

Тятя! Тятя! Нейросети заменили продавца!

Ппилиф Ульяновкин

Листочек 6: свёрточные сети

Урра! Отлично сработано, ребятаки. Давайте завтра не придем? Возьмем отгул на денек? Вы пробовали шаурму? В двух кварталах отсюда делают какую-то шаурму. Не знаю, что это, но мне хочется.

Тони Старк (Мстители, 2012)

Упражнение 1 (Свёртка своими руками)

У Маши есть картинка и свёртка, которую она хочет применить к этой картинке¹.

3	3	2	1	0
0	0	1	3	1
3	1	2	2	3
3	1	2	2	3
3	1	2	2	3

Картинка Маши

0	1	2
2	2	0
0	1	2

Свёртка Маши

- Сделайте свёртку картинки без сдвигов и дополнений. К тому, что получилось применить `max pooling` и `average pooling` размера 2×2 .
- Примените к исходной картинке свёртку с параметром сдвига (`stride`) равным 2.
- Примените к исходной картинке свёртку с дополнением нулями (`zero padding`) и параметром сдвига (`stride`) равным 1.
- Сделайте расширенную свёртку (`dilated convolution`) с параметром расширения равным 2 (`dilation factor`). Сравните такую свёртку с обычной свёрткой со сдвигом. Эквивалентны ли они?

¹Многие задачи сделаны по мотивам виньетки: <https://arxiv.org/pdf/1603.07285.pdf>

Упражнение 2 (Ядра)

У Маши есть куча ядер для свёрток. Догадайтесь, что именно они делают с картинкой.

0	0	0
0	1	0
0	0	0

-1	-1	-1
-1	8	-1
-1	-1	-1

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

1	0	0
0	1	0
0	0	1

-0.1	0.1	-0.1
0.1	0.5	0.1
-0.1	0.1	-0.1

-1	0	1
-1	0	1
-1	0	1

-1	-1	-1
0	0	0
1	1	1

0	1	2
-1	0	1
-2	-1	0

0	1	0
1	1	1
0	1	0

1	0	100
0	1	50

Упражнение 3 (Крестики и нолики)

Маша хочет научить компьютер играть в крестики и нолики. На первом шаге ей надо научить алгоритм распознавать есть ли крестик на картинке. Под ноликом понимается любая фигура с дырой в середине. Под крестиком понимается любая фигура из двух пересекающихся линий.

Алгоритм должен быть устроен следующим образом. На первом шаге одна или несколько свёрток проходят по картинке. На втором шаге по результатам свёрток принимается решение. Например, берётся максимальное получившееся число и сравнивается с каким-то порогом. Классификатор крестиков и ноликов должен работать безупречно. Помогите Маше придумать такой классификатор.

- а. В мире Маши на картинках могут быть нарисованы либо крестики либо нолики. Все картинки, подающиеся на вход алгоритма могут быть только размера 4×4 . Примеры крестиков и ноликов нарисованы ниже.

1	1	1	0
1	0	1	0
1	0	1	0
1	1	1	0

нолик

0	0	1	0
0	1	0	1
0	0	1	0
0	0	0	0

нолик

0	1	0	1
0	0	1	0
0	1	0	1
0	0	0	0

крестик

0	1	0	0
1	1	1	1
0	1	0	0
0	1	0	0

крестик

- б. Предположим, что теперь кроме крестиков и ноликов в нашем мире существуют ещё и другие любые картинки. Как можно модернизировать ваш алгоритм, чтобы он по-прежнему стабильно работал с безупречным качеством?

Упражнение 4 (Свёрточный и полносвязный)

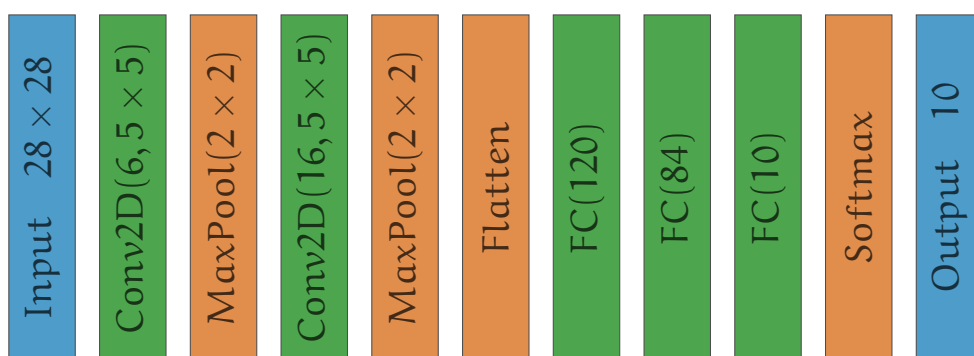
Маше рассказали, что свёрточный слой — это полносвязный слой с некоторыми ограничениями. Она хочет разобраться, что это за ограничения. На вход в слой идёт чёрно-белое изображение размера 4×4 . Каждый пиксель изображения — отдельная переменная.

