Известия вузов России. Радиоэлектроника. 2020. Т. 23, № 6. С. 6–16
Journal of the Russian Universities. Radioelectronics. 2020, vol. 23, no. 6, pp. 6–16

Телевидение и обработка изображений

УДК 004.931; 004.932 Оригинальная статья

https://doi.org/10.32603/1993-8985-2020-23-6-6-16

Метод автоматического определения ключевых точек объекта на изображении

И. Г. Зубов✉

ООО "НЕКСТ", Москва, Россия

✉ZubovIG@gmail.com

Аннотация

Введение. Внедрение систем технического зрения в повседневную жизнь становится все более популярным. Использование систем на основе монокулярной камеры позволяет решить большой спектр задач. Анализ монокулярных изображений является наиболее развивающимся направлением в области машинного зрения. Это обусловлено общедоступностью цифровых камер и больших наборов аннотированных данных, а также мощностью современной вычислительной техники. Для того чтобы система компьютерного зрения описывала объекты и предсказывала их действия в физическом пространстве сцены, необходимо интерпретировать анализируемое изображение с точки зрения базовой 3D-сцены. Этого можно достичь, анализируя жесткий объект как совокупность взаимно связанных частей, что представляет мощный контекст и структуру для рассуждений о физическом взаимодействии.

Цель работы. Разработка автоматического метода ключевых точек объекта интереса на изображении.

Методы и материалы. Предложен автоматический метод локализации ключевых точек транспортных средств на изображении, в частности номерного знака. Представленный метод позволяет зафиксировать ключевые точки объекта интереса на основе анализа сигналов внутренних слоев сверточных нейронных сетей, обученных для классификации изображений, и выделения объектов на изображении. Также метод позволяет детектировать части объекта без больших затрат на аннотацию данных и обучение.

Результаты. Эксперименты подтвердили корректность выделения ключевой точки объекта интереса на основе предложенного метода. Точность выделения ключевой точки на номерном знаке составила 97 %.

Заключение. Представлен новый метод выделения ключевых точек объекта интереса на основе анализа сигналов внутренних слоев сверточных нейронных сетей. Метод обладает точностью выделения ключевых точек объекта интереса на уровне современных методов, а в отдельных случаях превосходит их.

Ключевые слова: сверточные нейронные сети, анализ карт активаций, выделение ключевых точек

Для цитирования: Зубов И. Г. Метод автоматического определения ключевых точек объекта на изображении // Изв. вузов России. Радиоэлектроника. 2020. Т. 23, № 6. С. 6–16. doi: 10.32603/1993-8985-2020-23-6-6-16

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Статья поступила в редакцию 17.07.2020; принята к публикации после рецензирования 20.10.2020; опубликована онлайн 29.12.2020

© Зубов И. Г., 2020

Контент доступен по лицензии Creative Commons Attribution 4.0 License
This work is licensed under a Creative Commons Attribution 4.0 License

Television and Image Processing

Original article

An Automatic Method for Interest Point Detection

Ilya G. Zubov✉

Ltd "Next", Moscow, Russia

✉ZubovIG@gmail.com

Abstract

Introduction. Computer vision systems are finding widespread application in various life domains. Monocular-camera based systems can be used to solve a wide range of problems. The availability of digital cameras and large sets of annotated data, as well as the power of modern computing technologies, render monocular image analysis a dynamically developing direction in the field of machine vision. In order for any computer vision system to describe objects and predict their actions in the physical space of a scene, the image under analysis should be interpreted from the standpoint of the basic 3D scene. This can be achieved by analysing a rigid object as a set of mutually arranged parts, which represents a powerful framework for reasoning about physical interaction.

Objective. Development of an automatic method for detecting interest points of an object in an image.

Materials and methods. An automatic method for identifying interest points of vehicles, such as license plates, in an image is proposed. This method allows localization of interest points by analysing the inner layers of convolutional neural networks trained for the classification of images and detection of objects in an image. The proposed method allows identification of interest points without incurring additional costs of data annotation and training.

Results. The conducted experiments confirmed the correctness of the proposed method in identifying interest points. Thus, the accuracy of identifying a point on a license plate achieved 97%.

Conclusion. A new method for detecting interest points of an object by analysing the inner layers of convolutional neural networks is proposed. This method provides an accuracy similar to or exceeding that of other modern methods.

Keywords: convolutional neural networks, activation map analysis, interest point detection

For citation: Zubov I. G. An Automatic Method for Interest Point Detection. Journal of the Russian Universities. Radioelectronics. 2020, vol. 23, no. 6, pp. 6–16. doi: 10.32603/1993-8985-2020-23-6-6-16

Conflict of interest. The author declares no conflict of interest.

Submitted 17.07.2020; accepted 20.10.2020; published online 29.12.2020

Введение. Внедрение систем технического зрения в повседневную жизнь становится все более популярным. Использование систем на основе монокулярной камеры позволяет решить большой спектр задач. Анализ монокулярных изображений является наиболее развивающимся направлением в области машинного зрения. Это обусловлено общедоступностью цифровых камер и больших наборов аннотированных данных, а также мощностью современной вычислительной техники.

