

Сверточные нейронные сети

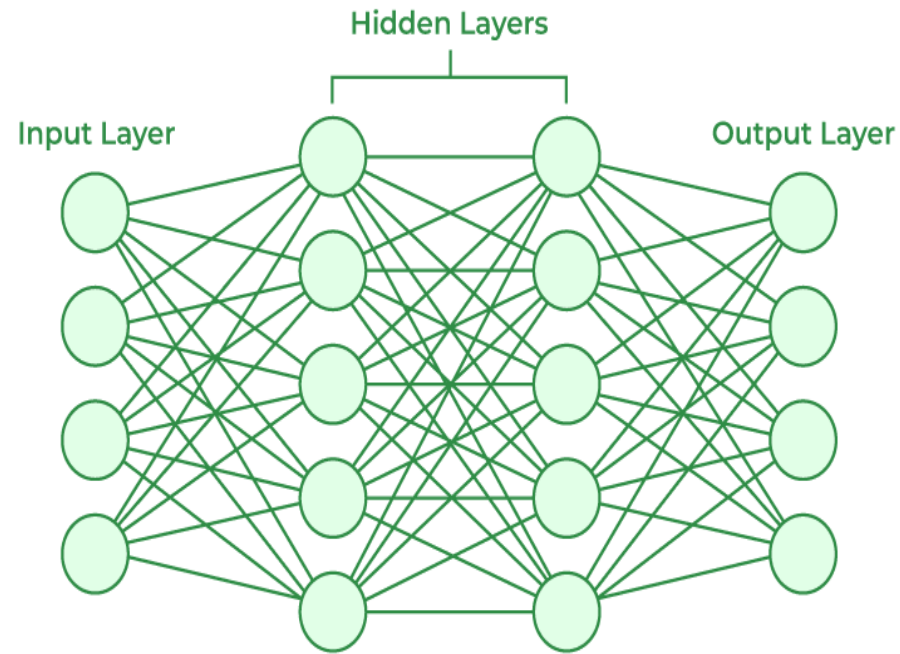
Студенты:
Скворцов В.С.
Горюнов М.Ю.

План

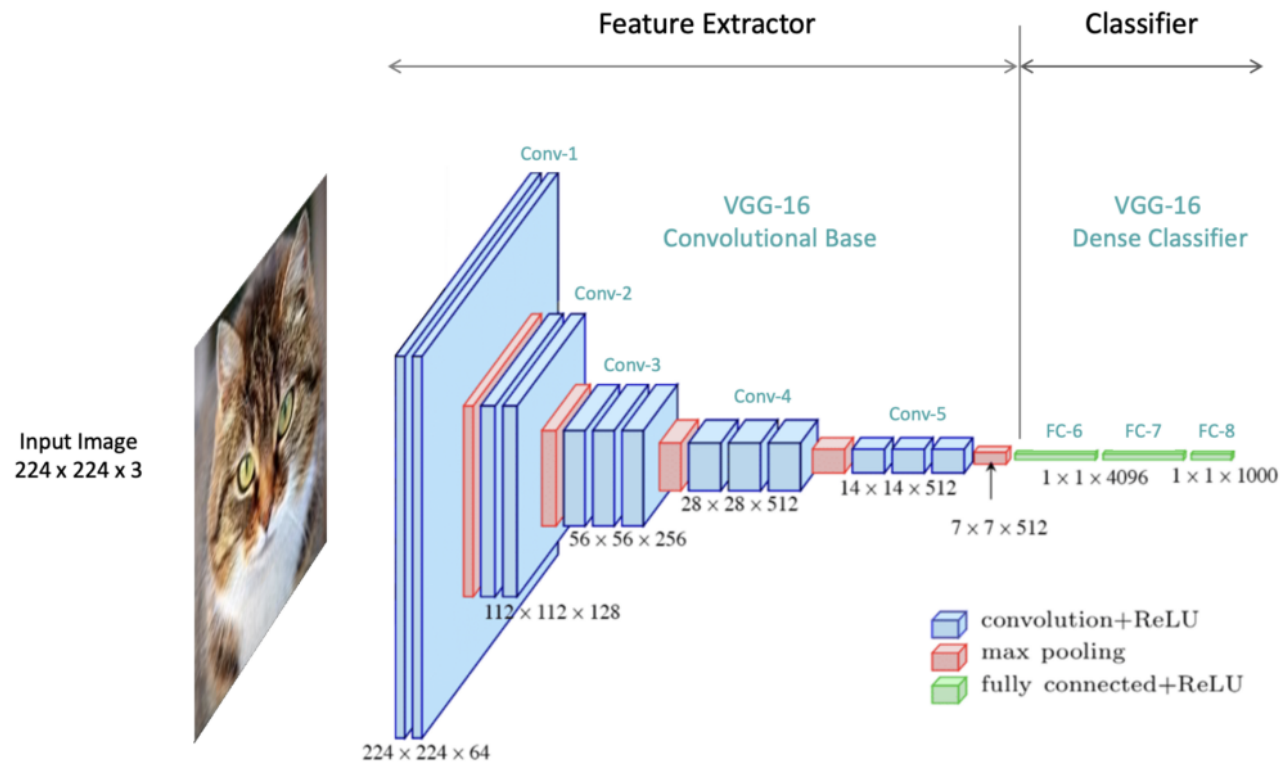
1. Введение. Общее описание нейронных сетей и их применения.
2. История создания и развития CNN.
3. Архитектура и принцип работы CNN, основные компоненты.
4. Адаптация обратного распространения ошибки в CNN.
5. Сравнение CNN с обычным многослойным перцептроном.
6. Области применения: распознавание изображений, обработка видео, распознавание речи.
7. Современные тенденции и перспективы развития CNN.
8. Заключение.