- Нарисуйте с помощью кругляшей и стрелочек полносвязный слой, который обрабатывает картинку. Подпишите все веса. Нарисуйте свёрточный слой в таком же формате. На картинке часть связей исчезнет. Часть весов станут одинаковыми.
- Запишите свёрточный слой с помощью перемножения матриц в виде $H = X \cdot W$. Как выглядит матрица W ? Как через свёрточный слой можно сделать шаг обратного распространения ошибки?

Упражнение 5 (Число параметров)

На вход в нейронную сетку идёт изображение рукописной цифры размера 28×28 .

- Маша вытягивает изображение в длинный вектор и использует полносвязную сетку для классификации изображений. В сетке идёт один полносвязный слой из 1000 нейронов. После идёт слой, который осуществляет классификацию изображения на 10 классов. Сколько параметров нужно оценить?
- Маша вместо полносвязной сетки использует свёрточную. На первом шаге делается 6 свёрток размера 5×5 . На втором шаге делается max-pooling размера 2×2 . На третьем 16 свёрток размера 5×5 . На четвертом max-pooling размера 2×2 . На пятом картинка вытягивается в длинный вектор. Далее идут три полносвязных слоя размеров 120, 84, 10. В конце делается softmax. После каждой свёртки и полносвязного слоя, кроме последнего, в качестве функции активации используется ReLU.



Сколько параметров необходимо будет оценить в такой модели? Какого размера будут выходы из каждого слоя?

- Маша использует архитектуру из пункта б, но все свёртки делает с дополнением нулями (zero padding). Как изменится число оцениваемых параметров? Какого размера будут выходы из каждого слоя?
- Маша использует архитектуру из пункта б, но все свёртки делает с параметром сдвига (stride) равным 2. Как изменится число оцениваемых параметров? Какого размера будут выходы из каждого слоя?

Упражнение 6 (Поле обзора)

Маша хочет найти котика размера 512×512 пикселей. Для этого она использует свёртки размера 5×5 без дополнения нулями (zero padding) После каждого свёрточного слоя Маша делает max-pooling размера 2×2 .

- а. Через сколько слоёв поле восприятия (receptive field) Машиной нейросетки впервые охватит котика?
- б. Маша хочет поменять max-pooling на свёртки со сдвигом (stride) так, чтобы котейка находился за такое же число слоёв. Какой размер сдвига ей надо выбрать?
- в. Пусть s — величина сдвига, k — размер свёртки, m — размер пулинга, n — номер слоя. Выпишите формулу, по которой можно найти размер поля видимости (receptive field).
- г. Маша делает расширенную свёртку (dilated convolution) с параметром расширения равным r . Пусть k — размер свёртки, m — размер пулинга, n — номер слоя. Выпишите формулу, по которой можно найти размер поля видимости (receptive field).

Упражнение 7 (Магия чисел)

- а. Маша использует свёртку размера $k \times k$. Какого размера должно быть дополнение (padding), чтобы размер картинки после свёртки не поменялся?
- б. Маша использует ту же свёртку со сдвигом (stride) размера s . Какого размера должно быть дополнение (padding), чтобы размер картинки после свёртки не поменялся?
- в. У Маши есть свёрточный слой. На вход в свёрточный слой идёт изображение с C_{in} каналами размера $W \times H$. Маша использует C_{out} фильтров размера $W_k \cdot H_k$. Сколько параметров ей предстоит оценить?
- г. Свёртка размера $k \times k$ называется сепарабельной, если её ядро можно представить в виде произведения двух векторов размеров $k \times 1$ и $1 \times k$. Пусть все свёртки, которые использует Маша — сепарабельные. Сколько параметров в условиях предыдущего пункта ей предстоит оценить?

Упражнение 8 (Скользящее среднее)

Временные ряды иногда сглаживают с помощью скользящего среднего. Этот процесс можно рассматривать как одномерную свёртку. Как выглядит ядро такой свёртки? Какой физический смысл стоит за размером такой свёртки и дополнением нулями? Опишите как она работает.