Для того чтобы система компьютерного зрения описывала объекты и предсказывала их действия в физическом пространстве сцены, необходимо интерпретировать анализируемое изображение с точки зрения базовой 3D-сцены. Анализ жесткого объекта как совокупности взаимно связанных частей представляет мощный контекст и структуру для рассуждений о физическом взаимодействии.

В настоящее время системы, основанные на сверточных нейронных сетях, обладают наибольшей точностью при решении задач обнаружения и классификации объектов на изображении. Одним из важных факторов для дальнейшего прогресса является понимание представлений, которые изучаются внутренними слоями нейронной сети. На данный момент существуют подходы, предназначенные для попыток визуальной интерпретации работы сверточных нейронных сетей.

Промежуточные выходы сигналов сверточных нейронных сетей часто могут быть связаны с семантическими частями объектов [1]. Как показано в [1], ядра сверток в начальных слоях похожи на

маски фильтров, используемые во многих популярных низкоуровневых дескрипторах объектов. Более глубокие слои чувствительны к более абстрактным узорам на изображении [1]. Эти шаблоны могут даже соответствовать целым объектам [2] или их частям [3].

Детекторы объектов реализуются в результате обучения сверточной нейронной сети для выполнения классификации сцен [4]. Одна и та же сеть может выполнять как распознавание сцен, так и выделение объектов в одном прямом проходе без какого-либо специального обучения понятию объектов.

Метод. Как правило, нейросетевые архитектуры, разработанные специально для классификации изображений, используются в качестве базовой структуры (Backbone) [5] для других задач, включая обнаружение и сегментацию.

Базовую структуру сверточной нейронной сети в общем виде можно представить как многоступенчатый каскад последовательно объединенных операций свертки и функций активации (рис. 1). Такая структура на основе предварительно обученных классификационных моделей используется без полносвязных слоев и предназначена для извлечения из изображения абстрактных карт признаков. Часто в качестве базовых структур используют такие предварительно обученные нейросетевые архитектуры, как ResNet [6], VGG [7], Mobile Net [8], SqeezNet [9] и т. д.

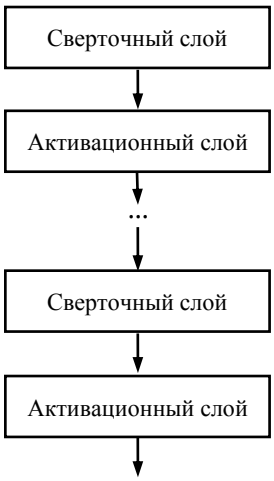

Рис. 1. Базовая структура сверточной нейронной сети

Fig. 1. The backbone of a convolutional neural network

В настоящей статье в качестве объекта интереса рассматривается транспортное средство. Для визуализации предложенного метода будет использована базовая структура сверточной нейронной сети VGG16, предназначенная для классификации изображений (рис. 2). Эта сеть оперирует со свертками 3×3 элемента, количество сверток указано в каждом слое.

Предложенный метод был исследован для сверточных нейронных сетей, использующих наиболее распространенные функции активации ReLU и Leaky ReLU. ReLU предусматривает установку пороговых значений на 0:

$$f(x) = \begin{cases} x, & x > 0; \\ 0, & x \le 0, \end{cases}$$

т. е. возвращает 0 при $x \le 0$ и линейную функцию при положительных значениях аргумента. Leaky ReLU описывается следующим образом:

$$f(x) = \begin{cases} x, & x \ge 0; \\ \alpha x, & x < 0, \end{cases}$$

где α – малая константа. В настоящей статье рассмотрена реализация YoloV3 [10] с константой $\alpha = 0.1$.

Основные шаги алгоритма.

Шаг 1. Изменение базовой структуры. Каждый слой нейронной сети может быть потенциально значимым и нести нужную информацию об интересующей семантической части объекта. Анализируя выходные сигналы каждого слоя нейронной сети, необходимо изменить ее архитектуру. Для анализа сигналов внутренних слоев к каждому активационному слою добавляется выход с последующим масштабированием выходной матрицы до разрешения входного изображения (рис. 3).

Преобразовав базовую структуру сверточной нейронной сети VGG16 предложенным методом, получим новую архитектуру (рис. 4).

Измененная базовая структура на выходе имеет матрицу с размерами $W \times H \times 4224$, где W