Нейронные сети

Нейронные сети — это модели машинного обучения, вдохновленные биологическими нейронными сетями. Эти модели используются для обучения компьютеров выполнению задач, которые обычно требуют человеческого вмешательства, таких как распознавание речи, распознавание изображений, прогнозирование и многие другие.



Сверточные нейронные сети (CNN)

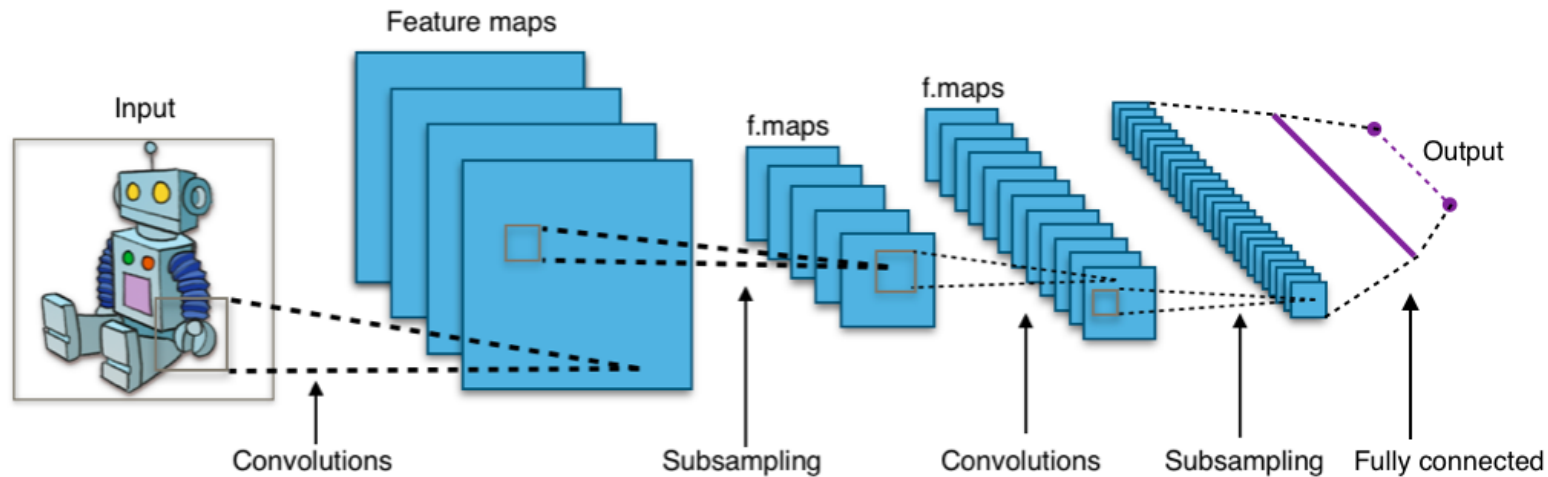


История создания

Сверточные нейронные сети, вдохновленные биологическими процессами, были впервые представлены в 1980-х годах **Янном ЛеКуном**. В 1990-х годах он разработал **LeNet**, что стало революционным шагом в области сверточных нейронных сетей, заменив стандартный подход к обучению нейронных сетей классификации изображений.



Архитектура и принцип работы



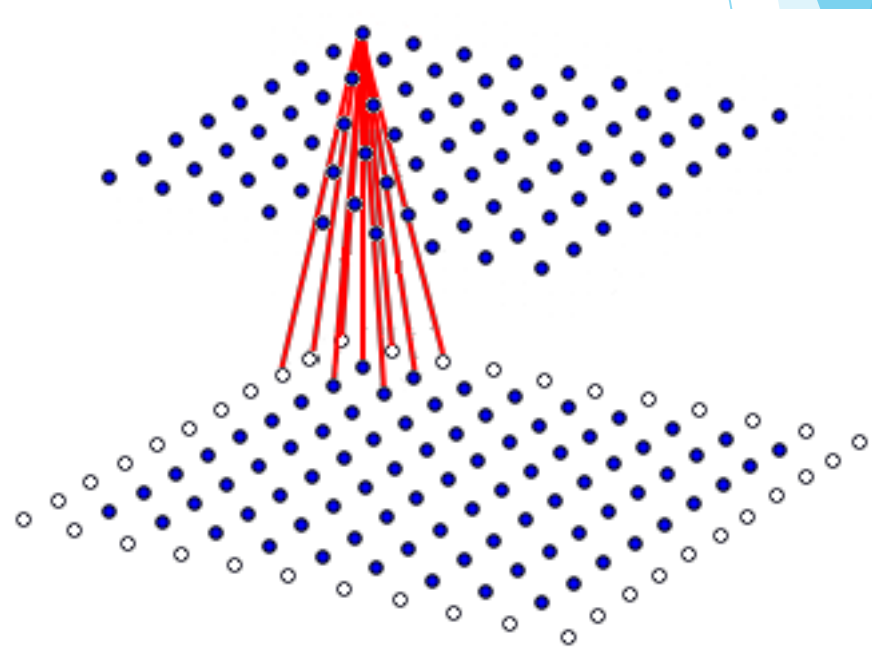
В CNN входной слой представляет собой тензор с формой: (количество входов) \times (высота входа) \times (ширина входа) \times (каналы входа).

В **операции свёртки** используется лишь ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою.

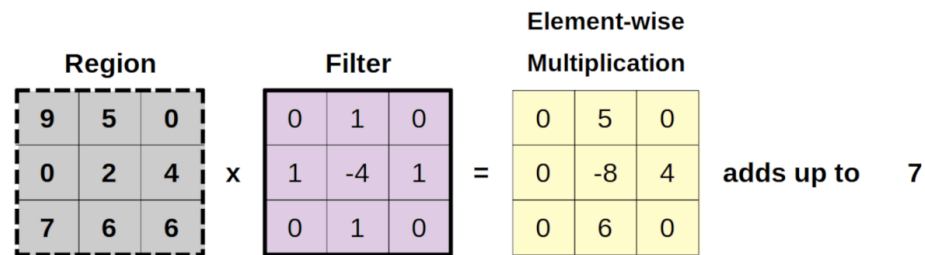
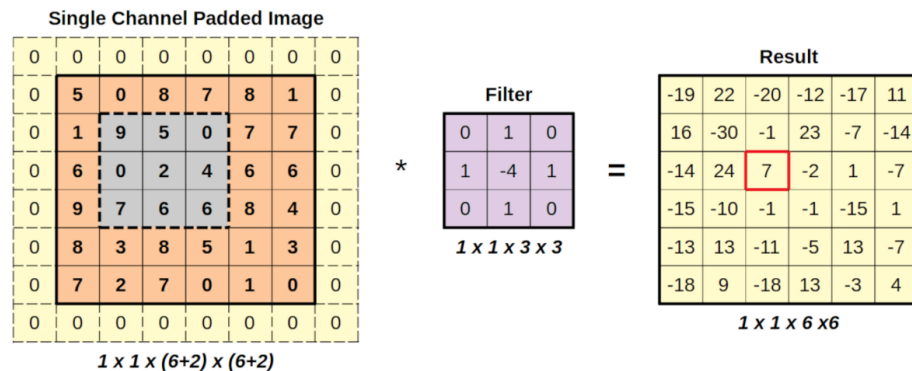
Архитектура и принцип работы

