

Заметки по сверточным нейронным сетям

Алексей Козловский

24 марта 2021 г.

1 Основные математические понятия

1.1 Свертка сигналов

Определение 1.1 *Свертка (в математике) - это мера перекрывания одной функции с отраженной и сдвинутой копией другой. В дискретном случае имеем*

$$y(t) = x(t)(*)h(t) = \sum_{n=-\infty}^{\infty} x(\tau)h(t - \tau).$$

В непрерывном случае свертка двух функций записывается в виде:

$$y(t) = x(t)(*)h(t) = \int_{\tau=-\infty}^{\infty} x(\tau)h(t - \tau).$$

В общем случае любую величину, испытывающую изменения во времени и/или пространстве можно рассматривать как сигнал. Сигналы могут быть аналоговыми и цифровыми. Мы будем работать с цифровыми сигналами, так как используем вычислительную технику, которая работает дискретно. Поэтому, нам придется переводить аналоговый сигнал в цифровой.

Определение 1.2 *Семплирование (sampling) - замерение амплитуды сигнала через определенные пространственные или временные интервалы фиксированной величины.*

Определение 1.3 *Квантование - разбиение диапазона отсчетных значений сигнала на конечное число уровней и округление этих значений до ближайших уровней*

В итоге семплирования мы получаем счётное множество значений сигнала, но значения его элементов могут принадлежать несчётному множеству, например \mathbb{R} . Поэтому следует применить квантование, чтобы эти элементы стали тоже принадлежать счетному множеству. Ну и так как наши множества изначально ограничены сверху и снизу и мы сделали их счетными, то теперь мы можем хранить информацию о сигнале.

После применения семплирования к аналоговому сигналу получается **дискретный** сигнал.

После применения квантования к дискретному сигналу получается **цифровой** сигнал.

Определение 1.4 Двумерная единичная дельта функция $\delta(n_1, n_2)$, где n_1 и n_2 - координаты горизонтали и вертикали определяется следующим образом: $\delta(0, 0) = 1$, во всех остальных случаях $\delta(n_1, n_2) = 0$.

Определение 1.5 Смещенная единичная дельта функция $\delta(n_1 - k_1, n_2 - k_2)$ определяется следующим образом: $\delta(n_1 - k_1, n_2 - k_2) = 0$ только если $n_1 = k_1$ и $n_2 = k_2$.

Утверждение 1.1 Любой двумерный дискретный сигнал можно представить в виде взвешенной суммы единичной дельта-функции с различными координатами.

1.2 Часто используемые фильтры изображений

Определение 1.6 Усредняющий фильтр (*mean filter*) - низкочастотный, фильтр, вычисляющий скользящее среднее интенсивности пикселей в любой заданной точке.

$$\begin{pmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{pmatrix}$$

Усредняющий фильтр используется главным образом для снижения шума на изображении.

Определение 1.7 Медианный фильтр (*median filter*) - заменяет каждый пиксель медианным значением интенсивности пикселей в его окрестности, покрываемой ядром фильтра.

Медианный фильтр хорошо подходит для удаления черно-белого крапчатого шума.

Определение 1.8 Гауссовский фильтр - это видоизменённая версия медианного фильтра, в которой веса функции импульсного отклика распределены по нормальному закону относительно начала координат.

Определение 1.9 Горизонтальный градиентный фильтр - фильтр, применяя который можно посчитать градиент по x .

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Определение 1.10 Вертикальный градиентный фильтр - фильтр, применяя который можно посчитать градиент по y .

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Определение 1.11 Фильтр Собеля - градиентный фильтр (может быть как горизонтальным, так и вертикальным), который учитывает градиент не только в текущий точке, но и градиент в соседних точках.

$$H_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$$

$$H_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Фильтры Собеля высокочастотны, так как ослабляют низкочастотные компоненты. Их используют для обнаружения краев на изображении.

Определение 1.12 Фильтр тождественного преобразования - не меняет изображение.

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

2 Сверточные нейронные сети

Определение 2.1 Сверточные нейронные сети - нейронные сети, базирующиеся на свертке изображений и обнаружении особенностей на основе фильтров, которым обучается CNN.

2.1 Компоненты CNN

Компоненты CNN:

- **Входной слой.** Используется для хранения интенсивности пикселей изображения
- **Слой свертки.** Принимает изображения от предыдущих слоев и выполняет их свертку с определенным количеством фильтров для создания изображений, называемых *выходными картами признаков*. Количество выходных карт признаков = количеству фильтров.
- **Слой активации.** Например ReLU. Добавляет в сеть нелинейность.
- **Слой субдискретизации (пулинга).** Выполняет снижение размерности карт признаков по высоте и ширине (глубина не меняется!).
- **Полносвязные слои.** Обычные полносвязные слои (многослойный перцептрон).