

Основы функционирования свёрточной нейронной сети

Свёрточная нейронная сеть (англ. *convolutional neural network*, *CNN*) - специальная архитектура искусственных нейронных сетей, впервые предложенная Яном Лекуном в 1988 году для решения задачи распознавание образов.

Идея свёрточной нейронной сети базируется на некоторых особенностях зрительной коры головного мозга человека, в которой есть простые и сложные клетки. Простые клетки реагируют на прямые линии под разными углами (сигналы, идущие из глаза человека), а сложные клетки затем активируют определенные наборы простых клеток. Поэтому и структура свёрточной нейронной сети (рис. 1) представляет собой многослойный ориентированный граф, в котором между входным и выходным слоями нейронов чередуются свёрточные слои и слои выборки. Функционирование свёрточной нейронной сети интерпретируется как переход от конкретных особенностей изображения к более абстрактным деталям, и далее к еще более абстрактным деталям вплоть до выделения понятий высокого уровня. При этом сеть самонастраивается и вырабатывает сама необходимую иерархию абстрактных признаков (последовательности карт признаков).

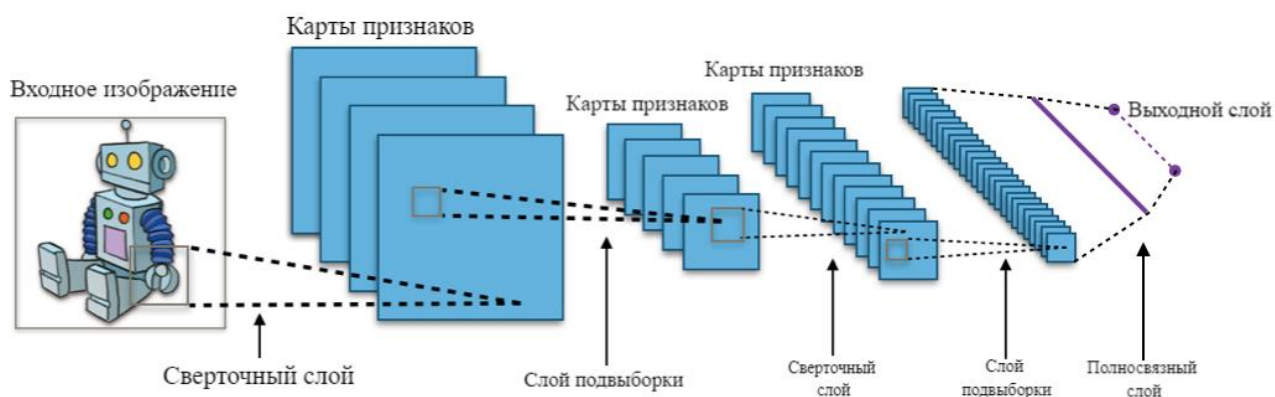


Рисунок 1 – Типовая архитектура CNN

Входной слой сети представляет собой матрицу признаков исходного изображения, в которой кодируются сигналы, приходящие из глаза человека. Они определяют входные и выходные сигналы нейронов во входном слое сети. В отличие от перцептрона, где у каждой дуги имеется свой вес, вычисление входного сигнала нейрона стока дуги определяется с помощью операции свёртки, для которой разработчик задает ядро свёртки (матрицу небольшой размерности, моделирующую некоторый вполне определенный признак

изображения). Как правило, разработчик задает несколько ядер свёртки Y^1, \dots, Y^k , где k – количество ядер свёртки. Отсюда и название таких нейронных сетей.

Построение карты признаков в свёрточном слое

Пусть имеется некоторая матрица M размерности $n_1 \times n_1$ из предыдущего слоя и ядро свёртки – матрица Y размерности $n_2 \times n_2$.

Свёртка – это операция над парой матриц M и Y , в результате которой мы получим новую матрицу M' для свёрточного слоя с размерностью:

$$n_1' = n_1 - n_2 + 1, \quad (1)$$

в которой каждый элемент вычисляется по следующей формуле:

$$m'(i, j) = f\left(\sum_{t=1}^{n_2} \sum_{p=1}^{n_2} m(i+t-1)(j+p-1) \cdot y(t, p)\right), \quad (2)$$

где f – функция активации нейронов из таблицы 1 (здесь x – входной сигнал нейрона).