Для различных нейронов выходного слоя используется одна и та же матрица весов, которую также называют **ядром свёртки**. Следующий слой, получившийся в результате операции свёртки такой матрицей весов, показывает наличие некоторого признака в обрабатываемом слое.

Проход каждым набором весов формирует свой собственный экземпляр карты признаков, делая нейронную сеть многоканальной (много независимых карт признаков на одном слое).

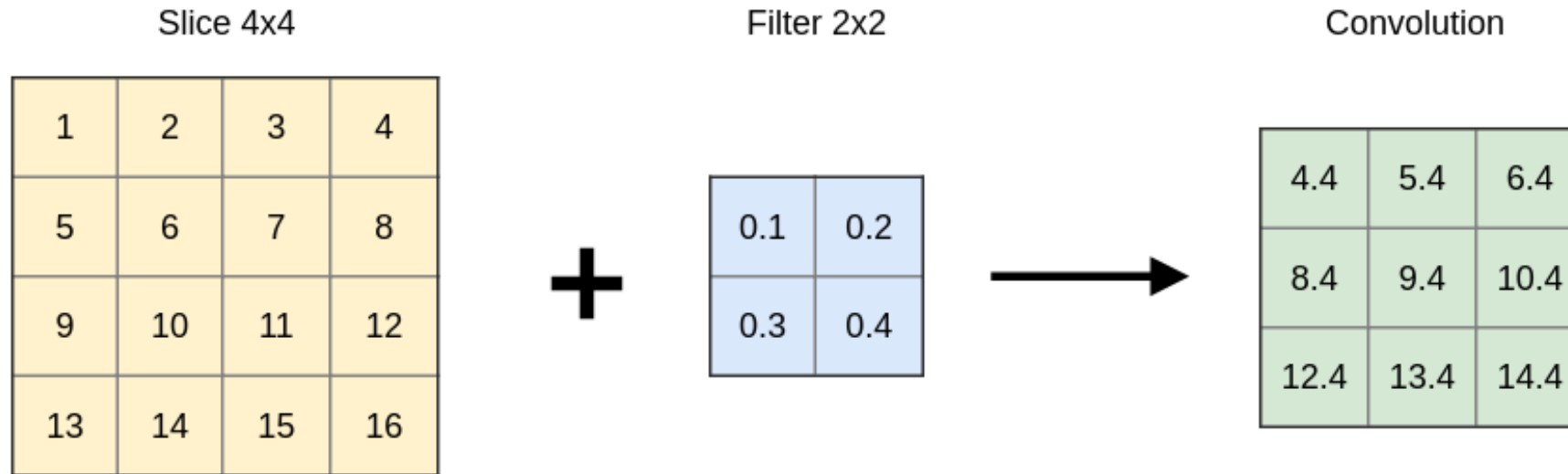


Дополнение изображения (padding)



Пиксели, расположенные по краям, участвуют в операции свертки только один раз. Для того, чтобы избежать недооценивания изображение обрамляется по краям набором пикселей с нулевыми значениями. Посредством этого исходные краевые пиксели будут участвовать в свертке то же число раз, что и любые другие пиксели из внутренней части изображения.