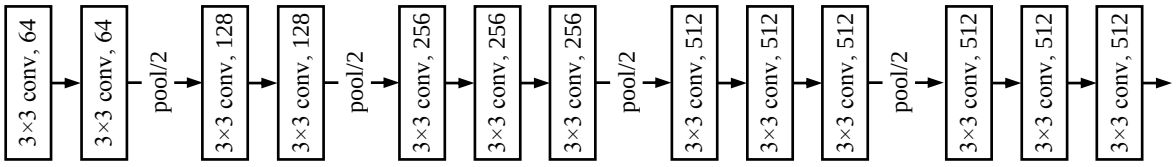

Рис. 2. Базовая структура сверточной нейронной сети VGG16

Fig. 2. The backbone of the convolutional neural network VGG16

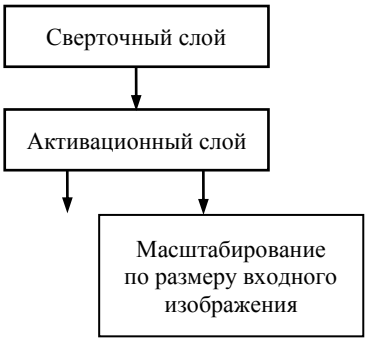

Рис. 3. Добавление выходного канала в базовую структуру

Fig. 3. Adding an output channel to the backbone

и H – высота и ширина входного изображения соответственно; 4224 – количество промежуточных сверток нейронной сети VGG16.

Визуализируя сигналы, полученные на промежуточных слоях сверточной нейронной сети, можно подтвердить сделанные в [1] предположения, что эти сигналы могут быть связаны с семантическими частями объектов. В качестве примера на рис. 5 представлены примеры набора изображений Carvana [11], а на рис. 6 – некоторые промежуточные сигналы сверточных слоев, активированные на семантические части автомобиля.

Шаг 2. Анализ сигналов внутренних слоев. Как видно из рис. 6, различные свертки могут активироваться на семантически одинаковые части объекта. Учитывая распространенные функции активации, автор настоящей статьи предлагает метод агрегации сигналов наиболее устойчивых сверток и их использования для выделения ключевых точек объекта интереса.

Для дальнейшего анализа активации отдельных сверток на определение семантической принадлежности к определенной части объекта интереса необходим набор аннотированных данных, указывающих положения интересующей части объекта.

Формально задачу выделения наиболее значимых сверточных слоев и каналов можно описать следующим образом: пусть X – множество изображений I и масок A с разрешением $H \times W$. Количество сверточных слоев равно T, $f_{ij}(I)$, $i \in [1:T]$, $j \in [1:T_i]$ – функция активации i-го слоя, j-го канала. Выделим N_1 сверток наиболее устойчивых в области интереса:

$$\hat{i}_j = \left\{ i : Q_{ij} \geq \mathrm{th}_1 \,\&\, Q_{ij} \geq Q_{(|Q|-N_1)} \right\},$$

и N_2 сверток, не активирующихся в интересующей области:

$$\tilde{i}_j = \left\{ i : Q_{ij} \leq \mathrm{th}_2 \,\&\, Q_{ij} \geq Q_{(N_2)} \right\},$$

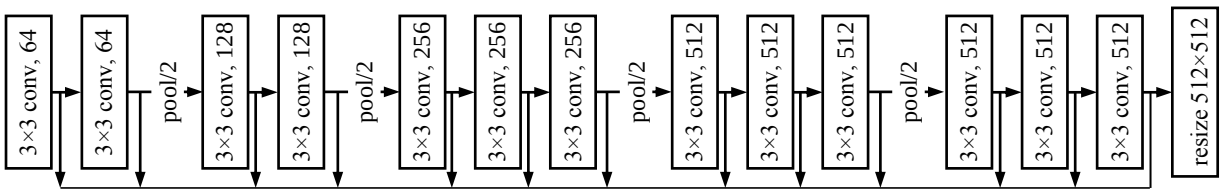

Рис. 4. Базовая структура сверточной нейронной сети VGG16 после преобразования

Fig. 4. The backbone of the convolutional neural network VGG16 after conversion

Рис. 5. Пример изображений набора данных Carvana

Fig. 5. Sample of the Carvana dataset images

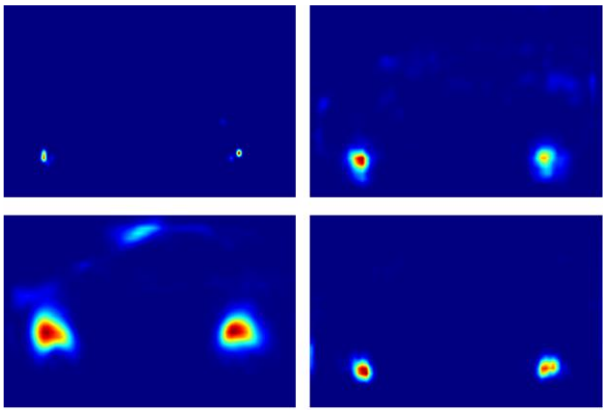

Рис. 6. Пример визуализации промежуточных выходов сверточных слоев нейронной сети VGG16

Fig. 6. Samples of visualization of the intermediate outputs of the convolutional layers of the neural network VGG16

где Q – доля сосредоточения активации в области интереса;

$$Q_{ij} = \sum_{\{I,\, A\} \in X} \frac{\sum\limits_{k \in 1,\, H} \sum\limits_{l \in 1,\, W} f_{ij}(I)(k,\, l) A(k,\, l)}{\sum\limits_{k \in 1,\, H} \sum\limits_{l \in 1,\, W} f_{ij}(I)(k,\, l)}.$$

Порог $0 \le \mathrm{th}_1 \le 1$ выбирается экспериментально: например, $\mathrm{th}_1 = 1$ будет указывать на полное сосредоточение активации в области интереса и отсутствие активации вне этой области. Порог $0 \le \mathrm{th}_2 \le 1$ выбирается в зависимости от функции активации, используемой в анализируемой архитектуре: например, при использовании функции активации RelU $\mathrm{th}_2 = 0$ будет указывать на полное отсутствие активации нейрона в области интереса.