Таблица 1 – Используемые функции активации

№	Функция активации	Формула
1	Линейная	$f(x) = x$
2	RELU (англ. <i>Rectified Linear Unit</i>) - линейный выпрямитель	$f(x) = x, \text{ если } x > 0$ $f(x) = 0, \text{ если } x \leq 0$

На рисунке 2 приведен пример вычисления выходного сигнала нейрона с линейной функцией активации.

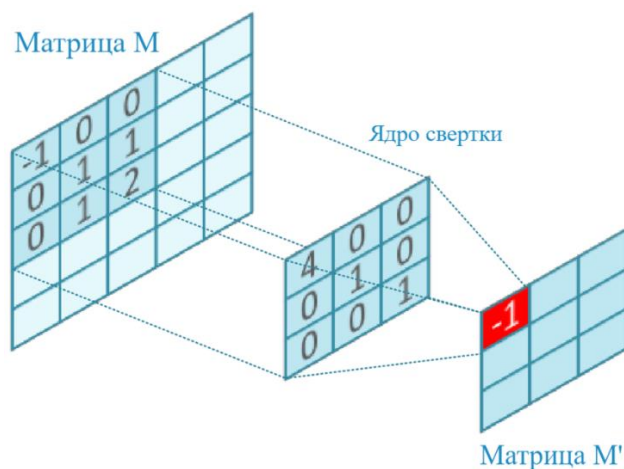


Рисунок 2 – Пример создания и заполнения матрицы M'

Здесь для матрицы M с размерностью 5×5 и ядра свертки – матрицы Y размерности 3×3 создана матрица M' для свёрточного слоя с размерностью 3×3 согласно формуле (1), а элемент $m'(1,1) = -1$ рассчитан по формуле (2), т.е. равен значению линейной функции активации (табл. 1) от суммы произведений соответствующих элементов выделенной в M подматрицы размерности 3×3 и матрицы Y .

Построение карты признаков в слое выборки

В процессе распространения сигнала через слой подвыборки происходит уменьшение размерности матриц предыдущего слоя свёртки за счет применения функции подвыборки $\mu(x)$ (табл. 2) над подматрицами A размерности $n_3 \times n_3$. Значение n_3 задается разработчиком, как правило, $n_3 = 2$, т.е. происходит уменьшение размерности матриц вдвое.

Таблица 2 – Используемые функции подвыборки

№	Функция подвыборки	Формула
1	Максимальное значение в подматрице A	$\mu(x) = \max(A)$
2	Среднее значение в подматрице A	$\mu(x) = \text{average}(A)$

Пусть имеется некоторая матрица M' размерности $n_1' \times n_1'$ из предыдущего свёрточного слоя. Тогда в следующем слое подвыборки для нее будет построена новая матрица M'' с размерностью:

$$n_1'' = n_1' / n_3, \quad (3)$$

в которой каждый элемент $m''(i, j)$ вычисляется как значение функции подвыборки (табл. 2) в подматрице A , выделенной в матрице M' , начиная с элемента:

$$m'(1 + (i - 1) \cdot n_3, 1 + (j - 1) \cdot n_3). \quad (4)$$

На рисунке 3 приведен пример построения матрицы M'' для $n_3 = 2$. Здесь подматрицы A и соответствующие им значения в матрице M'' показаны для наглядности разными цветами. Как видно из рисунка здесь в очередном слое подвыборки матрица была уменьшена в размерности в два раза: матрица M' имела размерность 4×4 , а построенная матрица M'' – 2×2 .

