Нейронные сети

Тема семинара: Свертки

Иванов И.Е.

кафедра Математической Теории Интеллектуальных Систем

11 октября 2024 г.





План семинара

- Скользящее окно VS перемножение матриц
- 2 Транспонированная свертка



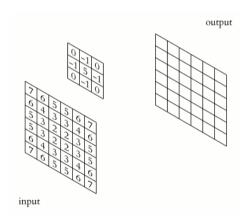


• Вообще, операция применения свертки — это последовательное вычисление скалярного произведения с помощью скользящего окна





• Вообще, операция применения свертки — это последовательное вычисление скалярного произведения с помощью скользящего окна



• А можно ли данную операцию делать более единообразно (например, с помощью обычных матричных операций навроде умножения матриц)?



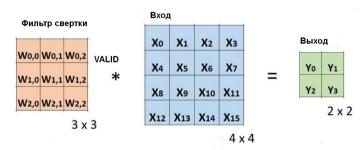


- А можно ли данную операцию делать более единообразно (например, с помощью обычных матричных операций навроде умножения матриц)?
- Для этого рассмотрим простой случай:
 - ullet Вход 4 imes 4, фильтр свертки 3 imes 3, тип свертки "VALID" (без добивки) \Rightarrow выход 2 imes 2





- А можно ли данную операцию делать более единообразно (например, с помощью обычных матричных операций навроде умножения матриц)?
- Для этого рассмотрим простой случай:
 - ullet Вход 4 imes 4, фильтр свертки 3 imes 3, тип свертки "VALID" (без добивки) \Rightarrow выход 2 imes 2







• Оказывается, последовательное вычисление скалярного произведения с помощью скользящего окна можно реализовать с помощью обычного матричного умножения!





- Оказывается, последовательное вычисление скалярного произведения с помощью скользящего окна можно реализовать с помощью обычного матричного умножения!
- Для этого:
 - ullet Вход представим как столбец X размера 16 imes 1





- Оказывается, последовательное вычисление скалярного произведения с помощью скользящего окна можно реализовать с помощью обычного матричного умножения!
- Для этого:
 - ullet Вход представим как столбец X размера 16 imes 1
 - ullet Выход представим как столбец Y размера 4 imes 1





- Оказывается, последовательное вычисление скалярного произведения с помощью скользящего окна можно реализовать с помощью обычного матричного умножения!
- Для этого:
 - ullet Вход представим как столбец X размера 16 imes 1
 - ullet Выход представим как столбец Y размера 4 imes 1
 - А фильтр хитрым образом (см. картинку на следующем слайде) разложим в разреженную матрицу C размера 4×16





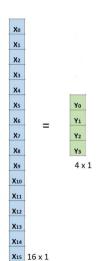
 \bullet Тогда $Y = C \cdot X$:





• Тогда $Y = C \cdot X$:







- Можно поступать и наоборот:
 - ullet Фильтр представить как строку F размера 1 imes N





- Можно поступать и наоборот:
 - ullet Фильтр представить как строку F размера 1 imes N
 - Вход хитрым образом (см. картинку на следующем слайде) разложить в матрицу X размера $N \times K$, где $K = H \cdot W$ размерность выхода





- Можно поступать и наоборот:
 - ullet Фильтр представить как строку F размера 1 imes N
 - \bullet Вход хитрым образом (см. картинку на следующем слайде) разложить в матрицу Xразмера $N \times K$, где $K = H \cdot W$ — размерность выхода
 - Выход строка $Y = F \cdot X$





- Можно поступать и наоборот:
 - ullet Фильтр представить как строку F размера 1 imes N
 - Вход хитрым образом (см. картинку на следующем слайде) разложить в матрицу X размера $N \times K$, где $K = H \cdot W$ размерность выхода
 - ullet Выход строка $Y = F \cdot X$
- Подобная операция называется "im2col", и применяется в низкоуровневых реализациях