Таким образом, элементы, не входящие в множества $\hat{i}_j$ и $\tilde{i}_j$, могут быть удалены из базовой архитектуры.

Шаг 3. Агрегация. Агрегируя выходные сигналы базовой структуры нейронной сети, получим тепловую карту интересующей части объекта

$$F = \sum_{i \in 1,\, T} \left(\sum_{i \in \hat{i}_j} f_{ij}[I] - \sum_{i \in \tilde{i}_j} f_{ij}[I] \right).$$

Шаг 4. Выделение маски и ключевой точки. Чтобы получить маску интересующей части объекта, используется пороговая фильтрация

$$B = F(x,\, y) \ge \mathrm{th},$$

где порог th выбирается экспериментально, максимизируя целевую метрику.

Координаты ключевой точки определяются выражением $x,\, y = \arg\max_{x,\, y} F(x,\, y)$.

Результаты. В настоящей статье в качестве объекта интереса выбран номерной знак автомобиля. Выделение номерного знака – одна из наиболее распространенных задач определения автомобиля. Для решения задачи выделения номерного знака разработано много различных подходов [12–14], а также сформированы наборы изображений с подробной аннотацией.

Задачу выделения номерного знака на изображении принято делить на 2 случая:

– поиск номерного знака на части изображения, содержащей только одно транспортное средство;

– поиск номерного знака на всем кадре без предварительного выделения автомобиля.

В настоящей статье рассматривается задача локализации номерного знака, а также определение его ключевой точки на части изображения, содержащей только одно транспортное средство. Результаты выделения по разработанному методу сравниваются с результатами, даваемыми тремя широко распространенными подходами к выделению номерных знаков:

1. Метод Виолы–Джонса [12]. Метод обнаружения объектов на изображении, предложенный Полом Виолой и Майклом Джонсом, основан на применении каскада простых классификаторов на основе признаков Хаара. Функции Хаара используются в качестве сверточного ядра при рассмотрении смежных прямоугольных областей. Несмотря на то, что метод был предложен в 2001 г., он до сих пор широко используется в задачах поиска номерных знаков на изображении в реальном времени. В настоящей статье представлена реализация [15].

2. Warped Planar Object Detection Network (WPOD-Net) [13] – сверточная нейронная сеть, разработанная для задачи выделения номерного знака и его ключевой точки. Нейронная сеть предсказывает восьмиканальную карту объектов, которая кодирует вероятности наличия объекта и параметры аффинного преобразования для выравнивания номерного знака по вертикали и по горизонтали. Архитектура сети имеет в общей сложности 21 сверточный слой, причем 14 находятся внутри остаточных блоков. В настоящей статье представлена реализация [15].

3. Mask-RCNN [14] – архитектура современной сверточной нейронной сети для сегментации объектов на изображениях. Mask-RCNN представляет собой конвейер "сегментации с использованием распознавания", который можно представить в виде следующих модулей:

– генератора особенностей (features extractor) на основе базовой структуры ResNet [6], формирующего трехмерную абстрактную матрицу признаков входного изображения;

– Region Proposal Network – сети генерации регионов. Извлекает потенциальные регионы, которые могут содержать объекты интересов;

– полносвязных слоев – сети, которая для каждого региона вырезает из матрицы особенностей соответствующую этому региону часть, выдавая класс объекта и уточненный описывающий объект – прямоугольник, а также формирует маску объекта интереса.

В настоящей статье представлена реализация [17].

В качестве тестируемых моделей используются две распространенные архитектуры сверточных нейронных сетей, предназначенных для классификации изображений, – VGG16 и SqueezeNet (1_0), а также распространенная модель выделения объектов YoloV3. В настоящей статье используется реализация [18].

Экспериментальная проверка предложенного метода проводилась с использованием двух наборов изображений:

1. Внутреннего набора, включающего более 13 000 изображений, на которых содержалось более 21 000 автомобилей. Данные были получены с камеры фотовидеофиксации на дорогах Москвы. Для каждого изображения вручную были размечены границы автомобиля и номерного знака. Разрешение каждого изображения 2448 × 2048. Изображения получены при различной освещенности, различных погодных условиях и представлены в градациях серого. Пример изображения с размеченной аннотацией приведен на рис. 7.

2. Chinese City Parking Dataset (CCPD) [19] – набора изображений от городской компании по управлению парковками в одной из провинциальных столиц Китая. CCPD предоставлял более 250 тыс. уникальных изображений, на которых присутствовали номерные знаки с подробными аннотациями. В качестве тестовой выборки было выбрано 70 000 изображений случайным образом. Разрешение каждого изображения составляло 720 × 1160 × 3 (ширина × высота × каналы). В аннотации содержались 4 точки углов номерного знака. CCPD содержало изображения при различной освещенности и окружающей среде в разную погоду. Изображения в CCPD получены с разных позиций и углов. Пример изображения с размеченной аннотацией приведен на рис. 8.