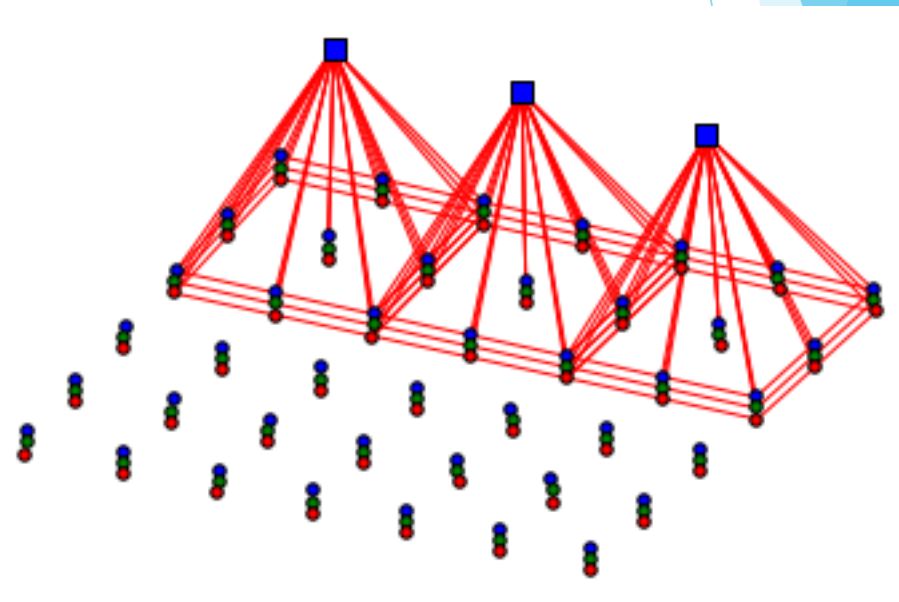
Слой свертки



Слой свёртки — это основной блок свёрточной нейронной сети. Слой свёртки включает в себя для каждого канала свой фильтр, ядро свёртки которого обрабатывает предыдущий слой по фрагментам. Весовые коэффициенты ядра свёртки неизвестны и устанавливаются в процессе обучения.

Слой свертки

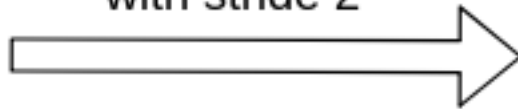
Особенностью свёрточного слоя является сравнительно небольшое количество параметров. Так например, если исходное изображение имеет размерность 100×100 пикселей по трём каналам (это значит 30 000 входных нейронов), а свёрточный слой использует фильтры с ядром 3×3 пикселя с выходом на 6 каналов, тогда в процессе обучения определяется только 9 весов ядра, однако по всем сочетаниям каналов, то есть $9 \times 3 \times 6 = 162$, в таком случае данный слой требует нахождения только 162 параметров.



Субдискретизация (pooling)

