

## Convolutional Neural Networks, CNN

Свёрточная нейронная сеть — специальная архитектура искусственных нейронных сетей, нацеленная на эффективное распознавание образов, входит в состав технологий глубокого обучения.

### Архитектура.

Основная идея CNN заключается в использовании сверточных слоев, которые применяют фильтры (ядра свертки) для извлечения признаков из входных данных. Эти фильтры могут обнаруживать различные особенности, такие как края, текстуры и узоры, что делает CNN особенно подходящими для анализа визуальной информации.

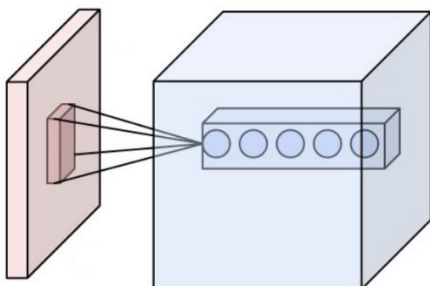
Рассмотрим основные компоненты архитектуры.

#### 1. Входной слой (Input Layer).

Принимает входные данные, обычно изображения. Данные представляются в виде тензора. Например, для RGB-изображения размером 28x28 размерность тензора будет (28, 28, 3).

#### 2. Сверточные слои (Convolutional Layers).

Основная задача сверточных слоев — извлечение признаков из входных данных. В операции свёртки используется ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (в самом начале — непосредственно по входным данным). То есть для различных нейронов выходного слоя используется одна и та же матрица весов, которую также называют ядром свёртки. Её интерпретируют как графическое кодирование какого-либо признака.



Для сверточного слоя задается несколько параметров:

Размер фильтра (ядро): например,  $3 \times 3$  или  $5 \times 5$ .

Шаг (stride): определяет, на сколько сдвигается фильтр после каждой операции свертки.

Количество фильтров: определяет, сколько признаков будет извлечено.

### 3. Функции активации

После каждой свертки применяется функция активации, которая представляет собой некоторую нелинейную функцию (обычно ReLU – Rectified Linear Unit).

### 4. Слои подвыборки (Pooling Layers)

Уменьшают размерность данных и помогают извлекать наиболее важные признаки, уменьшая вычислительную нагрузку. Для изображений преобразования затрагивают непересекающиеся прямоугольники или квадраты, каждый из которых ужимается в один пиксель, при этом выбирается пиксель, имеющий максимальное значение.

### 5. Полносвязные слои (Fully Connected Layers)

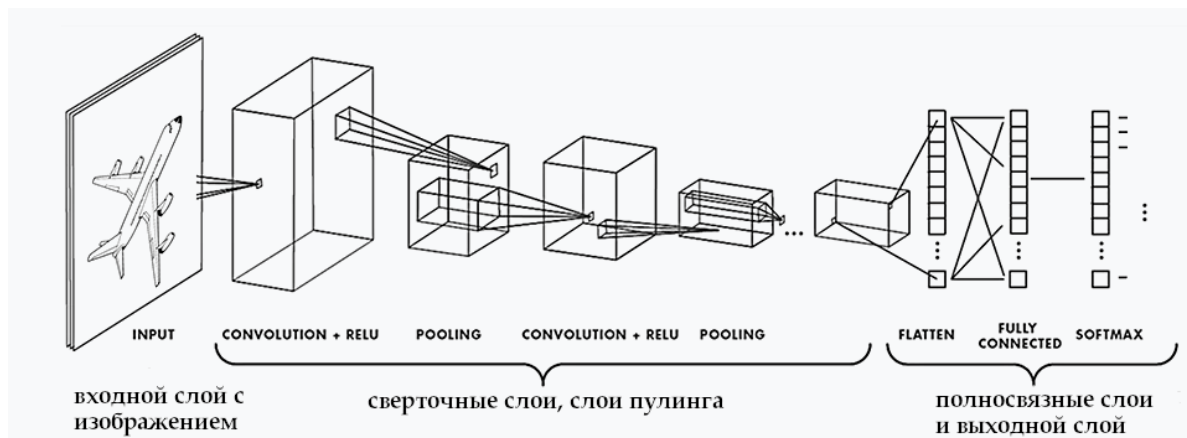
Используются в конце сети для классификации. Каждый нейрон в этом слое связан со всеми нейронами предыдущего слоя. Они принимают получившиеся ранее абстрактные признаки и комбинируют их для предсказания класса.