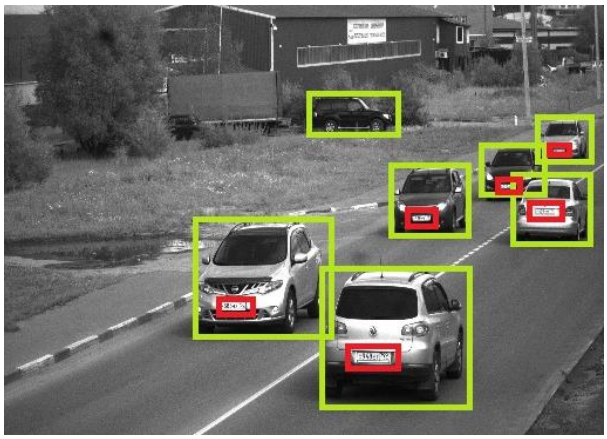

Рис. 7. Примеры визуализации аннотаций внутреннего набора изображений

Fig. 7. Samples of visualization of the annotations of an internal set of images

Рис. 8. Пример визуализации аннотаций набора изображений CCPD

Fig. 8. The sample of the visualization of annotations of CCPD image set

В качестве калибровочных данных (шаг 2) использовалось 100 изображений для каждого из наборов данных.

Пусть A, B – точная и вычисленная маски номерного знака соответственно $(A, B \subseteq S \in R_n)$. Для определения точности локализации ключевой точки номерного знака используется метрика качества

$$\text{acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},$$

где TP, TN, FP, FN – количества истинных положительных, истинных отрицательных, ложных положительных и ложных отрицательных решений соответственно. Решение считается истинно положительным, если $(x, y) \in A$. В табл. 1 приведена точность локализации ключевой точки номерного знака. Точность определения номерного знака при различных значениях фиксированного порога пересечения по объединению (intersection over union – IoU)

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|},$$

где A, B – точная и вычисленная маски номерного знака.

Табл. 1. Сравнительные результаты определения ключевой точки

Table 1. Comparative results of the key point detection

Метод определения	acc	
	CCPD	Внутренний набор
WPOD-NET	0.97	0.66
VGG19	0.97	0.95
Sqeeznet	0.93	0.93
YoloV3	0.93	0.96

Точность характеризуется метриками качества precision (точность) и recall (полнота):

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}; \text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

В табл. 2 приведена точность определения номерного знака распространенными методами для базы изображений CCPD и внутренней базы изображений.

В табл. 3 приведена точность определения номерного знака предложенным методом на изображениях внутренней базы и базы CCPD для трех различных сверточных нейронных сетей. Примеры выделения номерного знака разработанным методом представлены на рис. 9 и 10. На рис. 11 представлен результат работы предложенного метода на изображениях, содержащих несколько объектов интереса.

Табл. 2. Сравнительные результаты определения номерного знака известными методами по изображениям базы CCPD и внутренней базы

Table 2. Comparative results of the license plate detection using known methods on the images of CCPD database and internal database

Метод определения	I_0U							
	0.1		0.3		0.5		0.7	
	Метрика качества – precision							
	База изображений							
	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя
Mask-RCNN	0.98	0.92	0.98	0.92	0.98	0.92	0.86	0.91
Виолы–Джонса	0.75	0.89	0.67	0.87	0.26	0.65	0.02	0.49
WPOD-NET	0.98	0.96	0.98	0.96	0.92	0.95	0.39	0.89
Метод определения	I_0U							
	0.1		0.3		0.5		0.7	
	Метрика качества – recall							
	База изображений							
	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя
Mask-RCNN	0.76	0.91	0.76	0.91	0.75	0.91	0.66	0.91
Виолы–Джонса	0.25	0.71	0.24	0.71	0.09	0.66	0.01	0.49
WPOD-NET	0.97	0.77	0.96	0.77	0.90	0.77	0.38	0.72

Табл. 3. Сравнительные результаты определения номерного знака предложенным методом с использованием различных нейронных сетей по изображениям базы CCPD и внутренней базы

Table 3. Comparative results of the license plate detection using proposed method with various neural networks on the images of CCPD database and internal database

Нейронная сеть	I_0U							
	0.1		0.3		0.5		0.7	
	Метрика качества – precision							
	База изображений							
	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя
VGG16	0.95	0.98	0.93	0.96	0.85	0.86	0.75	0.69
SqeezNet	0.93	0.93	0.91	0.91	0.83	0.83	0.70	0.70
YoloV3	0.94	0.97	0.92	0.93	0.82	0.81	0.69	0.66
Нейронная сеть	I_0U							
	0.1		0.3		0.5		0.7	
	Метрика качества – recall							
	База изображений							
	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя	CCPD	Внутренняя
VGG16	0.95	0.98	0.93	0.95	0.85	0.86	0.75	0.68
SqeezNet	0.93	0.93	0.91	0.91	0.83	0.83	0.70	0.70
YoloV3	0.93	0.96	0.91	0.93	0.82	0.80	0.68	0.66

Времена анализа одного изображения описанными в статье нейросетевыми архитектурами представлены в табл. 4. Разрешение входного изображения составляло 512 × 512 пикселей. Анализ производился на видеокарте NVIDIA GeForce GTX 1650.