1	3	10	2
5	4	6	14
1	2	17	8
3	3	10	12

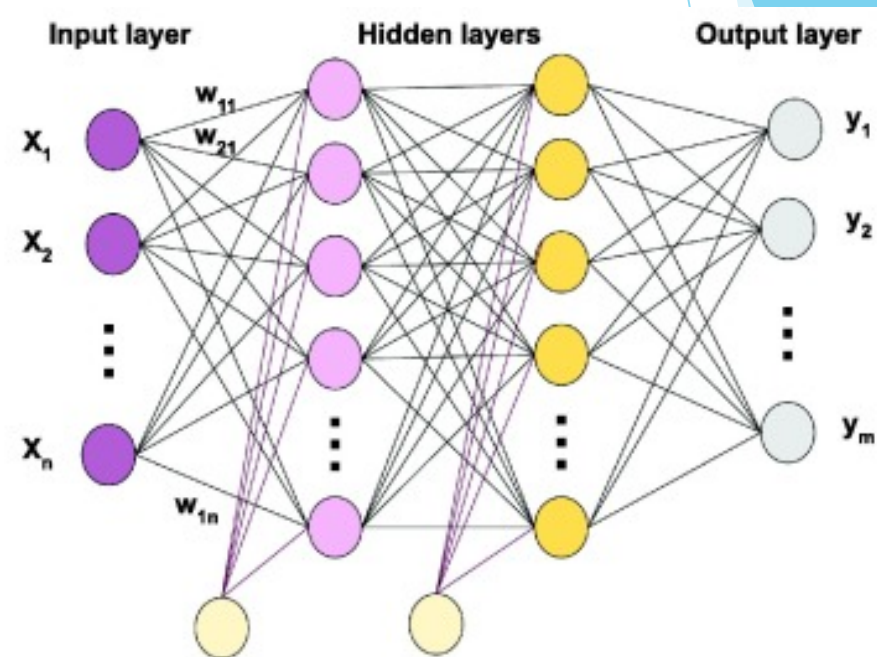
2x2 average pooling
with stride 2



3.25	8
2.25	11.75

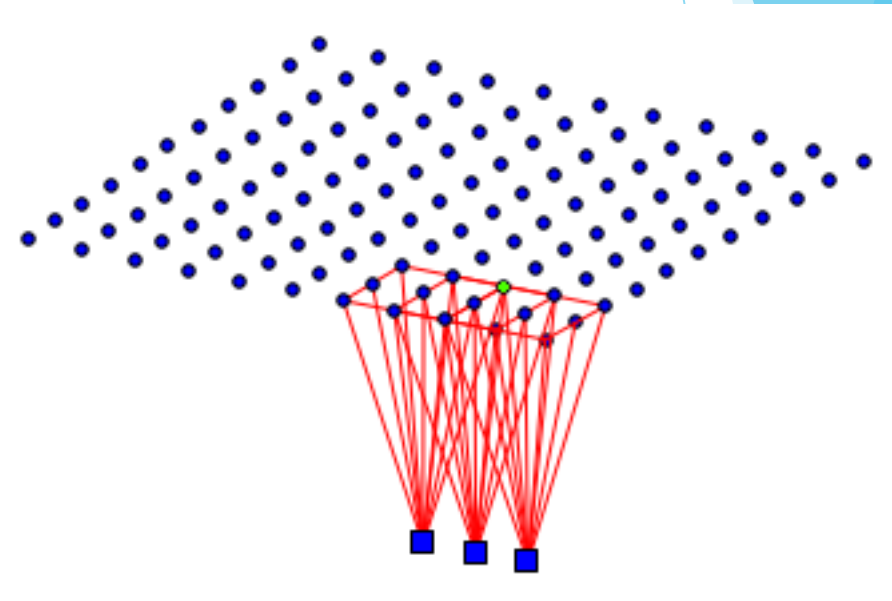
Полносвязная сеть

После нескольких проходов свёртки изображения и уплотнения с помощью субдискретизации система перестраивается от конкретной сетки к более абстрактным данным. Эти данные объединяются и передаются на обычную полносвязную нейронную сеть, которая тоже может состоять из нескольких слоёв.



Обратное распространение ошибки

Для каждого слоя сети вычисляются градиенты функции потерь по отношению к весам. Градиенты для весов свертки учитывают входные данные и градиенты на выходе, после чего происходит перевычисление весов. Нейроны распространяют градиент в обратном направлении только в своих группах, соответствующих фильтру. Обратный проход - это деконволюция.



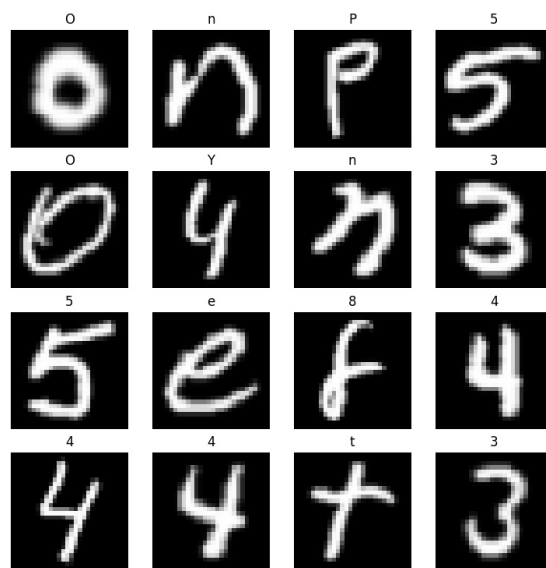
CNN и перцептрон

Главное отличие между CNN и перцептронами заключается в их способности обрабатывать изображения и другие входные данные с сетчатой структурой. В то время как прямые сети обрабатывают каждый входной элемент независимо, CNN учитывают пространственные взаимосвязи между входными элементами, что позволяет им выявлять локальные особенности, такие как линии, точки, края и углы на начальных этапах модели, а также более абстрактные объекты, такие как носы, глаза и колеса на более поздних этапах.