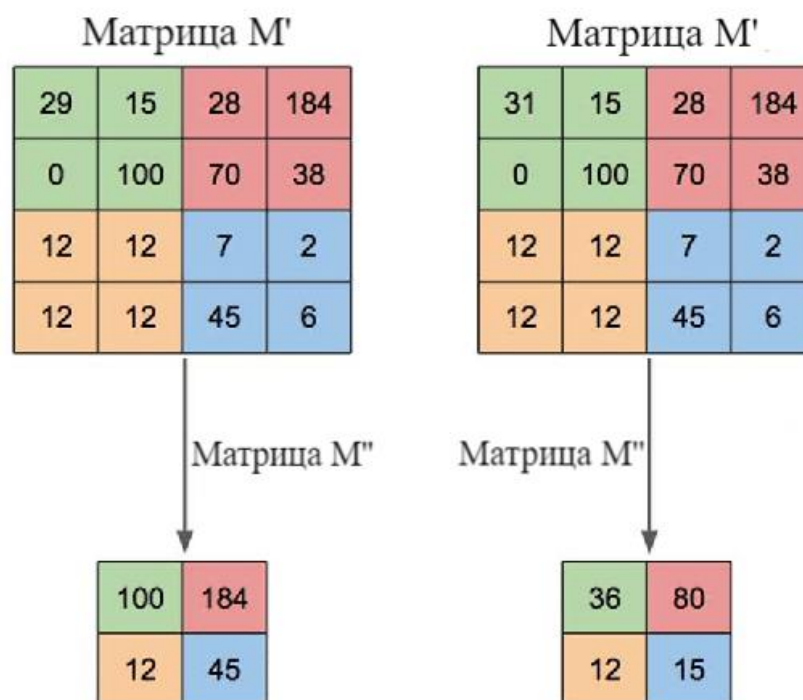


Рисунок 3 – Пример построения матрицы M'' в слое подвыборки с функцией подвыборки по максимальному значению (слева) и по среднему значению (справа)

Построение выходного слоя сети и оценка полученного решения

Процесс построения чередующихся слоев свёртки и слоев выборки завершается тогда, когда параметр n_1'' примет значение 1, т.е. когда в последнем слое выборки будут получены матрицы размерностью 1x1. Каждая такая матрица определяет входной сигнал соответствующего нейрона в выходном слое сети. Для определения выходного сигнала таких нейронов следует использовать заданную функцию активации (табл. 1). Для оценивания полученного решения используется **MSE** (mean squared error), которая рассчитывается аналогично тому, как и в перцептроне.

Пусть X_{OUT} - слой выходных нейронов, тогда:

$$MSE = \frac{1}{n} \sum_{\forall x_i \in X_{OUT}} (output(x_i) - output^*(x_i))^2, \quad (5)$$

где n – количество нейронов в выходном слое сети,

$output(x)$ – фактическое значение выходного сигнала в нейроне,

$output^*(x)$ – ожидаемое значение выходного сигнала в нейроне, которое определяется по следующему правилу:

$output^*(x) = 1$, если $output(x) > 0.5$;

$output^*(x) = 0$, если $output(x) \leq 0.5$

Алгоритм последовательного распространения сигнала в свёрточной нейронной сети

Пусть имеется матрица M размерности $n_1 \times n_1$ с признаками исходного изображения и набор из k -ядер свёртки – матрицы J_1, \dots, J_k , каждая с размерностью $n_2 \times n_2$. Определить выходные сигналы сети с использованием функций активации $f(x)$ и выборки $\mu(x)$ для подматриц размерности $n_3 \times n_3$, а также оценить полученное решение с помощью MSE. Введем счетчик свёрточных слоев – s .

1. Положить: $n_1'' = n_1, s=1$.
2. **ПОКА** $n_1'' > 1$ **ВЫПОЛНЯТЬ**:
 - 2.1. Рассчитать размерность матриц n_1' по формуле (1) для s -го свёрточного слоя.
 - 2.2. Построить для каждой исходной матрицы M набор матриц $M'_{s1}, M'_{s2}, \dots, M'_{sk}$ по формуле (2).
 - 2.3. Рассчитать размерность матриц слоя повыборки n_1'' по формуле (3).
 - 2.4. Построить для каждой матрицы M' s -го слоя свёртки матрицу M'' согласно заданной функции выборки $\mu(x)$ над подматрицей A , выделенной в M' по формуле (4).
 - 2.5. Определим: $s = s+1, n_1 = n_1''$.
 - 2.6. Положить в качестве исходных матриц матрицы $M''_{s1}, M''_{s2}, \dots, M''_{sk}$ и пронумеровать их по порядку как M_1, M_2, \dots .
3. Определить количество нейронов и их выходные сигналы в s -м слое (выходной слой сети).
4. Оценить полученное решение по формуле (5).
5. Округлить полученную оценку.

