

Родионова А.В.

Научный руководитель: к.т.н., доцент Вершинин В.В.

Владимирский государственный университет имени Александра Григорьевича и

Николая Григорьевича Столетовых

г. Владимир, ул. Горького 87

E-mail: nasty.rod@yandex.ru, vitaly.vershinin@gmail.com

Преимущества применения сверточных нейронных сетей в задачах классификации изображений

В современном мире большое количество задач решается программным способом. Если говорить о простых проектах, то с ними способна справиться обычная компьютерная программа, основанная на стандартных статистических методах. Но существуют и сложные задачи, требующие прогнозирования, анализа больших объемов данных, определения соответствий и т. д. Для решения таких нестандартных для компьютера задач были изобретены нейронные сети. Так, в работе [1] как раз показывается преимущество использования нейронной сети перед обычным шаблонным методом.

Нейронных сетей в настоящее время существует огромное множество, вот лишь некоторые известные:

- 1) Многослойный перцептрон
- 2) Сверточная нейронная сеть
- 3) Рекурсивная нейронная сеть
- 4) Рекуррентная нейронная сеть

Также стоит уточнить, что для различных задач могут использоваться различные виды нейронных сетей.

Например, рекуррентная нейронная сеть наиболее хорошо показала себя при решении задач обработки естественного языка (Natural Language Processing, NLP). В работе [2] создали рекуррентную нейросеть для классификации текста. Описанная модель превосходит по качеству традиционные методы для всех используемых наборов данных.

Мы решаем задачу классификации, поэтому следует выбрать ту нейронную сеть, которая наилучшим образом подойдет для решения именно этой задачи. Классификация изображений — это задача из области компьютерного зрения. В качестве входных данных принимается изображение, классифицируется и причисляется к предписанному классу.

На данный момент выдающиеся результаты в приложениях с распознаванием картинок показывают именно сверточные нейронные сети.

Сверточная нейронная сеть (Convolutional Neural Network, CNN) названа так, потому что содержит один или более соединенных сверточных слоев. CNN использует вариацию многослойного перцептрона. Сверточные слои применяют операцию свертки для входных данных и передают результат в следующий слой. Эта операция позволяет сети быть глубже при меньшем количестве параметров.

Например, в статье [3] проводится эксперимент распознавания черно-белых изображений с использованием обычных и сверточных нейронных сетей. Рассматривалось шесть простых нейронных сетей и две сверточных с различным числом параметров. В результате были сделаны выводы, что даже простая свёрточная сеть, у которой до 80 раз меньше коэффициентов, выдаёт качество на ~1% лучше, чем обычная. К тому же, сверточная сеть тренировалась до своей максимальной точности в три раза быстрее.

Также, в [4, 5] наилучшие результаты в области распознавания лиц (по результатам анализа публикаций) показала именно сверточная нейронная сеть. Успех обусловлен возможностью учета двумерной топологии изображения, в отличие от многослойного перцептрона.

Рассмотрим пример для лучшего понимания принципов работы CNN. При классификации картинок размером $m \times n = 6 \times 6$ необходимо выделять $k = 50$ признаков, тогда входной и выходной слои имеют $m \times n = 6 \times 6 = 36$ нейронов, а скрытый слой имеет $k = 50$ нейронов, количество весов равно $2 \times m \times n \times k = 2 \times 6 \times 6 \times 50 = 3600$. Но если картинки имеют размер $m \times n =$

100×100 , то количество весов будет $2 \times 100 \times 100 \times 50 = 1\,000\,000$. Почти в 278 раз больше, существенная цифра. В таком случае, чтобы размер нейронной сети не был большим и не зависел от размера обучаемых данных, и используется CNN [6]. При обучении для каждого образца на вход свёрточной нейронной сети попадает не вся картинка $m \times n$, а только ее часть $r \times c$, например, 2×2 , эта часть называется фильтром или ядром, и она сдвигается по большой картинке. Такая свёрточная нейронная сеть имеет фиксированное количество нейронов на входном и выходном слоях $r \times c = 2 \times 2 = 4$, а количество весов $2 \times r \times c \times k = 2 \times 2 \times 2 \times 50 = 400$.