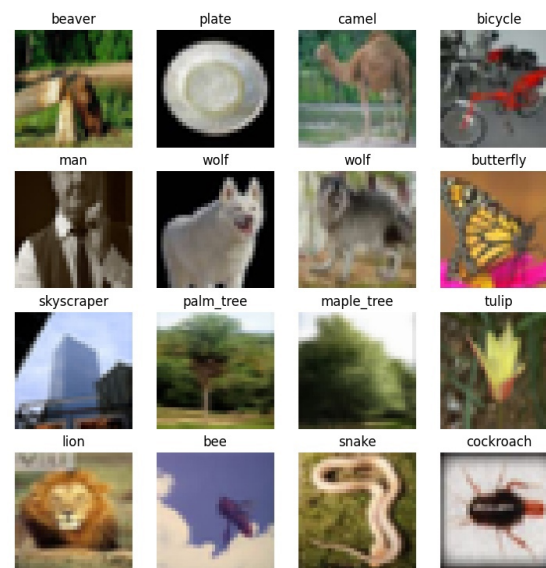
Датасеты



MNISTFashion



EMNIST



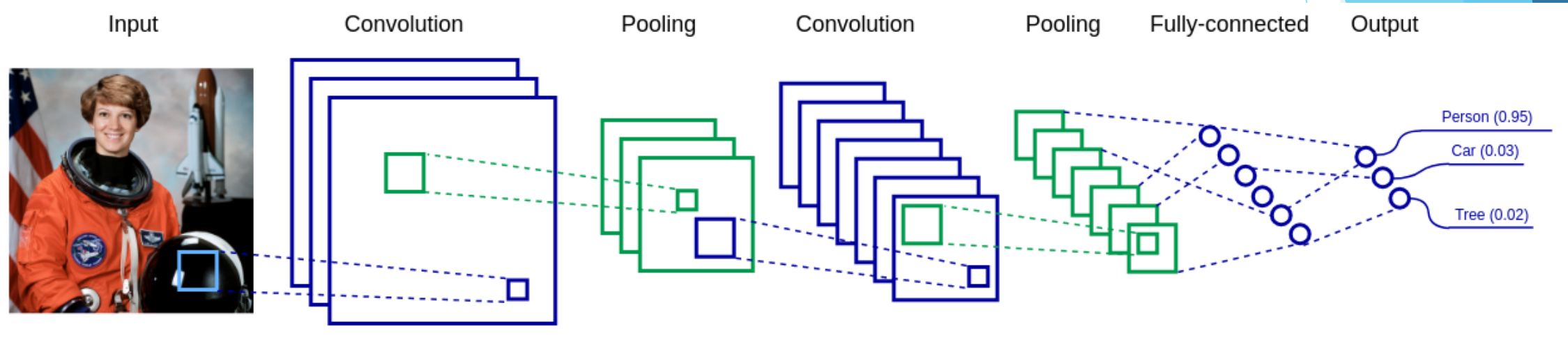
CIFAR1000

Результаты сравнения

Датасет	Модель	Число параметров	Loss	Accuracy
EMNIST	FF	1901615	0.710763	85.831207
EMNIST	CNN	1682895	0.529978	88.249362
FashionMNIST	FF	1863690	0.426498	90.455272
FashionMNIST	CNN	1644970	0.299659	92.501997
CIFAR100	FF	4298852	3.634229	21.605431
CIFAR100	CNN	2229316	3.554307	28.464457

