

АНАЛИЗ АРХИТЕКТУР ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

В.Н.Зуева¹⁾, И.А. Груднов²⁾

1) к.т.н., доцент кафедры внутризаводского электрооборудования и автоматики Армавирского механико-технологического института (филиал) ФГБОУ ВО «Кубанский государственный технологический университет», г. Армавир, Россия, Victoria_zueva@list.ru

2) студент Армавирского механико-технологического института (филиала) ФГБОУ ВО «Кубанский государственный технологический университет», г. Армавир, Россия, fidjerald63@mail.ru

Аннотация: Глубокие нейронные сети представляют перспективный инструмент машинного обучения и достигли значительных успехов во многих задачах компьютерного зрения и обработки изображений. В статье обсуждаются архитектуры новых структур глубоких нейросетей, которые имеют лучшую производительность, чем традиционная структура нейронной сети с прямой связью.

Ключевые слова: глубокие нейронные сети, нейронные сети с прямой структурой, оптимизация структуры, алгоритмы обучения.

ANALYSIS OF DEEP NEURAL NETWORK ARCHITECTURES

V. N. Zueva¹⁾, I.A. Grudnov²⁾

1) Ph. D., associate Professor, Armavir Institute of Mechanics and Technology (branch) of Federal State Budgetary Institution of Higher Education “Kuban State Technological University”, city of Armavir, Russia, Victoria_zueva@list.ru

1) the student Armavir Institute of Mechanics and Technology (branch) of Federal State Budgetary Institution of Higher Education “Kuban State Technological University”, city of Armavir, Russia, fidjerald63@mail.ru

Annotation: Deep neural networks represent a promising machine learning tool and have made significant advances in many computer vision and image processing tasks. The article discusses the architectures of new deep neural network structures that have better performance than the traditional feedforward neural network structure.

Keywords: deep neural networks, neural networks with direct structure, structure optimization, learning algorithms.

Глубокие нейронные сети (ГНС) пользуются большой популярностью при решении многих задач компьютерного зрения и обработки изображений таких, как классификация данных [1] и обнаружение необходимых объектов [2]. На сегодняшний день предложено множество вариантов ГНС. Например, проект ImageNet [3] был направлен на создание и сопровождение массивной

базы данных аннотированных изображений с иерархией WordNet, где каждый узел иерархии изображен сотнями и тысячами изображений. Для классификации изображений в ImageNet были предложены различные типы архитектур нейронных сетей, производительность которых постоянно улучшается. Так для распознавания визуальных объектов были разработаны нейронные сети VGG [4], остаточные сети (ResNets) [5], плотные сверточные сети (DenseNets) [6], в которых используются пропускающие соединения и плотные соединения для упрощения обучения и повышения производительности сети. Однако, основной проблемой ГНС является нахождение компромисса между производительностью и сложностью структуры нейронной сети. При выборе подходящей структуры сети необходимо учитывать параметры сети (глубина сети, тип каждого устройства и связь между уровнями), которые зависят от поставленной задачи и типа данных. Таким образом, от архитектуры ГНС зависит ее производительность [7]. Поэтому поиск новых эффективных методов оптимизации структуры сети является важной задачей для многих исследователей, поскольку не существует практических правил и часто оптимальная архитектура определяется либо опытным путем проб и ошибок, либо перекрестной проверкой, что требует большего времени на вычисления.

При решении множества задач компьютерного зрения и обработки изображений ключевым фактором для увеличения мощности вычислений ГНС является глубина сети. С увеличением данных за последние десятилетия глубина CNN постоянно увеличивалась, что привело к увеличению трудностей с оптимизацией структуры и обучением сети. Добавление большего количества слоев не гарантирует лучших результатов. Для решения этих проблем были разработаны ResNet и DenseNet, которые решают проблемы обучения ГНС. На рис. 1 представлены традиционные структуры перечисленных ГНС.

Традиционные CNN с прямой связью соединяют выход i -го слоя x_i со следующим $(i + 1)$ слоем после применения композиции операций $\Phi_i(\cdot)$, которая включает в себя операцию свертки или объединение слоев, пакетную нормализацию и функцию активации. Уравнение этого процесса имеет вид:

$$x_{i+1} = \Phi_i(\cdot)x_i. \quad (1)$$

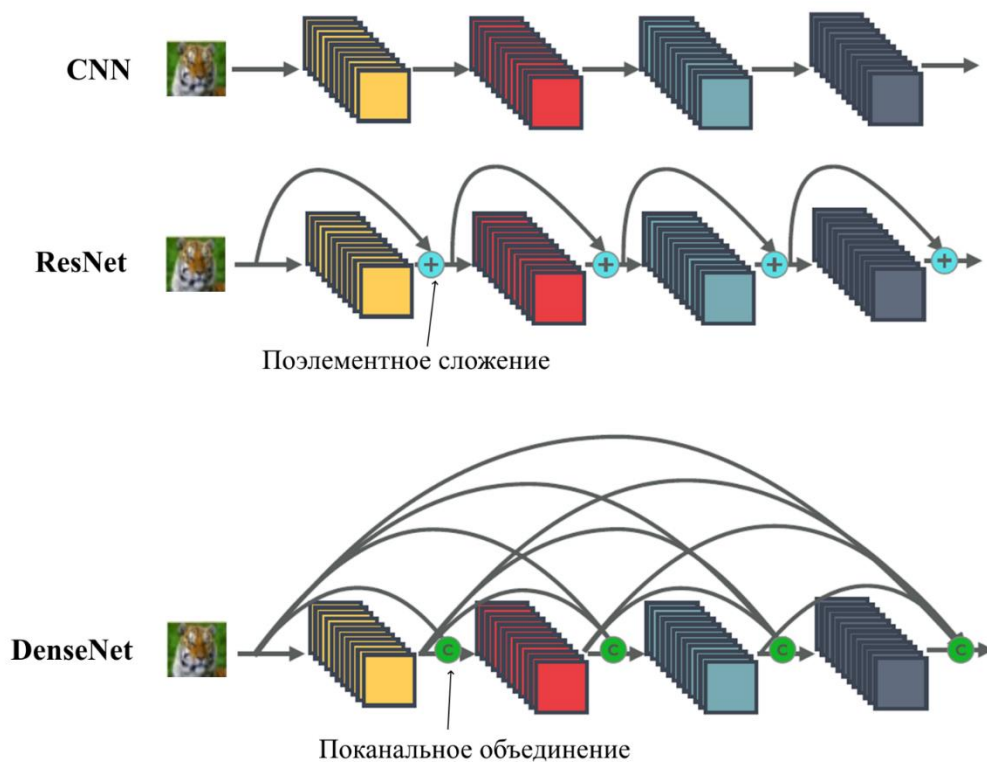


Рисунок 1 - Архитектуры основных ГНС, которые используются для классификации обработки изображений.

Как видно из рисунка, ResNet состоят из большого количества составных блоков. Поскольку ГНС извлекают различные по уровню признаки сквозным многослойным способом, то увеличение количества слоев в каждом блоке приводит к увеличению уровней признаков. Основная идея ResNet состоит в облегчении изучения сопоставлений идентичности для этих блоков. Это достигается за счет добавления к входным блокам его изученному выходу $F(x)+x$, что позволяет этим слоям в блоке соответствовать «остаточному». ResNets очень эффективны в изучении CNN и широко используются на практике для решения сложных задач, таких как обнаружение объектов и сегментирование экземпляров. Учитывая решающее значение ResNets для изучения глубоких представлений для визуального распознавания, в этой статье предлагается оптимизировать структуру ResNets для увеличения ее производительности. В ResNets уравнение (1) имеет следующий вид:

$$x_{i+1} = \Phi_i(\cdot)x_i + x_i. \quad (2)$$

Кроме того, согласно недавним исследованиям [8,9] увеличение глубины и эффективности CNN обучения можно также достичь за счет более коротких связей между слоями, близкими ко входу, и уровнями, близкими к выходу. DenseNet представляет собой сеть с плотными соединениями между слоями [7] (рис.1). DenseNet требует меньше параметров по сравнению с CNN, поскольку нет необходимости изучать избыточные карты признаков. В отличие от ResNet, у которой количество слоев велико из-за веса, которые нужно изучить для каждого слоя, в DenseNet количество сокращается и добавляются небольшой набор новых карт признаков. Также, DenseNet решает проблему с обучением из-за упомянутого потока информации и градиентов через прямой доступ к

градиентам из функции потерь и входного изображения. В отличие от ResNet, DenseNet не суммирует выходные карты признаков слоя с входящими картами признаков, а объединяет их, и уравнение (1) записывается в следующем виде:

$$x_{i+1} = \Phi_i([x_0, x_1, \dots, x_i]). \quad (3)$$

Независимо от того, являются ли группы с добавлением или объединением фильтров, то их количество постоянно меняется. Эти переходные слои зависят от понижающей дискретизации и применяют пакетную нормализацию, свертку 1x1 и слои объединения 2x2.

С объединением карты признаков размер канала увеличивается на каждом слое, следовательно, для создания k карт признаков для каждого раза будет обобщение для i -го слоя выглядеть как

$$k_i = k_0 + a(i - 1). \quad (4)$$

Где a - скорость роста, которая регулируется на каждом уровне. Каждый слой имеет доступ к предыдущим картам признаков и общей информации об объекте. Затем каждый слой добавляет новую информацию в общую информацию в виде конкретных k карт признаков.

В этой статье изучены особенности архитектур ГНС, модифицированные с помощью различных алгоритмов оптимизации и обучения для задач классификации и распознавания изображений. Результаты анализа экспериментальных данных показали, что оптимизация структур ResNet и DenseNet с помощью алгоритмов градиентного спуска и тяжелого шара, соответственно, приводит к увеличению точности результатов с увеличением глубины сети.

Список использованных источников:

1. Bashar, A. (2019). Survey on evolving deep learning neural network architectures. *Journal of Artificial Intelligence*, 1(02), 73-82.
2. Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234, 11-26.
3. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
4. Guan, Q., Wang, Y., Ping, B., Li, D., Du, J., Qin, Y., ...& Xiang, J. (2019). Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. *Journal of Cancer*, 10(20), 4876.
5. Zhang, K., Sun, M., Han, T. X., Yuan, X., Guo, L., & Liu, T. (2017). Residual networks of residual networks: Multilevel residual networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(6), 1303-1314.
6. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
7. Zhou, D. X. (2020). Universality of deep convolutional neural networks. *Applied and computational harmonic analysis*, 48(2), 787-794.

8. Шумков Е.А., Карлов Д.Н. Скоростной метод обучения многослойного персептрона // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2011. № 65. С. 153-161.

9. Карлов Д.Н. Алгоритм скоростного обучения многослойного персептрона // Электронный сетевой политематический журнал "Научные труды КубГТУ". 2018. № 3. С. 167-173.