

real-time network connection. The article provides results of TSN Ethernet network analysis in the form of a graph.

Keywords: *TSN, deterministic analysis, network calculus, end-to-end delay, boundary delays, arrival curve, service curve*

DOI: 10.18469/ikt.2023.21.2.06

Sudareva Maria Evgenievna, Samara National Research University, 34, Moskovskoe shosse, Samara, 443086, Russian Federation; Assistant Advanced Aerospace Engineering School. Tel. +7 937 171-12-01. E-mail: masha_sudareva@mail.ru

Yashin Vladimir Nikolaevich, Samara State Technical University, 244, Molodogvardeyskaya Street, Samara, 443100, Russian Federation; Associate Professor of Information Technologies Department, PhD in Technical Science. Tel. +7 987 913-40-52. E-mail: vlyashin@yandex.ru

References

1. Vorobiev S. TSN – time synchronized network. Part 1. *Sovremennye tehnologii avtomatizatsii*, 2020, no. 1, pp. 18–22. (In Russ.)
2. Vorobiev S. TSN – time synchronized network. Part 2. *Sovremennye tehnologii avtomatizatsii*, 2020, no. 2, pp. 22–27. (In Russ.)
3. Roslyakov A.V. et al. Time sensitive networking standardization. *Standarty i kachestvo*, 2021, no. 4, pp. 48–53. (In Russ.)
4. IEEE Standard for Local and Metropolitan Area Networks-Bridges and Bridged Networks. URL: https://standards.ieee.org/standard/802_1Q-2018.html (accessed: 08.11.2023).
5. Bina Sh., Bruckner D., Vasina A.S. OPC UA TSN as a technology for communication at all levels of automation. *Avtomatizatsiya v promyshlennosti*, 2019, no. 2, pp. 26–34. (In Russ.)
6. Improving the worst-case delay analysis of an AFDX network using an optimized trajectory approach. URL: <https://ieeexplore.ieee.org/document/5524098> (accessed: 01.11.2023).
7. Performance impact of the interactions between time-triggered and rate-constrained transmissions in TTEthernet. URL: <https://hal.archives-ouvertes.fr/hal-01255939> (accessed: 03.11.2023).
8. Timing analysis of TDMA-based networks using network calculus and integer linear programming. URL: <https://ieeexplore.ieee.org/document/703363> (accessed: 05.11.2023).
9. Timing analysis of AVB traffic in TSN networks using network calculus. URL: <https://ieeexplore.ieee.org/document/8430063> (accessed: 08.11.2023).
10. Formal worst-case timing analysis of Ethernet TSN's time-aware and peristaltic shapers. URL: <https://ieeexplore.ieee.org/document/7385584> (accessed: 09.11.2023).
11. A calculus for network delay. I. Network elements in Isolation. URL: <https://ieeexplore.ieee.org/document/61109> (accessed: 07.11.2023).
12. Bruckner D. et al. A new Solution for Industrial Communication. URL: <https://www.moxa.com/Moxa/files/66/6669d232-4227-440a-9ddf-477e70b11780.pdf> (accessed: 10.03.2021).

Received 09.11.2023

НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.896

ПРИМЕНЕНИЕ СОВРЕМЕННЫХ АЛГОРИТМОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПОДСЧЕТА ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Альгашев Г.А., Солдатова О.П.

Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, РФ

E-mail: algashev@live.com

Данная статья посвящена исследованию архитектур сверточных нейронных сетей для решения задачи подсчета объектов на изображении. В настоящее время для решения этой задачи применяют набирающие все большую популярность методы, использующие регрессию. В работе для решения задачи подсчета объектов как задачи с использованием регрессии были применены модификации эталонных сверточных нейронных сетей AlexNet, VGG16 и ResNet 50, которые предназначены для классификации изображений. Модификация заключалась в замене второй части нейронной сети, которая классифицирует изображения, на один полносвязный слой, состоящий из одного нейрона без функции активации. В экспериментах измененные архитектуры эталонных сверточных сетей инициализировались двумя способами: с помощью случайной инициализации весов и с помощью заранее подготовленных весов, обученных на наборе данных ImageNet. Представлены результаты экспериментов, в которых подтверждается работоспособность предложенных моделей и применения метода нейропластичности для решения задачи с использованием регрессии. В качестве данных для обучения и тестирования применялась база изображений бактериальных клеток.

Ключевые слова: сверточная нейронная сеть, регрессия, инициализация весов, нейропластичность, подсчет объектов, компьютерное зрение, анализ изображения

Введение

Задача подсчета однотипных объектов на изображении является одной из задач компьютерного зрения. Решение этой задачи необходимо во многих сферах деятельности, например, для подсчета клеток крови на микроскопических изображениях, мониторинга потока людей в системах видеонаблюдения или оценки лесного покрова на аэрофотоснимках. Для решения этих задач часто используют ручной подсчет объектов.

