



### Литература

- 1 Агафонова Ю.Д. Локализация области патологических изменений на изображениях МРТ головного мозга / Ю.Д. Агафонова, А.В. Гайдель // Перспективные информационные технологии (ПИТ 2019). – Самара, 2019. – С. 362-365.
- 2 He, K. Convolutional neural networks at constrained time cost / K. He, J. Sun // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 5353-5360.
- 3 Davis J., Goadrich M.. The Relationship Between Precision-Recall and ROC Curves, ICML 2006Gaidel, A. Method of automatic ROI selection on lung CT images / A. Gaidel // Procedia Engineering. – 2017. – Vol. 201. – P. 258-264.

Г.А. Альгашев, О.П. Солдатова

## РЕШЕНИЕ ЗАДАЧИ ПОДСЧЁТА КОЛИЧЕСТВА ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ С ПОМОЩЬЮ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

(Самарский университет)

Задача подсчёта объектов на изображении является одной из задач систем компьютерного зрения. Её суть состоит в определении или предсказывании количества однотипных объектов на изображении. Эта задача часто встречается в самых различных сферах, начиная от промышленности, заканчивая медициной.

Оценка количества объектов на изображении можно отнести к задачам обработки изображений. Сейчас большой популярностью обладают методы обработки изображений с помощью свёрточных нейронных сетей. Изначально с помощью них решали задачу классификации изображений, но сейчас свёрточные модели применяют и для решения задачи регрессии.

Свёрточная нейронная сеть — специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году и нацеленная на эффективное распознавание образов, входит в состав технологий глубокого обучения [1]. Сеть использует некоторые особенности зрительной коры, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток. Таким образом, идея свёрточных нейронных сетей заключается в чередовании свёрточных слоёв и слоёв подвыборки. Структура сети — однонаправленная (без обратных связей), принципиально многослойная. Для обучения используются стандартные методы, чаще всего метод обратного распространения ошибки. Функция активации нейронов (передаточная функция) — любая, по выбору исследователя.

Название архитектуры сеть получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу



(ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.

Архитектура сети представлена на рисунке 1.

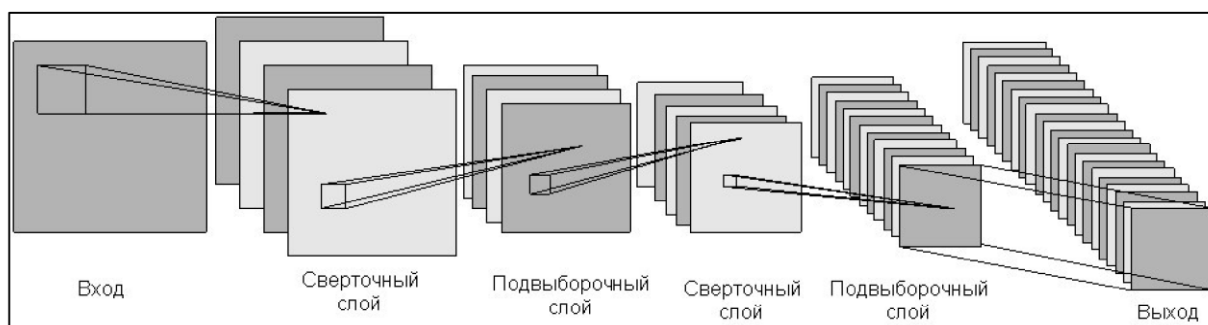


Рис. 1. Архитектура свёрточной нейронной сети

Для решения поставленной задачи было решено использовать глубокую свёрточную нейронную сеть и свести задачу подсчёта объектов на изображении к решению задачи регрессии, т.е. предсказывать значение числовой переменной на основе поступающих на вход данных об изображении (предикторов).

Для решения поставленной задачи в качестве данных для обучения была взята база изображений, содержащие бактериальные клетки, подсвеченные с помощью люминесцентной микроскопии в количестве 200 экземпляров (рисунок 2) и набор изображений, представляющий собой аннотацию к исходным данным, в которых выделены центры бактериальных клеток на изображении (рисунок 3) [2]. Исходные данные представляют из себя цветные изображения размеров 256 на 256 пикселей. Для обучения и тестирования сети было решено выделить на изображениях участки размеров 32 на 32 и подавать их на вход свёрточной нейронной сети.