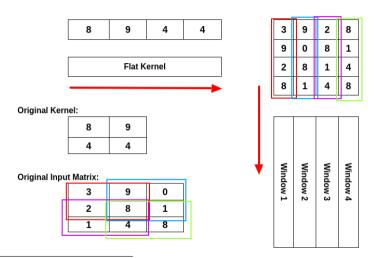




- Можно поступать и наоборот:
 - ullet Фильтр представить как строку F размера 1 imes N
 - Вход хитрым образом (см. картинку на следующем слайде) разложить в матрицу X размера $N \times K$, где $K = H \cdot W$ размерность выхода
 - ullet Выход строка $Y = F \cdot X$
- Подобная операция называется "im2col", и применяется в низкоуровневых реализациях
- Правда, в этом случае не удастся продемонстрировать, почему "транспонированную" свертку так называют



Иллюстрация $im2col^1$



¹https://towardsdatascience.com/

how-are-convolutions-actually-performed-under-the-hood-226523ce7fbf - - -



- \bullet Предположим, что мы хотим из входа X размера 4×1 получить выход Y большего размера 16×1
- ullet Применение фильтра F умножение слева на вход





- \bullet Предположим, что мы хотим из входа X размера 4 \times 1 получить выход Y большего размера 16 \times 1
- Применение фильтра F умножение слева на вход
- ullet Таким образом, фильтр должен быть матрицей 16 imes 4





- \bullet Предположим, что мы хотим из входа X размера 4×1 получить выход Y большего размера 16×1
- ullet Применение фильтра F умножение слева на вход
- ullet Таким образом, фильтр должен быть матрицей 16 imes 4
- Получается, что размер фильтра равен размеру транспонированной матрице фильтра C, которая была нужна, чтобы из 16×1 получить 4×1 (вход меняем с выходом)



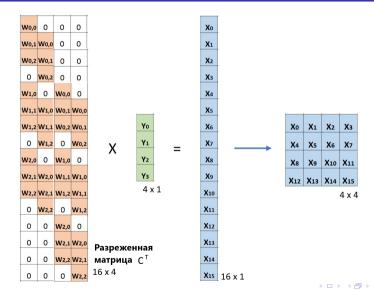


- \bullet Предположим, что мы хотим из входа X размера 4×1 получить выход Y большего размера 16×1
- Применение фильтра F умножение слева на вход
- ullet Таким образом, фильтр должен быть матрицей 16 imes 4
- Получается, что размер фильтра равен размеру транспонированной матрице фильтра C, которая была нужна, чтобы из 16×1 получить 4×1 (вход меняем с выходом)
- Вот и источник названия "траспонированная" свертка!



Иванов И Е 11 октября 2024 г. 9 / 20 Свертки

Иллюстрация транспонированной свертки как перемножения матриц





11 октября 2024 г.

Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется





Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

• Как и для обычной свертки, для транспонированной используются два главных параметра:





Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг s (stride)





Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг *s* (stride)
 - Добивка p (паддинг / padding)





Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг *s* (stride)
 - \bullet Добивка p (паддинг / padding)
- Однако в данном случае (как и в случае с перемножением матриц), здесь все наоборот



 Иванов И.Е.
 Свертки
 11 октября 2024 г.
 11/20

Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг *s* (stride)
 - \bullet Добивка p (паддинг / padding)
- Однако в данном случае (как и в случае с перемножением матриц), здесь все наоборот
- ullet При шаге s>1 для входной матрицы между каждыми элементами вставляется ровно (s-1) нулей



Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг *s* (stride)
 - \bullet Добивка p (паддинг / padding)
- Однако в данном случае (как и в случае с перемножением матриц), здесь все наоборот
- ullet При шаге s>1 для входной матрицы между каждыми элементами вставляется ровно (s-1) нулей
 - При этом фильтр $q \times q$ скользящим окном по преобразованному входу идет всегда с шагом 1



Иванов И.Е. Свертки 11 октября 2024 г. 11/20

Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг s (stride)
 - Добивка р (паддинг / padding)
- Однако в данном случае (как и в случае с перемножением матриц), здесь все наоборот
- ullet При шаге s>1 для входной матрицы между каждыми элементами вставляется ровно (s-1) нулей
 - При этом фильтр $q \times q$ скользящим окном по преобразованному входу идет всегда с шагом 1
- Далее выполняется добивка нулями ширины (q-1-p)



11 / 20

Иванов И.Е Свертки

Параметры транспонированной свёртки

Параметры следует рассматривать как параметры обычной свёртки для построения матрицы с той лишь разницей, что после построения матрица транспонируется

- Как и для обычной свертки, для транспонированной используются два главных параметра:
 - Шаг *s* (stride)
 - \bullet Добивка p (паддинг / padding)
- Однако в данном случае (как и в случае с перемножением матриц), здесь все наоборот
- ullet При шаге s>1 для входной матрицы между каждыми элементами вставляется ровно (s-1) нулей
 - \bullet При этом фильтр $q\times q$ скользящим окном по преобразованному входу идет всегда с шагом 1
- ullet Далее выполняется добивка нулями ширины (q-1-p)
 - Заметим: p не прибавляется, а отнимается!