Для автоматизации решения задачи подсчета объектов на изображении используют алгоритмы сегментации изображения [1–5] и алгоритмы детектирования объектов на изображении [6–10]. Однако эти алгоритмы показывают низкую эффективность в случаях, если объекты на изображении соприкасаются или пересекаются, расположены на разном расстоянии от предмета съемки, размыты или имеют малый размер.

В настоящее время большую популярность набирают методы, в которых задача подсчета объектов представляется как задача регрессии. В этих методах строится регрессионная функция зависимости количества объектов в некоторой области изображения от визуальных признаков объектов в данной области. К данным методам относятся алгоритмы, выделяющие структурные признаки [11; 12], градиентные признаки [13; 14; 15] и текстурные признаки [16; 17].

Как правило, для решения задачи регрессии используют нейронные сети, а именно, многослойный персептрон или сеть радиально-базисных функций. Эти сети хорошо работают, если входные данные представляют собой векторы признаков, которые являются упрощенными представлениями реальных данных и имеют невысокую размерность. Однако не все данные можно представить при помощи упрощенного вектора признаков. При работе с изображениями определение наиболее значимых характеристик является слож-

ной задачей, которую не всегда удается решить. Поэтому для работы с изображениями все чаще используют сверточные нейронные сети [18; 19; 20], которые для решения поставленной задачи анализируют изображения целиком.

Обзор современных методов подсчета объектов на изображениях

Существующие алгоритмы, которые решают задачу подсчета объектов на изображении, можно разделить на следующие категории:

- алгоритмы, использующие сегментацию изображения;
- алгоритмы, использующие детектирование объектов на изображении;
- алгоритмы, использующие регрессию.

Алгоритмы, использующие сегментацию изображения, являются самыми распространенными и применяются чаще других для подсчета объектов [1–5]. Идея данных алгоритмов заключается в решении задачи сегментации областей интереса на изображении и их дальнейшего анализа.

Анализ изображений состоит из трех этапов:

1. Сегментация областей изображения, представляющих интерес для анализа.
2. Локализация отдельных объектов на изображении.
3. Оценка количества объектов.

В качестве недостатков данных алгоритмов можно выделить недостаточно высокую эффективность оценки в случае, если объекты на изображении соприкасаются или пересекаются друг с другом, расположены на разном расстоянии от предмета съемки, размыты или имеют малый размер.

Алгоритмы детектирования объектов, используют визуальный детектор объектов, который локализует отдельные объекты на изображении [8–12]. После локализации всех объектов на изображении не составляет труда посчитать их количество. Такие методы ориен-

тированы на изображения, где объекты удалены друг от друга и являются однородными. Проблемы алгоритмов детектирования аналогичны проблемам алгоритмов сегментации.

При подсчете количества объектов с использованием методов регрессии [13–19] задача локализации объектов на изображении не решается. Вместо этого алгоритмы анализируют изображение на основе его глобальных характеристик, в результате чего система обучается выдавать оценку количества объектов. Стоит отметить, что методы регрессии не используют для обучения информацию о местонахождении объектов, а используют только информацию об их количестве. Для обучения таких систем требуется большое количество заранее подготовленных обучающих данных.

Модели сверточных нейронных сетей для решения задачи регрессии

Для решения задачи подсчета объектов на изображении как задачи регрессии в работе использовались известные архитектуры сверточных нейронных сетей [20], которые предназначены для обработки и анализа изображений.

Для разработки собственной архитектуры нейронной сети, подбора параметров сети и ее обучения может потребоваться много времени и вычислительных ресурсов. Кроме того, при таком подходе нельзя использовать эталонные модели, обученные на одинаковых стандартных наборах изображений, и как следствие, сравнение полученных результатов затруднено. Поэтому чаще всего для решения задач используют эталонные модели сетей с натренированными весами, которые затем модифицируют для решения конкретной задачи. Это позволяет сэкономить время на настройку весов сети, так как эталонная модель предварительно обучена и во время повторного обучения на новом наборе данных, веса будут только незначительно корректироваться. При необходимости, данные модели можно обучать с нуля случайно инициализированными весами, не используя натренированные веса.

Для решения задачи в работе были использованы три эталонных архитектуры сверточных нейронных сетей, а именно сеть AlexNet [21], VGG16 [22] и ResNet50 [23]. Использование этих сетей обусловлено тем, что модели можно инициализировать весами, обученными на одинаковом наборе данных ImageNet, для решения поставленной задачи. Данная особенность позволит сравнивать результаты обучения моделей в случае их инициализации натренированными весами.

Представленные эталонные модели в неизменном виде не могут быть использованы для решения задачи регрессии. Архитектура каждой модели сети состоит из двух частей:

1. Первая часть выделяет характерные признаки изображений и состоит из сверточных и подвыборочных слоев.

2. Вторая часть выполняет классификацию изображения по выделенным на предыдущих слоях признакам, и состоит из полносвязных слоев. Последний полносвязный слой использует функцию активации SoftMax [24].