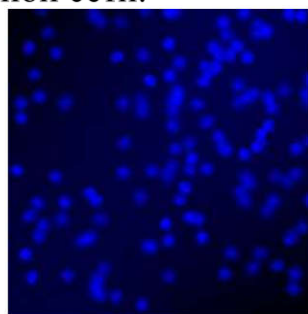


Рис. 2. Изображение, содержащее бактериальные клетки

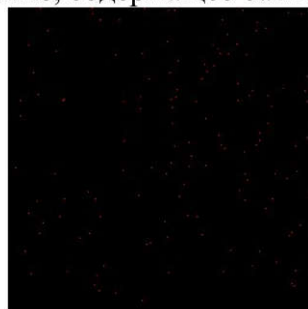


Рис. 3. Аннотация к исходным данным



Для обучения и тестирования сети было решено взять 1000 изображений размером 32 на 32 и разделить их в соотношении 80% обучающих данных и 20% данных для тестирования или 800 и 200 изображений соответственно.

Перед использованием в сети исходные данные и аннотации к ним были нормализованы.

В качестве модели нейронной сети было решено взять свёрточную нейронную сеть VGG16 [3]. Эта сеть характерна своей простой организацией, использует свёрточные слои с размером ядра свёртки 3x3, которые идут друг за другом. Уменьшение размерности производится слоями подвыборки, в которых происходит выбор максимального значения из квадрата 2x2. Затем следуют полносвязные слои, каждый из которых состоит из 4096 нейронов. Последний слой, который реализован с функцией активации SoftMax для классификации объекта, было решено заменить на слой, состоящий из одного нейрона, с помощью которого будет решаться задача регрессии. Номер 16 в названии модели означает количество слоёв, из которых состоит рассматриваемая модель. Сеть обучалась в течении 100 эпох. Архитектура сети представлена на рисунке 4.

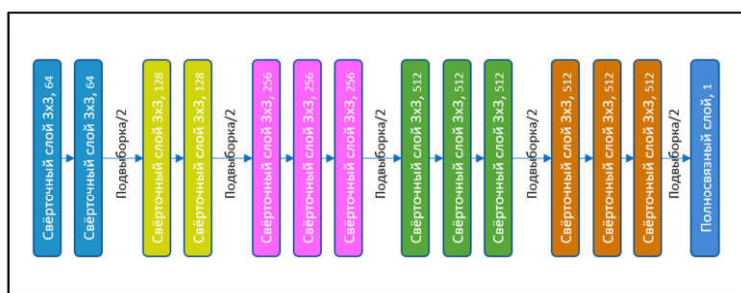


Рис. 4. Архитектура обучающей нейронной сети

В качестве метрики было решено использовать евклидово расстояние [4]. Это наиболее распространенная функция расстояния, которая представляет собой геометрическое расстояние в многомерно пространстве:

$$\rho(x, \hat{x}) = \sqrt{\sum_i^n (x_i - \hat{x}_i)^2}$$

В качестве оптимизатора используется метод адаптивной инерции [5]. Это оптимизационный алгоритм, который сочетает в себе идею накопления движения и идею слабого обновления весов для типичных признаков. Оценка первого момента вычисляется как скользящее среднее. В методе оценки первого и второго моментов инициализируются нулями, в результате чего используется небольшая корректировка, чтобы результирующие оценки не были смещены к нулю. Метод инвариантен к масштабированию градиентов.

В методе используются следующие параметры:



- 1)  $Y_1$ , и  $Y_2$  – окна накопления, могут принимать значения от 0 включительно до 1;
- 2)  $m_t$  – оценка математического ожидания на шаге  $t$ ;
- 3)  $g_t$  – оценка средней нецентрированной дисперсии на шаге  $t$ ;
- 4)  $\hat{m}_t$  и  $\hat{g}_t$  – скорректированные значения оценок математического ожидания и средней нецентрированной дисперсии на шаге  $t$ ;
- 5)  $\eta$  – скорость обучения;
- 6)  $\epsilon$  – коэффициент сглаживания.

При заданных параметрах  $Y_1$ ,  $Y_2$ ,  $\eta$ ,  $m_0=0$  и  $g_0=0$  правило перерасчёта следующее:

$$m_{t+1} = Y_1 m_t + (1 - Y_1) \nabla f_i(\theta_t),$$

$$g_{t+1} = Y_2 g_t + (1 - Y_2) \nabla f_i(\theta_t)^2,$$

$$\hat{m}_{t+1} = \frac{m_{t+1}}{1 - Y_1^{t+1}},$$

$$\hat{g}_{t+1} = \frac{g_{t+1}}{1 - Y_2^{t+1}},$$

$$\theta_{t+1} = \theta_t - \frac{\eta \hat{m}_{t+1}}{\sqrt{\hat{g}_{t+1} + \epsilon}}$$

В обучаемой модели скорость обучения  $\eta=0,00001$ ,  $\epsilon=10^{-8}$ , параметры  $Y_1=0,9$ ,  $Y_2=0,999$ .