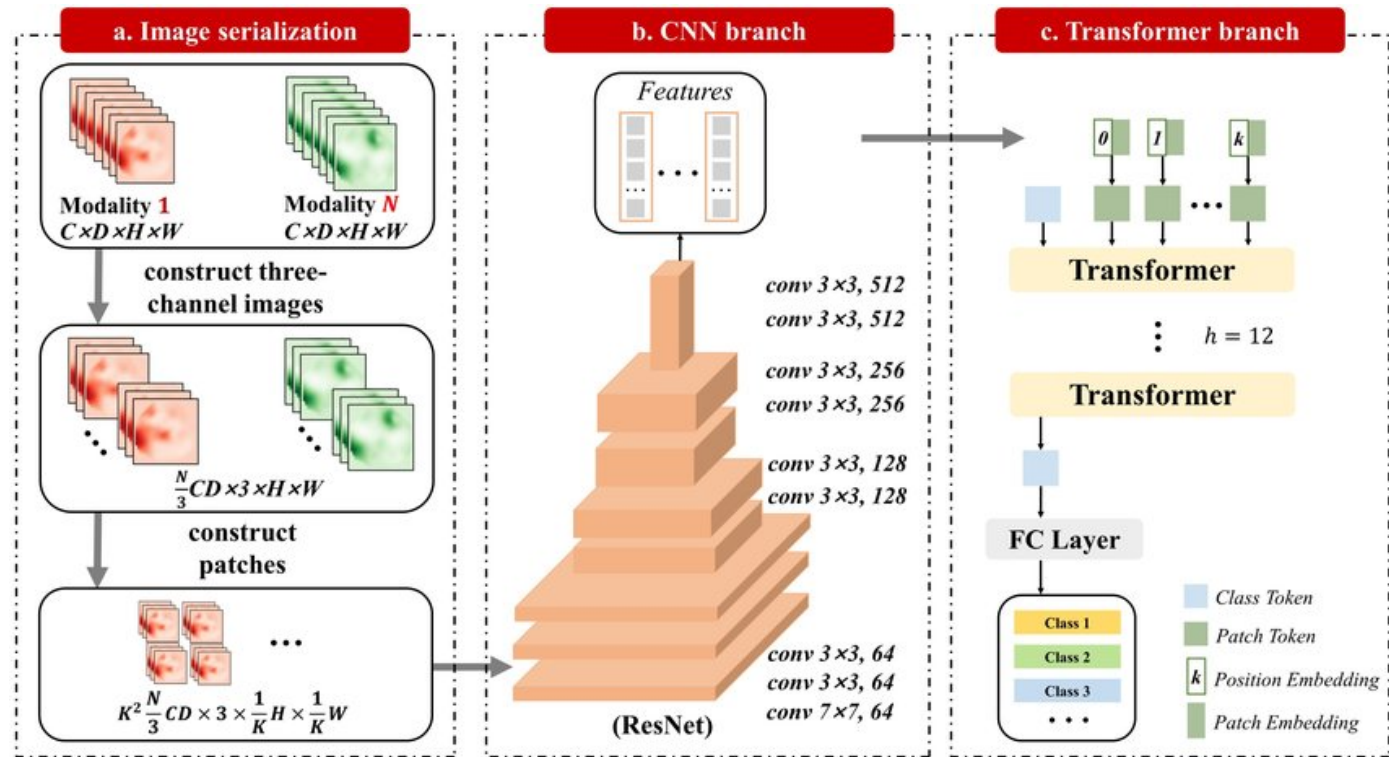


Применение CNN



CNN стали основой для многих современных систем компьютерного зрения, где они используются для распознавания образов, обработки видео, анализа изображений и многих других задач. Сверточные нейронные сети также используются в области распознавания речи, где они помогают распознавать и интерпретировать звуковые сигналы.

Применение CNN



Заключение

CNN продолжают развиваться и адаптироваться к новым вызовам и возможностям, предлагая более эффективные и инновационные решения в области обработки данных. Они стали основой для многих современных систем компьютерного зрения и продолжают преобразовывать многие области, включая автоматическое вождение, медицинскую диагностику, биометрию, робототехнику и многие другие.

В будущем мы ожидаем увидеть еще больше инноваций и прогресса в области сверточных нейронных сетей, поскольку исследователи продолжают исследовать новые методы, алгоритмы и архитектуры для улучшения их производительности и эффективности.