Для решения новой задачи регрессии были изменены вторые части всех трех моделей, в которых все полносвязные слои были заменены на один полносвязный слой нейронов без функции активации. Этот полносвязный слой будет состоять только из одного нейрона, который будет выдавать на выходе оценку количества объектов на изображении.

Измененные архитектуры сверточных сетей AlexNet, VGG16 и ResNet50 представлены на рисунке 1.

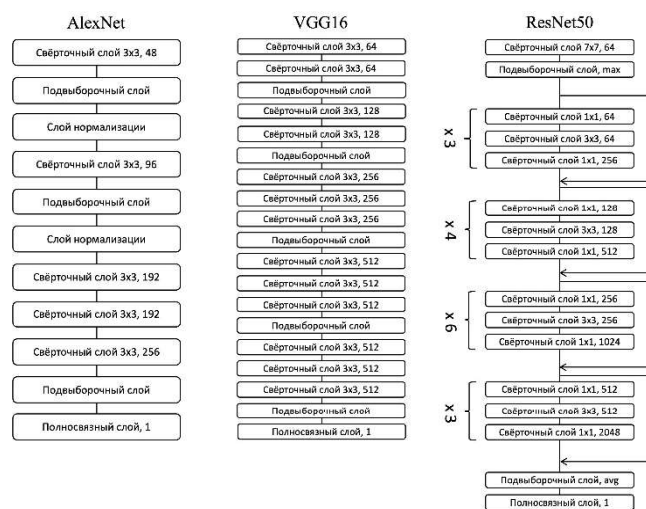


Рисунок 1. Архитектуры сетей AlexNet, VGG16 и ResNet50 для решения задачи регрессии

Данные для обучения и тестирования

Для решения задачи подсчета объектов на изображениях в качестве данных для обучения и тестирования была взята база изображений бактериальных клеток [25]. Набор состоит из изображений, на которых содержатся бактериальные клетки, подсвеченные с помощью люминесцентной микроскопии (рисунок 2а), и изображений, которые являются аннотациями к исходным данным (рисунок 2б). На аннотациях выделены центры бактериальных клеток исходных изображений.

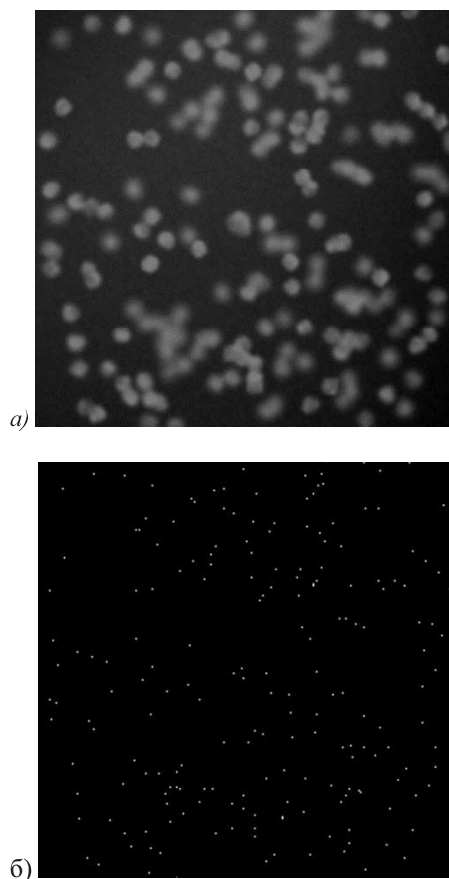


Рисунок 2. Пример изображения, содержащего бактериальные клетки, и аннотация к изображению

Набор данных состоит из 200 исходных изображений и 200 аннотаций к ним. Каждое изображение имеет размер 256 на 256 пикселей и является цветным. В исходных изображениях может содержаться от 74 до 317 объектов.

Анализировать исходные изображения целиком сложно из-за их высокого разрешения, поэтому исходные изображения и аннотации были разделены на фрагменты размером 32 на 32 пикселей, которые и подавались на вход сети.

Выбор размера фрагментов изображений обусловлен следующими особенностями. При увеличении размера фрагментов время обучения сверточных нейронных сетей и значение погрешности будет увеличиваться. При уменьшении размера фрагментов время обучения и значение погрешности будет уменьшаться. Но в таком случае для обработки исходного изображения потребуется посылать большее количество фрагментов, что увеличит время анализа изображения.

Обучение моделей сетей происходило в облачном сервисе Google Colaboratory, который имеет следующие ограничения:

- использование только одного графического процессора для обучения сетей;
- объем доступной видеопамати изменяется со временем и ограничен 12 Гб;

- максимальное время сеанса не превышает 12 часов;

- высокая вероятность аварийного завершения сеанса обучения при превышении лимита памяти или наличии нестабильного интернет-соединения.

Наличие данных ограничений не позволяет производить длительные и высокопроизводительные вычисления для обучения моделей сетей.

На основе исходной базы данных были сгенерированы 10 000 фрагментов изображений, содержащих бактериальные клетки, и 10 000 фрагментов аннотаций к ним. Для проведения экспериментов использовано 9 000 фрагментов для обучения, 500 фрагментов для промежуточной проверки работы сети во время обучения и 500 изображений для тестирования.