Параметры транспонированной свертки: Stride

s=1

При s=1 транспонированная свёртка является свёрткой





Параметры транспонированной свертки: Stride

s=1

При s=1 транспонированная свёртка является свёрткой

Переход от s=1 к s>1 для обыкновенной свёртки

Если мы построили матрицу для обыкновенной свёртки при s=1, то матрица для s>1 получается из исходной путём удаления строк по некоторой регулярной решетке.





Параметры транспонированной свертки: Stride

s=1

При s=1 транспонированная свёртка является свёрткой

Переход от s=1 к s>1 для обыкновенной свёртки

Если мы построили матрицу для обыкновенной свёртки при s=1, то матрица для s>1 получается из исходной путём удаления строк по некоторой регулярной решетке.

Переход от s=1 к s>1 для транспонированной свёртки

Если мы построили матрицу для обыкновенной свёртки при s=1, то матрица для s>1 получается из исходной путём удаления столбцов по некоторой регулярной решетке.

Иванов И.Е. Свертки 11 октября 2024 г. 12 / 20

Параметры транспонированной свертки: Stride

s=1

При s=1 транспонированная свёртка является свёрткой

Переход от s=1 к s > 1 для обыкновенной свёртки

Если мы построили матрицу для обыкновенной свёртки при s=1, то матрица для s>1 получается из исходной путём удаления строк по некоторой регулярной решетке.

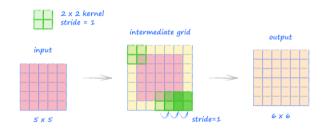
Переход от s=1 к s>1 для транспонированной свёртки

Если мы построили матрицу для обыкновенной свёртки при s=1, то матрица для s>1 получается из исходной путём удаления столбцов по некоторой регулярной решетке.

То есть для перехода от s>1 к s=1 необходимо разбавить вход нулями по той же регулярной решетке, а потом применить обычную свёртку.

Иванов И.Е. Свертки 11 октября 2024 г. 12 / 20

Транспонированная свертка: $s = 1, p = 0^2$



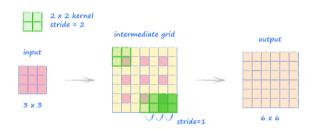


13 / 20

Иванов И.Е. Свертки 11 октября 2024 г.

²https://makeyourownneuralnetwork.blogspot.com/2020/02/calculating-output-size-of-convolutions.html

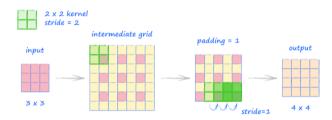
Транспонированная свертка: s = 2, p = 0







Транспонированная свертка: s = 2, p = 1







ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p





- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- ullet Тогда можно легко посчитать размер выхода w imes w:





- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- Тогда можно легко посчитать размер выхода $w \times w$:
 - ullet Размер входа со вставленными нулями: $(h-1)\cdot s+1$





- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- Тогда можно легко посчитать размер выхода $w \times w$:
 - ullet Размер входа со вставленными нулями: $(h-1)\cdot s+1$
 - ullet Размер суммарной добивки: $2\cdot (q-1-p)$





- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- Тогда можно легко посчитать размер выхода $w \times w$:
 - ullet Размер входа со вставленными нулями: $(h-1) \cdot s + 1$
 - Размер суммарной добивки: $2 \cdot (q 1 p)$
 - ullet Итого преобразованный вход размерности $w' imes w', w' = (h-1) \cdot s + 1 + 2 \cdot (q-1-p)$





- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- Тогда можно легко посчитать размер выхода $w \times w$:
 - Размер входа со вставленными нулями: $(h-1) \cdot s + 1$
 - ullet Размер суммарной добивки: $2\cdot (q-1-p)$
 - ullet Итого преобразованный вход размерности $w' imes w', w' = (h-1) \cdot s + 1 + 2 \cdot (q-1-p)$
 - ullet Размер выхода w imes w для обычной свертки размера q imes q с шагом 1 и входа размера w' imes w' равен: w=w'-(q-1)



16 / 20



- ullet Пусть вход размера h imes h, фильтр размера q imes q, шаг s и добивка p
- Тогда можно легко посчитать размер выхода $w \times w$:
 - ullet Размер входа со вставленными нулями: $(h-1)\cdot s+1$
 - ullet Размер суммарной добивки: $2\cdot (q-1-p)$
 - ullet Итого преобразованный вход размерности $w' imes w', w' = (h-1) \cdot s + 1 + 2 \cdot (q-1-p)$
 - ullet Размер выхода w imes w для обычной свертки размера q imes q с шагом 1 и входа размера w' imes w' равен: w = w' (q-1)
 - Объединяя, получим:

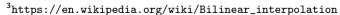
$$w = (h-1) \cdot s - 2 \cdot p + q$$



Иванов И.Е. Свертки 11 октября 2024 г. 16/20

Билинейная интерполяция

• Если с помощью транспонированной свертки мы можем увеличивать размер выхода, то можно ли реализовать классический алгоритм билинейной интерполяции³?





17 / 20

 Иванов И.Е.
 Свертки
 11 октября 2024 г.

Билинейная интерполяция

- Если с помощью транспонированной свертки мы можем увеличивать размер выхода, то можно ли реализовать классический алгоритм билинейной интерполяции³?
- Формула линейной интерполяции для единичного квадрата (считаем, что нам известны значения функции в углах этого квадрата с координатами (0,0),(1,0),(0,1) и (1,1)):

$$f(x,y)pprox \left[egin{array}{ccc} f(x,y)pprox \left[egin{array}{ccc} 1-x & x
ight] egin{bmatrix} f(0,0) & f(0,1) \ f(1,0) & f(1,1) \ \end{bmatrix} egin{bmatrix} 1-y \ y \ \end{bmatrix}$$



Иванов И.Е. Свертки 11 октября 2024 г. 17/20

 Рассмотрим простейший случай: увеличение размерности ровно в два раза с помощью билинейной интерполяции



- Рассмотрим простейший случай: увеличение размерности ровно в два раза с помощью билинейной интерполяции
- Тогда нужно вставлять ровно одно значение между изначальными узлами решетки \Rightarrow шаг s=2





- Рассмотрим простейший случай: увеличение размерности ровно в два раза с помощью билинейной интерполяции
- Тогда нужно вставлять ровно одно значение между изначальными узлами решетки \Rightarrow шаг s=2
- В этом случае нам нужно вставить точки с координатами (0.5,0),(0,0.5),(0.5,0.5),(1,0.5) и (0.5,1)





- Рассмотрим простейший случай: увеличение размерности ровно в два раза с помощью билинейной интерполяции
- Тогда нужно вставлять ровно одно значение между изначальными узлами решетки \Rightarrow шаг s=2
- В этом случае нам нужно вставить точки с координатами (0.5,0),(0,0.5),(0.5,0.5),(1,0.5) и (0.5,1)
- Значит, при вычислении центральной точки (0.5,0.5) по формуле выше нам нужно использовать все четыре узла вокруг $(\mathsf{B}\ (0,0),(1,0),(0,1)\ \mathsf{u}\ (1,1))$





- Рассмотрим простейший случай: увеличение размерности ровно в два раза с помощью билинейной интерполяции
- Тогда нужно вставлять ровно одно значение между изначальными узлами решетки \Rightarrow шаг s=2
- В этом случае нам нужно вставить точки с координатами (0.5,0),(0,0.5),(0.5,0.5),(1,0.5) и (0.5,1)
- Значит, при вычислении центральной точки (0.5,0.5) по формуле выше нам нужно использовать все четыре узла вокруг $(\mathsf{B}\ (0,0),(1,0),(0,1)\ \mathsf{u}\ (1,1))$
- Это можно сделать с помощью фильтра размера 3 × 3:

$$\begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$$

Задание. Доказать (и обратить внимание на добивку!).



Время для вопросов







Спасибо за внимание!





11 октября 2024 г.