Для каждой большой картинке $m \times n$ берем маленькие картинки $r \times c$ для обучения, их количество равно $(m - r + 1) \times (n - c + 1)$ (рис. 1) [7].

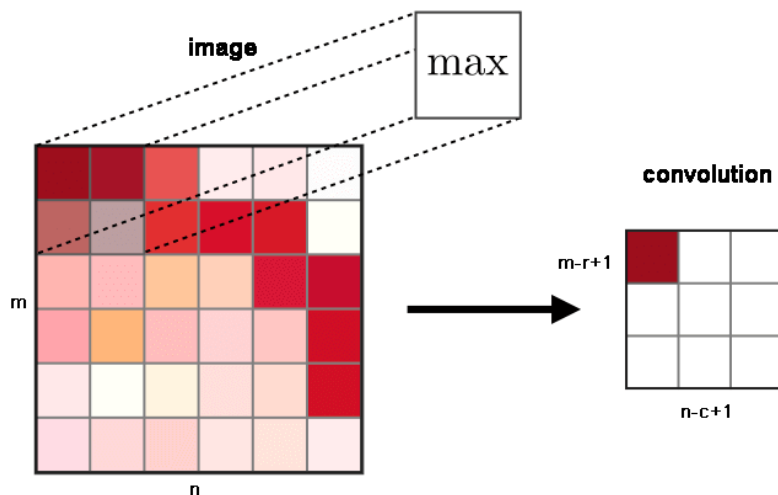


Рис. 1. Процесс свертки

Таким образом, мы можем достичь той же точности распознавания, что и с применением обычных нейронных сетей, но с меньшим количеством параметров.

В итоге выделены следующие преимущества CNN:

- 1) Повышение скорости обучения и уменьшение количества параметров в отличие от обычной нейронной сети
- 2) Возможность применения параллельных вычислений и создание алгоритмов обучения на графических процессорах (GPU).
- 3) Устойчивость к искажению входных данных, шуму, искажению, смене позиции объекта.

Например, собака на двух картинках имеет разные позиции, но свёрточная нейронная сеть выделяет одни и те же признаки (рис. 2) [7]. Это происходит потому, что при обучении CNN сдвигается по частям объекта, поэтому обучаемые признаки не зависят от позиции «важных частей». Подобное свойство CNN помогает повышать качество классификации [8].



Рис. 2. Устойчивость к сдвигу позиции объекта

Литература

1. Михалевич Ю. С., Ткаченко В. В. ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ АВТОМОБИЛЬНЫХ НОМЕРОВ. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ ПО СРАВНЕНИЮ С ШАБЛОННЫМ МЕТОДОМ // Научный журнал КубГАУ, - №120(06), - 2016 г.
2. Siwei Lai, Liheng Xu, Kang Liu, Jun Zhao Recurrent convolutional neural networks for text classification // AAAI'15: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, - January 2015, - pp. 2267–2273
3. Сборник статей [Электронный ресурс] // 21.10.2017, – Режим доступа: https://vbystricky.github.io/2017/10/mnist_cnn.html
4. Huriel Khalajzadeh, Mohammad Manthouri, Mohammad Teshnehlab. (2014) Face Recognition Using Convolutional Neural Network and Simple Logistic Classifier // Advances in Intelligent Systems and Computing 223:197-207
5. Steve Lawrence, Andrew Back. Face Recognition: A Convolutional Neural Network Approach // IEEE Xplore: https://www.researchgate.net/publication/3302251_Face_Recognition_A_Convolutional_Neural_Network_Approach
6. LeCun Y. LeNet-5, convolutional neural networks. Retrieved 16 November 2013.
7. Ле Мань Ха, Свёрточная нейронная сеть для решения задачи классификации // Сборник ТРУДЫ МФТИ, - 2016
8. Bengio Y., LeCun Y. Scaling learning algorithms towards AI. eds. L. Bottou, O. Chapelle, D. DeCoste, J. Weston. Large Scale Kernel Machines. MIT Press, 2007.