Перед использованием в сети фрагменты аннотаций были преобразованы в числовой вектор, в котором записано количество объектов для каждого фрагмента изображения путем суммирования выделенных пикселей. Фрагменты изображений были нормализованы перед подачей на вход сети.

Нормализация входных данных заключалась в делении значения каждого пикселя на 255 для каждого канала RGB, с целью преобразования данных в диапазон от 0 до 1.

Обучение моделей сверточных сетей и анализ результатов экспериментов

Для разработки и обучения программной реализации моделей сверточных нейронных сетей был использован язык высокого уровня Python и нейросетевые библиотеки TensorFlow и Keras, которые предназначены для работы с данными и нейронными сетями.

Для исследования были проведены две серии экспериментов:

1. Веса сверточных слоев инициализировались случайным образом, и модели обучались с нуля.
2. Веса сверточных слоев инициализировались тренированными весами, обученными на наборе данных ImageNet, и происходило дообучение сетей.

В качестве алгоритма обучения использовался метод адаптивной инерции (Adam) [26]. Метод вычисляет адаптивные скорости обучения для каждого веса сети из оценок первого и второго моментов градиентов. Алгоритм вычисляет экспоненциальную скользящую среднюю градиента и квадрата градиента, а параметры β_1 и β_2 управляют скоростью затухания скользящих средних.

Правила перерасчета метода:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \Delta L_t(w_{t-1}),$$

$$v_t = \beta_2 g_{t-1} + (1 - \beta_2) \Delta L_t(w_{t-1}),$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t},$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t},$$

$$w_t = w_{t-1} = \frac{\eta \hat{m}_t}{\sqrt{\hat{v}_t + \varepsilon}},$$

где L – функция погрешности, m_t – оценка первого момента, а v_t – оценка второго момента. Скорректированные оценки $\hat{m}_t$ и $\hat{v}_t$ используются для предотвращения долгого накопления m_t и v_t .

Метод имеет следующие задаваемые параметры:

- η – коэффициент скорости обучения;
- β_1 – коэффициент - показатель экспоненциальной скорости затухания для первого момента;
- β_2 – коэффициент - показатель экспоненциального спада для оценок второго момента;
- ε – коэффициент сглаживания, для предотвращения деления на ноль.

Для обучения моделей сетей были заданы параметры $\eta = 0,0001$, $\beta_1 = 0,9$, $\beta_2 = 0,999$, $\varepsilon = 0,0000001$.

В качестве функции погрешности используется евклидово расстояние и средняя квадратическая ошибка.

Евклидово расстояние представляет собой геометрическое расстояние в многомерном пространстве:

$$L = \sqrt{\sum_i^n (y_i - y'_i)^2},$$

где n – количество нейронов в выходном слое сети, y – ожидаемое значение выходного сигнала нейрона, y' – фактическое значение выходного сигнала нейрона.

Средняя квадратическая ошибка – оценка среднего квадратичного отклонения, которая вычисляется по формуле:

$$L = \frac{1}{n} \sum_i^n (y_i - y'_i)^2.$$

Использование этих функций погрешностей обусловлено тем, что они являются самыми распространенными при решении задачи регрессии.

В таблице 1 и таблице 2 приведены значения погрешностей обучения и обобщения сетей AlexNet, VGG16 и ResNet для каждой 10-ой эпохи обучения в первой серии экспериментов. По-

грешность обучения рассчитывалась на основе данных из обучающей выборки, а погрешность обобщения рассчитывалась на основе данных для промежуточной проверки, которые не участвовали в обучении сетей.

Анализ результатов экспериментов показал, что значение погрешности обучения во всех трех моделях уменьшалось в процессе обучения и к концу обучения сетей достигло значения 0,1045 (0,005) для сети AlexNet, значения 0,1442 (0,001) для сети VGG16 и значения 0,2280 (0,0026) для сети ResNet50. Значение погрешности обобщения напротив, уменьшалось неравномерно и скачкообразно и к концу обучения достигло значения 0,3844 (0,0074) для сети AlexNet, значения 0,3117 (0,0049) для сети VGG16 и значения 0,4489 (0,0101) для сети ResNet50.

На тестовой выборке модели показали значение погрешности 0,4313 (0,0093) для сети AlexNet, значение 0,2378 (0,0028) для сети VGG16 и значение 0,4491 (0,0101) для сети ResNet50. По этим данным можно сделать вывод, что сеть VGG16 показала лучшие результаты на выборке для промежуточной проверки и на тестовой выборке.

Для второй серии экспериментов веса моделей инициализировались подготовленными весами, которые были получены путем обучения моделей на наборе данных ImageNet [27]. Идея использовать подготовленные веса в сети основана на свойстве нейропластичности сверточных сетей. Под нейропластичностью понимают использование подготовленных весов для дообучения сети на новом наборе данных, таким образом, модели изменяют свои веса незначительно для решения новой задачи [28; 29].

