10.1109/CVPR.2018.00193.

19. *Zaidi S., Ansari M., Aslam A., Kanwal N., Asghar M., Lee B.* A Survey of modern deep learning based object detection models // Digital Signal Processing. 2022, vol. 126. DOI: 10.1016/j.dsp.2022.103514.

20. *Kai Feng, Han Hong, Ke Tang, Jingyuan Wang* Decision making with machine learning and ROC curves // SSRN Electronic Journal. 2019/06/01. DOI: 10.2139/ssrn.3382962. ГИАБ

REFERENCES

1. Arena F., Pau G., Severino A. An overview on the current status and future perspectives of smart cars. *Infrastructures*. 2020, vol. 5, no. 7, article 53. DOI: 10.3390/infrastructures5070053.

2. Temkin I., Myaskov A., Deryabin S., Konov I., Ivannikov A. Design of a digital 3D model of transport—technological environment of open-pit mines based on the common use of telemetric and geospatial information. *Sensors*. 2021, vol. 21, no. 18, article 6277. DOI: 10.3390/s21186277.

3. Temkin I. O., Klebanov D. A., Deryabin S. A., Konov I. S. Construction of intelligent geoinformation system for a mine using forecasting analytics techniques. *MIAB. Mining Inf. Anal. Bull.* 2020, no. 3, pp. 114—125. [In Russ]. DOI: 10.25018/0236-1493-2020-3-0-114-125.

4. Zhang H., Tao P., Meng X., Liu M., Liu X. An optimum deployment algorithm of camera networks for open-pit mine slope monitoring. *Sensors*. 2021, vol. 21, no. 4, article 1148. DOI: 10.3390/s21041148.

5. Brooks C. N., Dobson R. J., Dean D. B., Banach D., Oommen T., Havens T. C., Ahlborn T., Cook S. J., Clover A. *Evaluating the use of unmanned aerial vehicles for transportation purposes. Report No. RC-1616*. Michigan Department of Transportation, Michigan, USA, 2015, 201 p.

6. Bolotova Y. A., Spitsyn V. G., Rudometkina M. N. License plate recognition algorithm on the basis of a connected components method and a hierarchical temporal memory model. *Computer optics*. 2015, vol. 39, no. 2, pp. 275—280. [In Russ]. DOI: 10.18287/0134-2452-2015-39-2-275-280.

7. Poltavskii A. V., Yurushkina T. G., Yurushkin M. V. Automatic license-plate recognition. *Advanced Engineering Research*. 2020, vol. 20, no. 1, pp. 93—99. [In Russ]. DOI: 10.23947/1992-5980-2020-20-1-93-99.

8. Yang X., Wang X. Recognizing license plates in real-time. *arXiv*: 2019. 1906.04376. DOI: 10.48550/arXiv.1906.04376.

9. Wang J., Hu X. Convolutional neural networks with gated recurrent connections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2022, vol. 44, pp. 3421—3435. DOI: 10.1109/TPAMI.2021.3054614.

10. Wang J., Hu Y. Gated recurrent convolution neural network for OCR. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 334—343.

11. Atienza R. Vision transformer for fast and efficient scene text recognition. *Document Analysis and Recognition — ICDAR*. 2021, pp. 319—334. DOI: 10.1007/978-3-030-86549-8_21.

12. Kissos I., Dershowitz N. OCR error correction using character correction and feature-based word classification. *12th IAPR Workshop on Document Analysis Systems (DAS)*. 2016. DOI: 10.1109/DAS.2016.44.

13. Mokhtar K., Bukhari S., Dengel A. OCR Error Correction: State-of-the-art vs An NMT Based Approach. *13th IAPR International Workshop on Document Analysis Systems (DAS)*. 2018. DOI: 10.1109/DAS.2018.63.

14. Jatzkowski I., Wilke D., Maurer M. A deep-learning approach for the detection of overexposure in automotive camera images. *21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018. DOI: 10.1109/ITSC.2018.8569692.

15. Singh M., Tiwari R., Swami K., Vijayvargiya A. Detection of glare in night photography. *23rd International Conference on Pattern Recognition (ICPR)*. 2016. DOI: 10.1109/ICPR.2016.7899744.

16. Chen Y., Liu F., Pei K. Self-supervised sun glare detection CNN for self-aware autonomous driving. *Machine Learning for Autonomous Driving Workshop at the 35th Conference on Neural Information Processing System*. 2021.

17. Zotin A. G., Pakhirka A. I., Damov M. V., Savchina E. I. Visual quality enhancement of the image based on infrared data in video surveillance systems. *Software and Systems*. 2016, no. 3, pp. 109—120. [In Russ]. DOI: 10.15827/0236-235X.115.109-120.

18. Xin Yang, Ke Xu, Yibing Song, Qiang Zhang, Xiaopeng Wei, Rynson Lau Image correction via deep reciprocating HDR transformation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018. DOI: 10.1109/CVPR.2018.00193.

19. Zaidi S., Ansari M., Aslam A., Kanwal N., Asghar M., Lee B. A Survey of modern deep learning based object detection models. *Digital Signal Processing*. 2022, vol. 126. DOI: 10.1016/j.dsp.2022.103514.

20. Kai Feng, Han Hong, Ke Tang, Jingyuan Wang Decision making with machine learning and ROC curves. *SSRN Electronic Journal*. 2019/06/01. DOI: 10.2139/ssrn.3382962.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Епифанов Владислав Александрович[1,2] — аспирант,
e-mail: epifanov.vld@yandex.ru,
Темкин Игорь Олегович[1] — д-р техн. наук,
профессор, e-mail: igortemkin@yandex.ru,
Краснояружский Сергей Евгеньевич[2] — разработчик,
e-mail: blazzer4267@gmail.com,
[1] НИТУ «МИСиС»,
[2] ООО «Ситилабс».
Для контактов: Темкин И.О., e-mail: igortemkin@yandex.ru.

INFORMATION ABOUT THE AUTHORS

V.A. Epifanov[1,2], Graduate Student,
e-mail: epifanov_vld@yandex.ru,
I.O. Temkin[1], Dr. Sci. (Eng.), Professor,
e-mail: igortemkin@yandex.ru,
S.E. Krasnoyaruzhskiy[2], Developer,
e-mail: blazzer4267@gmail.com,
[1] National University of Science and Technology «MISiS»,
119049, Moscow, Russia,
[2] «Citylabs» LLC, Moscow, 121205, Russia.
Corresponding author: I.O. Temkin, e-mail: igortemkin@yandex.ru.

Получена редакцией 02.03.2023; получена после рецензии 03.04.2023; принята к печати 10.05.2023.
Received by the editors 02.03.2023; received after the review 03.04.2023; accepted for printing 10.05.2023.