Табл. 4. Сравнение времен анализа нейронных сетей

Table 4. Comparison of analysis time of neural networks

Метод определения	Время обработки, мс
WPOD-NET	13
Mask-RCNN	166
VGG19	66
SqeezNet	9
YoloV3	50

Рис. 9. Результат применения предложенного метода к изображениям базы CCPD

Fig. 9. The result of applying of the proposed method to the images of the CCPD database

Рис. 10. Результаты применения метода к внутренней базе изображений

Fig 10. The results of applying the proposed method to the internal image database

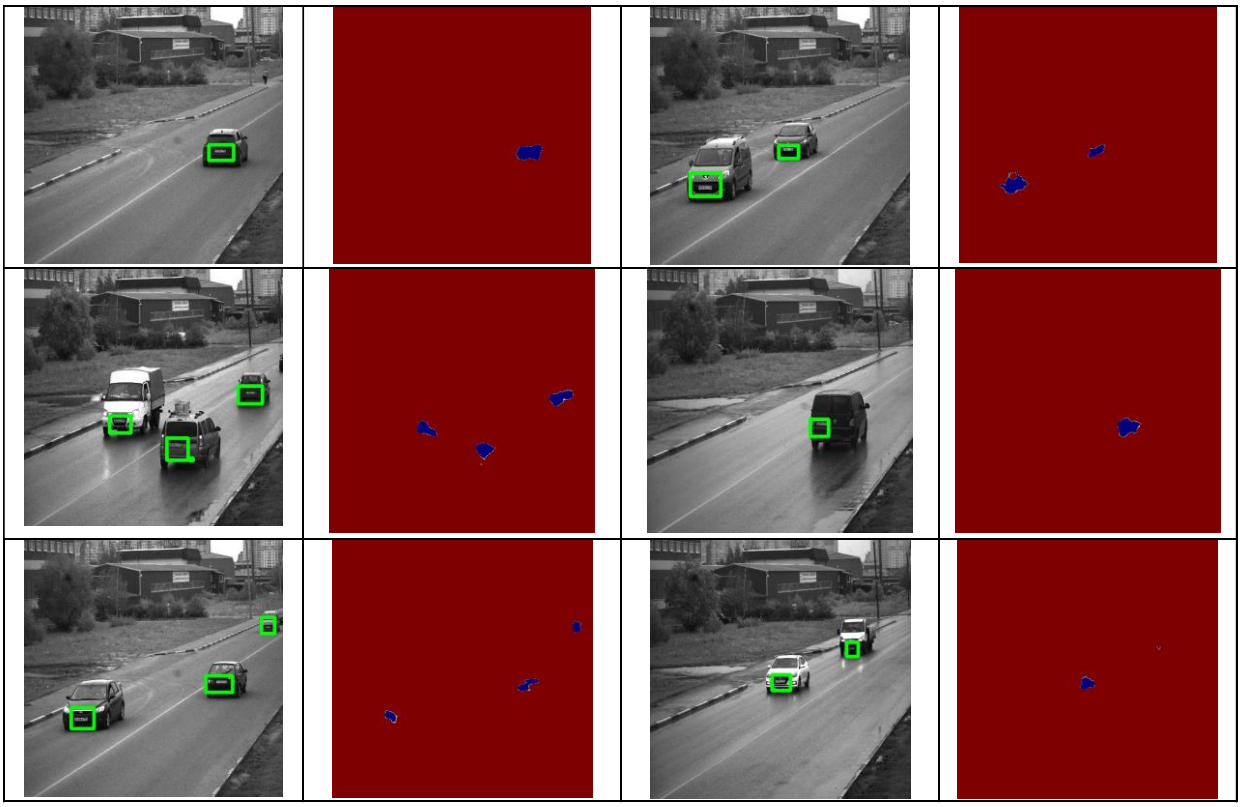

Рис. 11. Результат применения предложенного метода к изображениям, содержащим несколько автомобилей

Fig 11. The result of applying of the proposed method to images containing several vehicles

Вывод. В статье представлен новый метод автоматической локализации ключевых точек объекта интереса на изображении. Разработанный метод не требует больших наборов аннотированных данных и обучения нейросетевой модели, позволяет повысить информативность методов анализа изображений на основе сверточных нейронных сетей. Работоспособность и конкурентоспособность метода по отношению к известным алгоритмам выделения номерного знака WPOD-NET, Haar Cascad, Mask-RCNN проверены тестированием на открытой базе изображений CCPD, а также на внутренней базе изображений.