В таблице 3 и таблице 4 приведены значения погрешностей обучения и обобщения сетей AlexNet, VGG16 и ResNet для каждой 10-ой эпохи обучения во второй серии экспериментов.

Аналогично первой серии экспериментов, значение погрешности обучения во всех моделях равномерно уменьшалось во время обучения, а значение погрешности обобщения уменьшалось неравномерно. Отсюда можно сделать вывод, что характер изменения погрешностей во всех моделях не изменился.

К завершению обучения значения погрешности обучения и погрешности обобщения составляли 0,1129 (0,0006) и 0,2607 (0,0034) соответственно, для сети AlexNet, значения 0,1198 (0,0007) и 0,1943 (0,0019) для сети VGG16, и значения 0,1346 (0,0009) и 0,2731 (0,0037) для сети ResNet50. На тестовой выборке значение погрешности составляет 0,2913 (0,0042) для сети

Таблица. 1. Результаты обучения сверточных моделей при случайной инициализации весов и использовании в качестве функции погрешности евклидового расстояния

Номер эпохи	AlexNet		VGG16		ResNet50	
	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения
10	0,3461	0,3892	0,3832	0,5003	0,8559	0,9491
20	0,2412	0,4329	0,2868	0,3862	0,7758	5,7209
30	0,1827	0,3323	0,2220	0,5605	0,7582	0,7123
40	0,1583	0,5207	0,1994	0,3665	0,5825	1,2585
50	0,1461	0,3086	0,1905	0,5107	0,3982	0,3027
60	0,1322	0,3255	0,1704	0,6531	0,4080	1,4437
70	0,1234	0,3956	0,1857	0,6504	0,3173	0,5217
80	0,1125	0,5153	0,1612	0,2137	0,2683	0,3210
90	0,1073	0,3841	0,1535	0,4938	0,2153	0,3861
100	0,1045	0,3844	0,1442	0,3117	0,2280	0,4489

Таблица. 2. Результаты обучения сверточных моделей при случайной инициализации весов и использовании в качестве функции погрешности средней квадратичной ошибки

Номер эпохи	AlexNet		VGG16		ResNet50	
	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения
10	0,0060	0,0076	0,0073	0,0125	0,0366	0,0450
20	0,2412	0,0094	0,0041	0,0075	0,0301	1,6364
30	0,0017	0,0055	0,0025	0,0157	0,0287	0,0254
40	0,0013	0,0136	0,0020	0,0067	0,0170	0,0792
50	0,0011	0,0048	0,0018	0,0130	0,0079	0,0046
60	0,0009	0,0053	0,0015	0,0213	0,0083	0,1042
70	0,0008	0,0078	0,0017	0,0212	0,0050	0,0136
80	0,0006	0,0133	0,0013	0,0023	0,0036	0,0052
90	0,0006	0,0074	0,0012	0,0122	0,0023	0,0075
100	0,0005	0,0074	0,0010	0,0049	0,0026	0,0101

Таблица. 3. Результаты обучения сверточных моделей при инициализации сетей подготовленными весами и использовании в качестве функции погрешности евклидового расстояния

Номер эпохи	AlexNet		VGG16		ResNet50	
	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения
10	0,3724	0,3958	0,3731	0,3641	0,6255	0,7265
20	0,2629	0,4676	0,2967	0,2459	0,6379	0,6914
30	0,2008	0,4127	0,2644	0,2600	0,4918	0,9943
40	0,1738	0,3631	0,2490	0,4613	0,3812	0,4514
50	0,1535	0,2465	0,1869	0,5286	0,3298	0,4314
60	0,1438	0,4144	0,1773	0,3217	0,2823	0,4678
70	0,1322	0,4340	0,1965	0,2679	0,2271	0,3874
80	0,1246	0,4773	0,1595	0,6322	0,2037	0,4748
90	0,1199	0,4488	0,1436	0,3281	0,1702	0,3301
100	0,1129	0,2607	0,1198	0,1943	0,1346	0,2731

Таблица. 4. Результаты обучения сверточных моделей при инициализации сетей подготовленными весами и использовании в качестве функции погрешности средней квадратичной ошибки

Номер эпохи	AlexNet		VGG16		ResNet50	
	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения	Погрешность обучения	Погрешность обобщения
10	0,0069	0,0078	0,0070	0,0066	0,0196	0,0264
20	0,0035	0,0109	0,0044	0,0030	0,0203	0,0239
30	0,0020	0,0085	0,0035	0,0034	0,0121	0,0494
40	0,0015	0,0066	0,0031	0,0106	0,0073	0,0102
50	0,0012	0,0030	0,0017	0,0140	0,0054	0,0093
60	0,0010	0,0086	0,0016	0,0052	0,0040	0,0109
70	0,0009	0,0094	0,0019	0,0036	0,0026	0,0075
80	0,0008	0,0114	0,0013	0,0200	0,0021	0,0113
90	0,0007	0,0101	0,0010	0,0054	0,0014	0,0054
100	0,0006	0,0034	0,0007	0,0019	0,0009	0,0037

AlexNet, 0,1489 (0,0011) для сети VGG16 и 0,3923 (0,0076) для сети ResNet50.