На рисунке 5 изображен график изменения евклидовой метрики сети после каждой эпохи обучения. На рисунке 6 изображён график изменения достоверности сети после каждой эпохи обучения. Под достоверностью понимают долю правильно распознанных данных, подаваемых на вход сети.

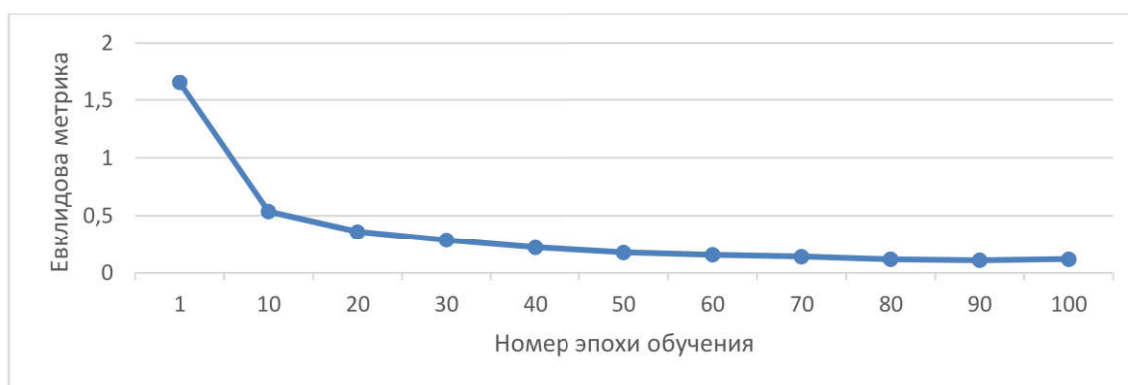


Рис. 5. График изменения евклидовой метрики сети

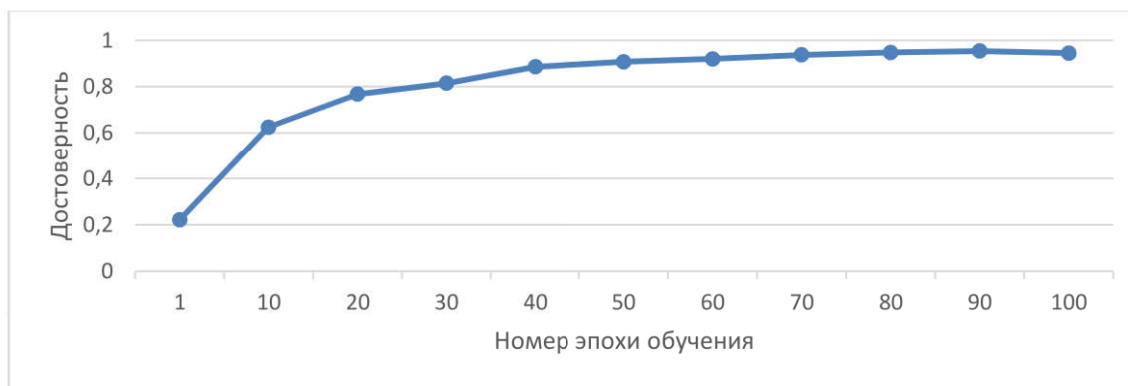


Рис. 6. График изменения достоверности сети

На тестовой выборке модель выдала данные с достоверностью в 0,73.

Как видно из графиков, обученная свёрточная нейронная сеть показала хороший результат в решении задачи регрессии. С увеличением количества эпох обучения сеть показывается всё большую достоверность и достигает своего максимума в диапазоне 90 – 100 эпох. Таким образом, используемый метод можно использовать для решения задачи подсчёта однотипных объектов на изображениях.

### Литература

1. LeCun Y., Boser B. Backpropagation Applied to Handwritten Zip Code Recognition [Текст] / Y. LeCun, B. Boser // Neural Computation / – 1989. 541-551 с.
2. Lempitsky V., Zisserman A. Learning To Count Objects in Images [Текст] / V. Lempitsky, A. Zisserman // Advances in Neural Information Processing Systems 23 (NIPS 2010) / – 2010.
3. Simonyan, K. Very Deep Convolutional Networks for Large-Scale Image Recognition [Текст] / K. Simonyan, A. Zisserman // Computer Vision and Pattern Recognition / – 2014. – 14 с.
4. Евклидова Метрика [Электронный ресурс] // Википедия – свободная библиотека. - URL: [https://ru.wikipedia.org/wiki/Евклидова\\_метрика](https://ru.wikipedia.org/wiki/Евклидова_метрика) (дата обращения 10.12.2019).
5. Kingma D, Ba J. Adam: A Method for Stochastic Optimization [Текст] / D. Kingma, J. Ba // <https://arxiv.org/abs/1412.6980> / - 2015. – 15 с.