Примечание:

1. Округление в п.5 следует проводить до второго знака после запятой по следующему правилу. Если значение после третьего знака после запятой больше, чем 0.005, то округлить значение второго знака после запятой в большую сторону, если меньше или равно 0.005, то оставить второй знак после запятой без изменения. Например: $0.475 \approx 0.47$, $0.4785 \approx 0.48$, $0.4705 \approx 0.47$.

ПРИМЕР

Дано: $k=2$, $n_1=7$, $n_2=2$, $n_3=2$.

Функция активации – линейная. Функция подвыборки – по максимальному значению в подматрице выборки.

Матрица признаков исходного изображения M :

0.15	0.36	0	0.28	0.57	0.62	0.87
0.49	0.84	0.51	0.48	0.64	0.2	0.43
0.74	0	0.75	0.46	0.23	0.52	0.64
0.37	0.67	0.62	0.35	0.92	0.68	0.34
0.13	0.98	0.26	0.94	0.08	0.42	0.34
0.38	0.73	0.17	0.56	0.64	0.89	0.06
0.95	0.28	0.39	0.66	0.51	0.4	0.01

Ядра свёртки Y_1 и Y_2 :

1	0	-1	0
0	1	0	-1

Решение:

1. $n_1'' = 7$, $s=1$.
2. Определим размерность матриц в слое свёртки: $n_1' = n_1 - n_2 + 1 = 7 - 2 + 1 = 6$.
3. Создадим для каждого ядра свёртки матрицы M' 1-го слоя свертки размерностью 6×6 и заполним их следующим образом:

0.15	0.36	0	0.28	0.57	0.62	0.87
0.49	0.84	0.51	0.48	0.64	0.2	0.43
0.74	0	0.75	0.46	0.23	0.52	0.64
0.37	0.67	0.62	0.35	0.92	0.68	0.34
0.13	0.98	0.26	0.94	0.08	0.42	0.34
0.38	0.73	0.17	0.56	0.64	0.89	0.06
0.95	0.28	0.39	0.66	0.51	0.4	0.01

+

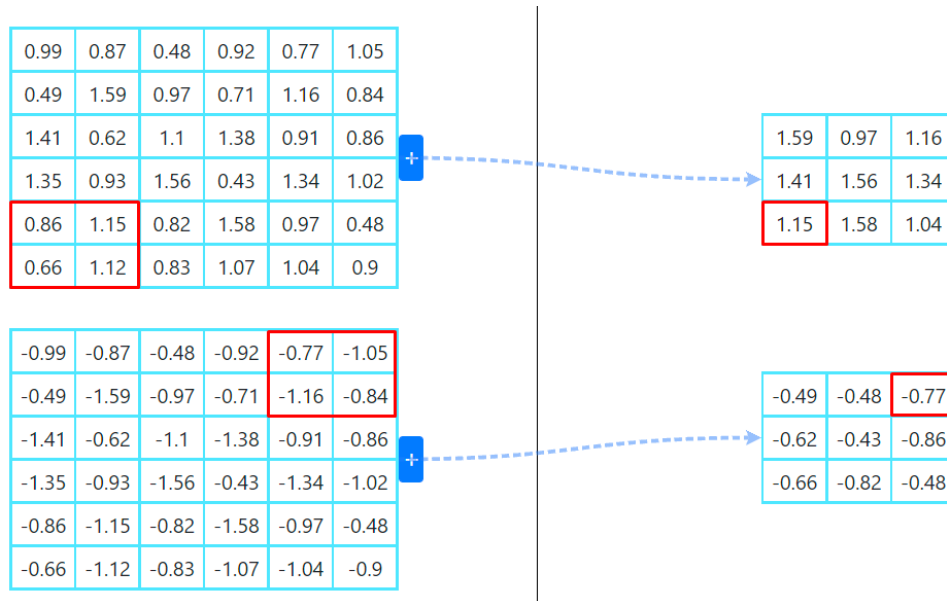
0.99	0.87	0.48	0.92	0.77	1.05
0.49	1.59	0.97	0.71	1.16	0.84
1.41	0.62	1.10	1.38	0.91	0.86
1.35	0.93	1.56	0.43	1.34	1.02
0.86	1.15	0.82	1.58	0.97	0.48
0.66	1.12	0.83	1.07	1.04	0.9