Список литературы

1. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks // Proc. of the 13th Europ. Conf. on Computer Vision, Zurich, Switzerland, 6–12 Sept. 2014. Berlin: Springer, 2014. P. 818–833. doi: 10.1007/978-3-319-10590-1_53

2. Simonyan K., Vedaldi A., Zisserman A. Deep inside convolutional networks: Visualising image classification models and saliency maps // Proc. of the ICLR Intern. Conf. on Learning Representations, Banff, Canada, Apr. 2014. URL: https://arxiv.org/abs/1312.6034 (дата обращения 15.11.2020)

3. Simon M., Rodner E., Denzler J. Part Detector Discovery in Deep Convolutional Neural Networks // Proc. of the ACCV Asian Conf. on Computer Vision, Singapore, 1–5 Nov. 2014. Berlin: Springer, 2014. Pt. 2. P. 162–177. doi: 10.1007/978-3-319-16808-1_12

4. Object Detectors Emerge in Deep Scene CNNs / B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba // Intern. Conf. on Learning Representations, San Diego, USA, May 2015. URL: http://hdl.handle.net/1721.1/96942 (дата обращения 15.11.2020)

5. MMDetection: open MMLab Detection Toolbox and Benchmark / K. Chen, J. Wang, J. Pang, Yu. Cao, Yu Xiong, X. Li, Sh. Sun, W. Feng, Z. Liu, J. Xu, Zh. Zhang, D. Cheng, Ch. Zhu, T. Cheng, Q. Zhao, B. Li, X. Lu, R. Zhu, Y. Wu, J. Dai, J. Wang, J. Shi, W. Ouyang, Ch. Change Loy, D. Lin. 13 p. URL: https://arxiv.org/pdf/1906.07155.pdf (дата обращения 02.06.2020)

6. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun // Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, USA, 27–30 June 2016. Piscataway: IEEE, 2016. Art. 16541111. doi: 10.1109/CVPR.2016.90

7. Simonyan K., Zisserman A. Very Deep Convolutional Networks for large-Scale Image Recognition. Apr 2015. 14 p. URL: https://arxiv.org/abs/1409.1556 (дата обращения 02.06.2020)

8. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications / A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam. URL: https://arxiv.org/abs/1704.04861 (дата обращения 02.06.2020)

9. SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5MB Model Size / F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer. URL: https://arxiv.org /abs/1602.07360 (дата обращения 02.06.2020)

10. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. URL: https://arxiv.org/abs/1804.02767 (дата обращения 02.06.2020)

11. Carvana Image Masking Challenge. URL: https://www.kaggle.com/c/carvana-image-masking-challenge (дата обращения 02.06.2020)

12. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features / Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Kauai, USA, 8–14 Dec. 2001. Piscataway: IEEE. Art. 7176899. doi: 10.1109/CVPR.2001.990517

13. Silva S. M., Jung C. R. License Plate Detection and Recognition in unconstrained Scenarios // Computer Vision – ECCV 2018. 15th Europ. Conf., Munich, Germany, 8–14 Sept. 2018. Berlin: Springer, 2018. P. 593–609. doi: 10.1007/978-3-030-01258-8_36

14. Mask R-Cnn / K. He, G. Gkioxari, P. Dollar, R. Girshick // IEEE Intern. Conf. on Computer Vision (ICCV), Venice, Italy, 22–29 Oct. 2017. Piscataway: IEEE, 2017. Art. 17467816. doi: 10.1109 /ICCV.2017.322

15. Haar Cascad license plate detection. Weights for the model. URL: https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_licence_plate_rus_16stages.xml (дата обращения 15.11.2020)

16. Silva S. M., Jung C. R. License Plate Detection and Recognition in Unconstrained Scenarios. URL: http://sergiomsilva.com/pubs/alpr-unconstrained/ (дата обращения 31.05.2020)

17. Nomeroff Net. A Open Source Python License Plate Recognition Framework. URL: https://nomeroff.net.ua/ (дата обращения 31.05.2020)

18. PyTorch implementation of YOLOv3. URL: http://docs.openvinotoolkit.org/2019_R2/_intel_models_person_vehicle_bike_detection_crossroad_1016_description_person_vehicle_bike_detection_crossroad_1016.html (дата обращения 31.05.2020).

19. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline / Z. Xu, W. Yang, A. Meng, N. Lu, H. Huang, C. Ying, L. Huang // Computer Vision. ECCV 2018. 15th Europ. Conf., Munich, Germany, 8–14 Sep. 2018. Berlin: Springer, 2018. P. 261–277. doi: 10.1007/978-3-030-01261-8_16

Информация об авторе

Зубов Илья Геннадьевич – магистр техники и технологий (2016), программист-алгоритмист компании ООО "НЕКСТ". Автор шести научных публикаций. Сфера научных интересов – цифровая обработка изображений; прикладные телевизионные системы.
Адрес: ООО "НЕКСТ", ул. Рочдельская, д. 15, стр. 13, Москва, 123022, Россия
E-mail: ZubovIG@gmail.com