Заключение

Предложенные модификации архитектур сверточных нейронных сетей позволили эффективно решить задачу подсчета объектов на изображении как задачу регрессии. По результатам анализа экспериментов в первой и второй серии, можно сделать вывод, что все модели сетей показали лучший результат на данных для промежуточной проверки и на тестовых данных при начальной инициализации сетей подготовленными весами. Экспериментально было показано, что применение свойства нейропластичности в сверточных сетях позволяет уменьшить значение погрешности при решении задачи регрессии.

В результате проведения исследований выяснилось, что самое большое значение погрешности тестирования 0,4491 (0,0101), показала модель сети ResNet50, обученная со случайно инициализированными весовыми коэффициентами.

Для решения поставленной задачи подсчета однотипных объектов на изображении лучшей является модификация модели нейронной сети VGG16 для решения задачи регрессии, инициализируемая весами, подготовленными для решения задачи классификации на наборе ImageNet. Такая модель показала значение погрешности на тестовых данных 0,1489 (0,0011), что в 3 раза меньше наихудшего результата сети ResNet50.

Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
2. Segmentation and Classification of SAR Data with Co-Occurrence Matrix for Texture Features / N.G. Kasapoglu [et al.] // Proceedings EUSAR 2002. Germany: Cologne, 2002. P. 717–720.
3. Fang S., Wen H., Shiyi M. Classification of SAR images Based on Simplified Segmentation // Proceedings EUSAR 2002. Germany: Cologne, 2002. P. 705–708.
4. Миронов Б.М., Малов А.Н. Сегментация изображений кластерным методом и алгоритмом случайных скачков: сравнительный анализ // Компьютерная оптика. 2010. Т. 34, №1. С. 132–137.
5. Белим С.В., Ларионов С.Б. Алгоритм сегментации изображений, основанный на поиске сообществ на графах // Компьютерная оптика. 2016. Т. 40, № 6. С. 904–910.
6. Shneier M. Road sign detection and recognition // Proceedings IEEE Computer Society International Conference on Computer Vision and Pattern Recognition. 2005. P. 215–222.
7. Распознавание тест-объектов на тепловизионных изображениях / А.В. Мингалев [и др.] // Компьютерная оптика. 2019. Т. 43, № 3. С. 402–411.
8. Viola P., Jones M. Robust Real-time Object Detection // International Journal of Computer Vision. 2001. no. 57(2). P. 137–154.
9. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // IEEE Conference on Computer Vision and Pattern Recognition. 2001. Vol. 1. P. 511–518.
10. Rich feature hierarchies for accurate object detection and semantic segmentation / R. Girshick [et al.] // 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 580–587.
11. Estimating pedestrian counts in groups / P. Kilambi [et al.] // Computer Vision and Image Understanding. 2008. Vol. 110. P. 43–59.
12. Crowd counting using group tracked and local features / D. Ryan [et al.] // 7th IEEE International Conference on Advanced Video and Signal Based Surveillance. 2010. P. 218–224.
13. Estimating crowd density with Minlowski fractal dimension / A.N. Marana [et al.] // 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. 1999. Vol. 6. P. 3521–3524.
14. Feature mining for localized crowd counting / K. Chen [et al.] // Proceedings of British Machine Conference. 2012. Vol. 21. P. 1–11.
15. Ma W., Huang L., Liu C. Crowd density analysis using co-occurrence texture features // Proceedings of International Conference on Computer Sciences and Convergence Information Technologies. 2010. P. 170–175.
16. Automatic estimation of crowd density using texture / A.N. Marana [et al.] // Safety Science. 1998. Vol. 28. P. 165–175.
17. Rahmalan H., Nixon M., Carter J. On crowd density estimation for surveillance // The Institution of Engineering and Technology Conference on Crime and Security. 2006. P. 540–545.
18. Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun [et al.] // Neural Computation. 1989. Vol. 1, no. 4. P. 541–551.
19. Cell Counting by Regression Using Convolutional Neural Network / Y. Xue [et al.]. Springer International Publishing, 2016. 17 p.
20. Deep Spatial Regression Model for Image Crowd Counting / H. Yao [et al.]. URL: <https://arxiv.org/pdf/1710.09757.pdf> (дата обращения: 26.08.2020).