### 6. Слой регуляризации (Regularization Layer)

Этот слой случайным образом отключает часть нейронов во время обучения, чтобы предотвратить переобучение. В данном слое устанавливается вероятность отключения нейронов.

### 7. Выходной слой (Output Layer)

Последний слой, который выдает предсказание. Для многоклассовой классификации используется softmax, который преобразует выходные значения в вероятности принадлежности к каждому классу.



### Пример структуры CNN.

Вход: 224x224x3 (RGB-изображение размером 224x224)

Сверточный слой: 32 фильтра 3x3, ReLU → 224x224x32

Max Pooling: 2x2 → 112x112x32

Сверточный слой: 64 фильтра 3x3, ReLU → 112x112x64

Max Pooling: 2x2 → 56x56x64

Полносвязный слой: 128 нейронов, ReLU

Dropout:  $p = 0.5$

Выходной слой: softmax (количество классов)

## Области применения CNN, актуальность.

### 1. Обработка изображений и классификация.

Примеры применения:

Классификация изображений. Определение объектов на изображениях (например, классификация животных в фото).

Обнаружение объектов. Использование CNN для нахождения и маркировки объектов на изображениях.

Успешность: CNN демонстрирует выдающиеся результаты в соревнованиях по классификации изображений, таких как ImageNet.

Актуальность: CNN остаются актуальными в данной области. Однако для задач, требующих реального времени, исследуются более легковесные архитектуры (например, MobileNet).

### 2. Обработка видео.

Примеры применения:

Анализ видеопотока. Определение действий людей на видео (например, распознавание жестов).

Распознавание лиц. Использование CNN для идентификации и верификации лиц в видеозаписях.

Успешность: CNN в сочетании с рекуррентными нейронными сетями (RNN) или 3D-свертками (например, C3D) показали хорошие результаты в задачах анализа видео.

Актуальность: с ростом данных и требований к скорости обрабатываются более оптимизированные архитектуры и подходы.

### 3. Медицинская диагностика

Примеры применения:

Анализ медицинских изображений. Использование CNN для диагностики заболеваний по снимкам (например, рентгеновские снимки, МРТ).

Автоматическое выявление опухолей. CNN помогают в раннем обнаружении рака на основании изображений.

Успешность: исследования показывают, что CNN могут достигать или даже превышать точность профессиональных радиологов в некоторых задачах.

Актуальность: актуальность CNN в медицине высока. Однако с увеличением объёма данных и необходимости объяснимости моделей начинают использоваться более интерпретируемые методы.

#### 4. Обработка аудио и сигналов

Примеры применения:

Аудиоклассификация. Использование CNN для классификации аудиофайлов (например, музыка, шумы).

Распознавание речи. CNN могут быть частью систем, которые переводят речь в текст.

Успешность: CNN показывают высокую точность в задачах распознавания речи и классификации звуков.

Актуальность: хотя CNN и используются, для задач обработки сигналов начинают применяться другие подходы, такие как трансформеры, которые показывают высокую эффективность.

#### Вывод.

CNN – один из лучших алгоритмов по распознаванию и классификации изображений. По сравнению с полносвязной нейронной сетью (типа перцептрона) в CNN гораздо меньшее количество настраиваемых весов, так как одно ядро весов используется целиком для всего изображения

вместо того, чтобы делать для каждого пикселя входного изображения свои персональные весовые коэффициенты.

Сверточные нейронные сети остаются актуальными благодаря своей эффективности в решении задач, связанных с изображениями. Они активно развиваются, и новые архитектуры, основанные на CNN, предлагают улучшенные характеристики производительности и точности.