References

1. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks. Proc. of the 13th Europ. Conf. on Computer Vision, Zurich, Switzerland, 6–12 Sept. 2014. Berlin: Springer, 2014, pp. 818–833. doi: 10.1007/978-3-319-10590-1_53

2. Simonyan K., Vedaldi A., Zisserman A. Deep inside convolutional networks: Visualising image classification models and saliency maps. Proc. of the ICLR Intern. Conf. on Learning Representations, Banff, Canada, apr. 2014. Available at: https://arxiv.org/abs/1312.6034 (accessed 15.11.2020)

3. Simon M., Rodner E., Denzler J. Part Detector Discovery in Deep Convolutional Neural Networks. Proc. of the ACCV Asian Conf. on Computer Vision, Singapore, 1–5 Nov. 2014. Berlin: Springer, 2014. Pt. 2, pp. 162–177. doi: 10.1007/978-3-319-16808-1_12

4. Zhou B., Khosla A., Lapedriza A., Oliva A., Torralba A. Object Detectors Emerge in Deep Scene CNNs. Intern. Conf. on Learning Representations, San Diego, USA, May 2015. Available at: http://hdl.handle.net/1721.1/96942 (accessed 15.11.2020)

5. Chen K., Wang J., Pang J., Cao Yu., Xiong Yu, Li X., Sun Sh., Feng W., Liu Z., Xu J., Zhang Zh., Cheng D., Zhu Ch., Cheng T., Zhao Q., Li B., Lu X., Zhu R., Wu Y., Dai J., Wang J., Shi J., Ouyang W., Change Loy Ch., Lin D. MMDetection: open MMLab Detection Toolbox and Benchmark, 13 p. Available at: https://arxiv.org/pdf/1906.07155.pdf (accessed 02.06.2020)

6. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, USA, 27–30 June 2016. Piscataway: IEEE, 2016, art. 16541111. doi: 10.1109/CVPR.2016.90

7. Simonyan K., Zisserman A. Very Deep Convolutional Networks for large-Scale Image Recognition. Apr. 2015, 14 p. Available at: https://arxiv.org/abs/1409.1556 (accessed 02.06.2020)

8. Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. MobileNets: Efficient Convolutional Neural Networks for

Mobile Vision Applications. Available at: https://arxiv.org/abs/1704.04861 (accessed 02.06.2020)

9. Iandola F. N., Han S., Moskewicz M. W., Ashraf K., Dally W. J., Keutzer K. SqueezeNet: AlexNet-Level Accuracy with 50x Few-er Parameters and <0.5MB Model Size. Available at: https://arxiv.org/abs/1602.07360 (accessed 02.06.2020)

10. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. Available at: https://arxiv.org/abs/1804.02767 (accessed 02.06.2020)

11. Carvana Image Masking Challenge. Available at: https://www.kaggle.com/c/carvana-image-masking-challenge (accessed 02.06.2020)

12. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Kauai, USA, 8–14 Dec. 2001. Piscataway: IEEE, 2001, art. 7176899. doi: 10.1109/CVPR.2001.990517

13. Silva S. M., Jung C. R. License Plate Detection and Recognition in unconstrained Scenarios // Computer Vision, ECCV 2018, 15th Europ. Conf., Munich, Germany, 8–14 Sept. 2018. Berlin: Springer, 2018, pp. 593–609. doi: 10.1007/978-3-030-01258-8_36

14. He K., Gkioxari G., Dollar P., Girshick R. Mask R-Cnn. IEEE Int. Conf. on Computer Vision (ICCV). Venice, Italy, 22–29 Oct. 2017. Piscataway: IEEE, 2017, art. 17467816. doi: 10.1109 /ICCV.2017.322

15. Haar Cascad license plate detection. Weights for the model. Available at: https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_licence_plate_rus_16stages.xml (accessed 15.11.2020)

16. Silva S. M., Jung C. R. License Plate Detection and Recognition in Unconstrained Scenarios. Available at: http://sergiomsilva.com/pubs/alpr-unconstrained/ (accessed 31.05.2020)

17. Nomeroff Net. A Open Source Python License Plate Recognition Framework. Available at: https://nomeroff.net.ua/ (accessed 31.05.2020)

18. PyTorch implementation of YOLOv3. Available at: https://docs.openvinotoolkit.org/latest/omz_models_intel_person_vehicle_bike_detection_crossroad_1016_description_person_vehicle_bike_detection_crossroad_1016.html (accessed 31.05.2020)

19. Xu Z., Yang W., Meng A., Lu N., Huang H., Ying C., Huang L. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline. Com-puter Vision, ECCV 2018, 15th Europ. Conf. Munich, Germany, 8–14 Sept. 2018. Berlin: Springer, pp. 261–277. doi: 10.1007/978-3-030-01261-8_16

Information about the author

Ilya G. Zubov, Master of Engineering and Technology (2016), Ltd "Next" algorithm programmer. The author of 6 scientific publications. Area of expertise: digital image processing; applied television systems.
Address: Ltd "Next", 15 Rochdelskaya St., bld. 13, Moscow 123022, Russia
E-mail: ZubovIG@gmail.com