21. Krizhevsky A., Sutskever I., Hinton G.E. Image Net classification with deep convolution neural networks // Advances in neural information processing systems (NIPS). 2012. P. 1097–1105.
22. Accelerating very deep convolutional networks for classification and detection / X. Zhang [et al.] // IEEE transactions on pattern analysis and machine intelligence. 2016. no. 38. P. 1943–1955.
23. Deep Residual Learning for Image Recognition / K. He [et al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778.
24. Bishop C.M. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, 2006. 738 p.
25. Lempitsky V., Zisserman A. Learning to count objects in images // Neural Information Processing Systems Foundation. 2010. P. 1324–1332.
26. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization // International Conference on Learning Representations. 2015. P. 1–13.
27. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolution neural networks // Advances in neural information processing systems. 2012. P. 1097–1105.
28. Альгашев Г.А., Солдатова О.П. Нейропластичность сверточных нейронных сетей // Перспективные информационные технологии (ПИТ-2018): материалы Международной научно-технической конференции. Самара: СНЦ РАН, 2018. С. 385–389.
29. Альгашев Г.А., Солдатова О.П. Исследование эффективности использования свойства нейропластичности в сверточных сетях // Информационные технологии и нанотехнологии (ИТНТ-2019): материалы V Международной конференции и молодежной школы. Самара: Новая техника, 2019. С. 711–720.

Получено 07.11.2023

Альгашев Геннадий Андреевич, ассистент кафедры информационных систем и технологий (ИСТ) Самарского национального исследовательского университета имени академика С.П. Королева (Самарский университет). 443086, Российская Федерация, г. Самара, ул. Московское шоссе, 34. Тел. +7 927 722-48-73. E-mail: algashev@live.com

Солдатова Ольга Петровна, к.т.н., доцент кафедры ИСТ Самарского университета. 443086, Российская Федерация, г. Самара, ул. Московское шоссе, 34. Тел. +7 927 654-29-33. E-mail: op-soldatova@yandex.ru

APPLICATION OF MODERN COMPUTER VISION ALGORITHMS TO MANAGE WITH THE COUNTING OF IMAGE OBJECTS

Algashev G.A., Soldatova O.P.

Samara National Research University, Samara, Russian Federation

E-mail: algashev@live.com

This article is devoted to the research of convolutional neural network architectures in counting objects in an image. Currently, methods using regression are gaining popularity. In this article in order to solve an object counting task using regression method, modifications of the reference convolutional neural networks AlexNet, VGG16, and ResNet 50, which were developed for image classification, were used. Modification presented by replacing the second part of the neural network, which classifies images, with one fully connected layer, consisting of one neuron without activating function. In experiments, modified architectures of the reference convolutional networks were initialized as follows: using random initialization of the weights and using pretrained weights trained on the ImageNet dataset. The results of experiments, which confirm the performance of the proposed models and the use of the neuroplasticity method to solve the problem using regression are presented. The database of images of bacterial cells was used as training and testing material.

Keywords: *convolutional neural network, regression, initialization of weights, neuroplasticity, object counting, computer vision, image analysis*

DOI: 10.18469/ikt.2023.21.2.07

Algashev Gennady Andreevich, Samara National Research University, 34, Moskovskoe shosse,

Samara, 443086, Russian Federation; Teacher of Information Systems and Technologies Department. Tel. +7 927 722-48-73. E-mail: algashev@live.com

Soldatova Olga Petrovna, Samara National Research University, 34, Moskovskoe shosse, Samara, 443086, Russian Federation; Associated Professor of Information Systems and Technologies Department, PhD in Technical Science. Tel. +7 927 654-29-33. E-mail: op-soldatova@yandex.ru

References

1. Gonzalez R. *Digital image processing*. Moscow: Tekhnosfera, 2005, 1072 p.
2. Kasapoglu N.G. et al. Segmentation and Classification of SAR Data with Co-Occurance Matrix for Texture Features. *Proceedings EUSAR 2002*, Germany: Cologne, 2002, pp. 717–720.
3. Fang S., Wen H., Shiyi M. Classification of SAR images Based on Simplified Segmentation. *Proceedings EUSAR 2002*, Germany: Cologne, 2002, pp. 705–708.
4. Mironov B.M., Malov A.N. Image segmentation by the cluster method and the random hopping algorithm: a comparative analysis. *Komp'yuternaya optika*, 2010, vol. 34, no. 1, pp. 132–137. (In Russ.)
5. Belim S.V., Larionov S.V. An algorithm of image segmentation based on community detection in graphs. *Komp'yuternaya optika*, 2016, vol. 40, no. 6, pp. 904–910. (In Russ.)
6. Shneier M. Road sign detection and recognition. *Proceedings IEEE Computer Society International Conference on Computer Vision and Pattern Recognition*, 2005, pp. 215–222.
7. Mingalev A.V. et al. Test-object recognition in thermal images. *Komp'yuternaya optika*, 2019, vol. 43, no. 3, pp. 402–411. (In Russ.)
8. Viola P., Jones M. Robust Real-time Object Detection. *International Journal of Computer Vision*, 2001, no. 57(2), pp. 137–154.
9. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. *IEEE Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 511–518.
10. Girshick R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587.
11. Kilambi P. et al. Estimating pedestrian counts in groups. *Computer Vision and Image Understanding*, 2008, vol. 110, pp. 43–59.
12. Ryan D. et al. Crowd counting using group tracked and local features. *7th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2010, pp. 218–224.
13. Marana A.N. et al. Estimating crowd density with Minlowski fractal dimension. *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1999, vol. 6, pp. 3521–3524.
14. Chen K. et al. Feature mining for localized crowd counting. *Proceedings of British Machine Conference*, 2012, vol. 21, pp. 1–11.
15. Ma W., Huang L., Liu C. Crowd density analysis using co-occurrence texture features. *Proceedings of International Conference on Computer Sciences and Convergence Information Technologies*, 2010, pp. 170–175.
16. Marana A.N. et al. Automatic estimation of crowd density using texture. *Safety Science*, 1998, vol. 28, pp. 165–175.
17. Rahmalan H., Nixon M., Carter J. On crowd density estimation for surveillance. *The Institution of Engineering and Technology Conference on Crime and Security*, 2006, pp. 540–545.
18. LeCun Y. et al. Backpropagation Appiled to Handwritten Zip Code Recognition. *Neural Computation*, 1989, vol. 1, no. 4, pp. 541–551.
19. Xue Y. et al. *Cell Counting by Regression Using Convolutional Neural Network*. Springer International Publishing, 2016. 17 p.
20. Yao H. et al. Deep Spatial Regression Model for Image Crowd Counting. URL: <https://arxiv.org/pdf/1710.09757.pdf> (accessed: 26.08.2020).
21. Krizhevsky A., Sutskever I., Hinton G.E. Image Net classification with deep convolution neural networks. *Advances in neural information processing systems (NIPS)*, 2012, pp. 1097–1105.