-0.99	-0.87	-0.48	-0.92	-0.77	-1.05
-0.49	-1.59	-0.97	-0.71	-1.16	-0.84
-1.41	-0.62	-1.10	-1.38	-0.91	-0.86
-1.35	-0.93	-1.56	-0.43	-1.34	-1.02
-0.86	-1.15	-0.82	-1.58	-0.97	-0.48
-0.66	-1.12	-0.83	-1.07	-1.04	-0.9

Здесь слева – исходная матрица M , а справа – набор матриц M'_{11} , M'_{12} . Например, тут для расчета выделенного элемента матрицы $m'_{11}(4,2)$ была использована выделенная в матрице M подматрица и 1-ое ядро свёртки:

$$m'_{11}(4,2) = f(0.67*1 + 0.62*0 + 0.98*0 + 0.26*1) = f(0.93) = 0.93.$$

4. Определим размерность матриц в слое подвыборки: $n_1'' = n_1' / 2 = 6 / 2 = 3$.
5. Создадим для каждой матрицы M' матрицу M'' для 1-го слоя подвыборки размерностью 3×3 и заполним их следующим образом:



Здесь слева – матрицы M'_{11} , M'_{12} , а справа – соответствующие им матрицы M''_{11} , M''_{12} . Например, тут для расчета выделенного элемента матрицы $m''_{11}(3,1)$ была использована подматрица 2×2 , выделенная в матрице M'_{11} , начиная с элемента $m'_{11}(5,1)$, а для расчета выделенного элемента матрицы $m''_{12}(1,3)$ была использована подматрица 2×2 , выделенная в матрице M'_{12} , начиная с элемента $m'_{12}(1,5)$, см. формулу (4):

$$m''_{11}(3,1) = \max(0.86, 1.15, 0.66, 1.12) = 1.15$$

$$m''_{12}(1,5) = \max(-0.77, -1.05, -1.16, -0.84) = -0.77$$

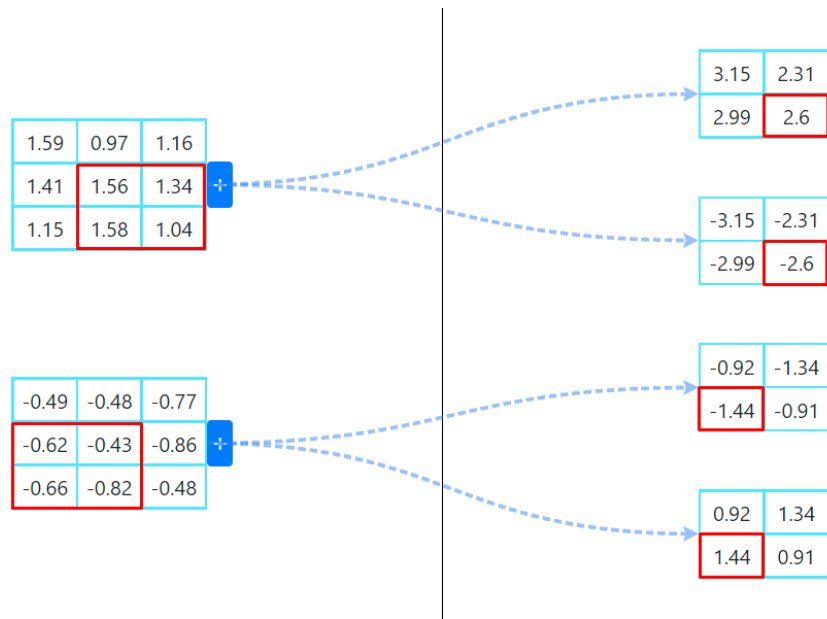
6. Определим $s = 2$, $n_1 = 3$.
7. Будем считать в качестве исходных матриц для построения 2-го слоя свёртки набор матриц M''_{11} , M''_{12} . Обозначим их как M_1 и M_2 .
8. Так как $n_1'' = 3$, то перейдем к следующей итерации по построению 2-го слоя свёртки.
9. Определим размерность матриц в слое свёртки: $n_1' = n_1 - n_2 + 1 = 3 - 2 + 1 = 2$.
10. Создадим для каждого ядра свёртки матрицы размерностью 2×2 : для исходной матрицы M_1 - матрицы M'_{21} , M'_{22} , а для исходной матрицы M_2 - матрицы M'_{23} , M'_{24} . Заполним их аналогичным образом, например:

$$m'_{21}(2,2) = f(1.56 * 1 + 1.34 * 0 + 1.58 * 0 + 1.04 * 1) = f(2.6) = 2.6$$

$$m'_{22}(2,2) = f(1.56 * (-1) + 1.34 * 0 + 1.58 * 0 + 1.04 * (-1)) = f(-2.6) = -2.6$$

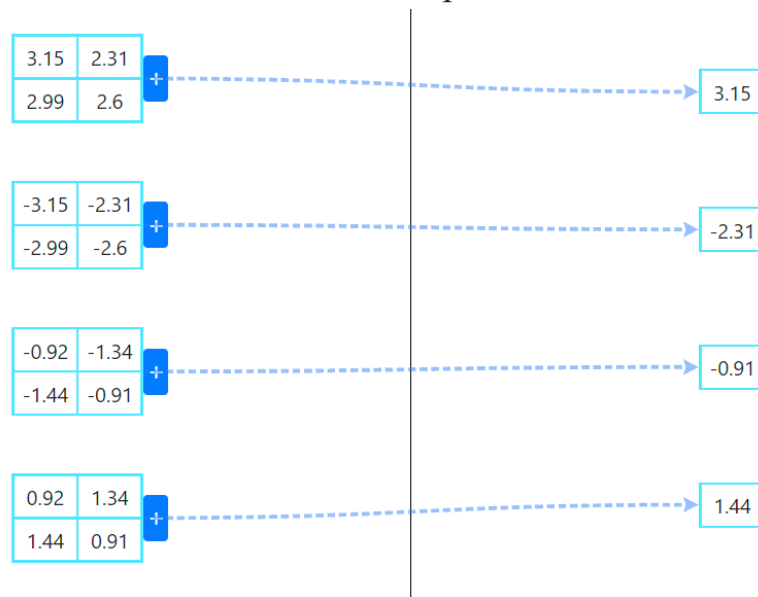
$$m'_{23}(1,2) = f(-0.62 * 1 + (-0.43) * 0 + (-0.66) * 0 + (-0.82) * 1) = f(-1.44) = -1.44$$

$$m'_{24}(1,2) = f(-0.62 * (-1) + (-0.43) * 0 + (-0.66) * 0 + (-0.82) * (-1)) = f(1.44) = 1.44$$



11. Определим размерность матриц в слое подвыборки: $n_1'' = n_1' / 2 = 2 / 2 = 1$.

12. Создадим для каждой матрицы M' матрицу M'' для 2-го слоя подвыборки размерностью 1×1 и заполним их аналогичным образом с выбором максимального значения в каждой матрице:



13. Определим $s = 3$, $n_1 = 1$.

14. Будем считать в качестве исходных матриц для построения 3-его слоя свёртки набор матриц $M''_{21}, M''_{22}, M''_{23}, M''_{24}$. Обозначим их как M_1, M_2, M_3, M_4 .

15. Так как $n_1'' = 3$, то закончим цикл по построению чередующихся слоев свёртки и слоев подвыборки и перейдем к построению выходного слоя сети.

16. Количество нейронов в выходном слое $n = 4$ (по количеству матриц). Их выходные сигналы: $output(1) = f(3.15) = 3.15$, $output(2) = f(-2.31) = -2.31$, $output(3) = f(-0.91) = -0.91$, $output(4) = f(1.44) = 1.44$.

17. Рассчитаем оценку полученного решения по формуле (5) и округлим ее значение до второго знака после запятой:

$$\mathbf{MSE} = ((3.15 - 1)^2 + (-2.31)^2 + (-0.91)^2 + (1.44 - 1)^2) / 4 = \mathbf{2,745075 \sim 2.75}$$