22. Zhang X. et al. Accelerating very deep convolutional networks for classification and detection. *IEEE transactions on pattern analysis and machine intelligence*, 2016, no. 38, pp. 1943–1955.
23. He K. et al. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
24. Bishop C.M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 2006, 738 p.
25. Lempitsky V., Zisserman A. Learning to count objects in images. *Neural Information Processing Systems Foundation*, 2010, pp. 1324–1332.
26. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*, 2015, pp. 1–13.
27. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolution neural networks. *Advances in neural information processing systems*, 2012, pp. 1097–1105.
28. Algashev G.A., Soldatova O.P. Neuroplasticity of convolutional neural networks. *Perspektivnye informacionnye tekhnologii (PIT-2018): materialy Mezhdunarodnoj nauchno-tekhnicheskoj konferencii*, Samara: SNTs RAN, 2018, pp. 385–389. (In Russ.)
29. Algashev G.A., Soldatova O.P. Study of the effectiveness of using the property of neuroplasticity in convolutional. *Informacionnye tekhnologii i nanotekhnologii (ITNT-2019): materialy V Mezhdunarodnoj konferencii i molodezhnoj shkoly*. Samara: Novaya tekhnika, 2019, pp. 711–720. (In Russ.)

Received 07.11.2023

УДК 004.942

ИТЕРАТИВНЫЕ АДАПТАЦИОННЫЕ АЛГОРИТМЫ В МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧАХ

Бельгер И.С., Солдатова О.П.

Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, РФ
E-mail: irina.belger@mail.ru, op-soldatova@yandex.ru

Данная статья рассматривает использование итеративных адаптационных алгоритмов для решения задачи определения места измерения комплекса интима-медиа сонной артерии. Приведена постановка многокритериальной задачи принятия решений, на основании которой происходит определение корректного критерия для выбора и успешного распознавания требуемого объекта на ультразвуковом снимке. В работе рассматриваются принципы построения каскадных классификаторов, применение каскадного классификатора Хаара и каскадного классификатора LBP (Local Binary Patterns), в качестве основы для которых использованы примитивы Хаара и локальные бинарные шаблоны. Приведены результаты экспериментальных исследований эффективности использования разных алгоритмов бустинга для решения данной задачи. Лучшие результаты показал каскадный классификатор Хаара, построенный с помощью итеративного адаптационного алгоритма, который успешнее справляется с решением многокритериальной задачи на данной обучающей выборке и определяет наиболее подходящие участки для измерения толщины комплекса интима-медиа общей сонной артерии.

Ключевые слова: адаптационные алгоритмы, бустинг, многокритериальная задача, каскадный классификатор, каскад Хаара, каскад LBP, признаки Хаара, Adaboost, локальные бинарные паттерны

Введение

Необходимость распознавания изображений возникает во множестве областей – военное дело, медицина, системы безопасности, оцифровка аналоговых сигналов. Одним из вариантов распознавания образов является детектирование объектов на изображениях, то есть определение наличия объекта на изображении и нахождение его положения в системе координат пикселей исходного изображения. Для того, чтобы корректно

определить наличие объекта на изображении и найти его положение, необходимо составить множество критериев, на основании которых будет производиться поиск объекта. Однако характеристики критериев имеют сильную зависимость от выбора алгоритма детектирования, поскольку положение объекта может определяться координатами прямоугольника, окаймляющего объект, либо контуром этого объекта, либо координатами точек, наиболее характерных для объекта.