

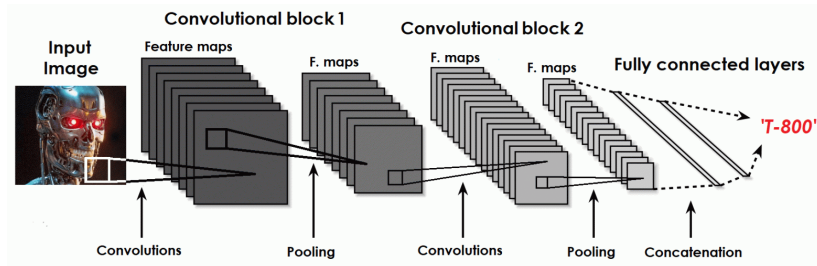
# Свёрточные нейронные сети для обработки изображений: классификация и сегментация

Санкт-Петербургский государственный университет  
Кафедра статистического моделирования

28 октября 2025

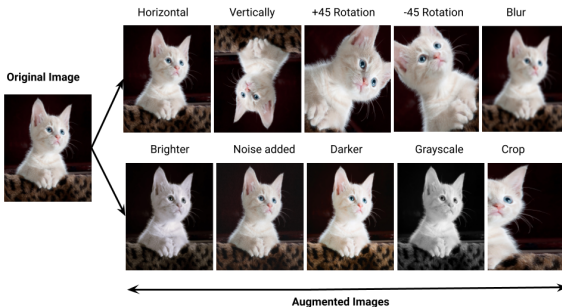
# Свёрточные нейронные сети (CNN)

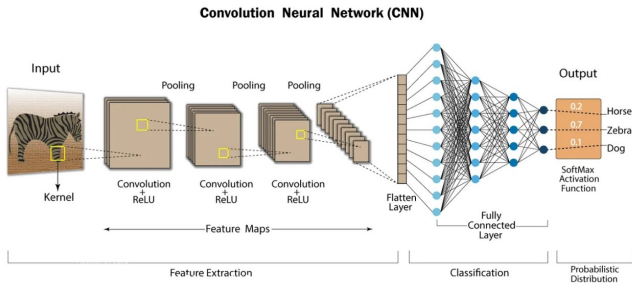
- Класс нейронных сетей, эффективно работающих с изображениями (и другими объектами, в которых важна пространственная связь);
- Используют свёртки (свёрточные слои) для извлечения признаков;
- Основные применения: классификация, сегментация, детекция объектов.



# Проблемы обработки изображений полносвязными сетями

- Изображения содержат пространственную структуру, которую теряют обычные полносвязные сети;
- Если изображение имеет высокое разрешение, то полносвязная сеть содержит слишком много параметров;
- При небольшом изменении изображения (сдвиг, поворот) входной слой обычной сети полностью меняется, хотя суть осталась прежней.



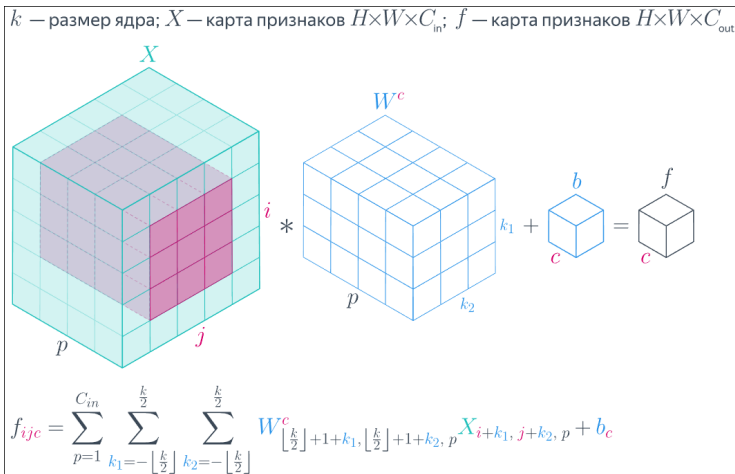


- Свёрточные слои (Conv);
- Пулинг (Pooling, часто MaxPooling или AvgPooling);
- Нелинейности (ReLU, sigmoid, tanh и т.д.);
- Полносвязные слои.

С помощью свёрточных слоёв и пулинга изображение сводится до вектора размерности 1, то есть формируется вход для полносвязной сети.

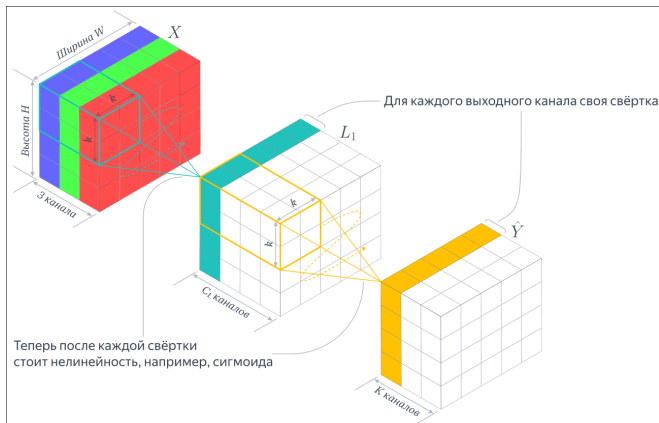
# Свёрточная операция

- Свертка — это скользящее умножение ядра на участок изображения.
- Позволяет выделить локальные признаки (границы, текстуры).



# Каналы и свёртки

- Изначально в цветном изображении содержится три канала для каждого цвета;
- С каждой свёрткой мы уменьшаем размер изображения, но увеличиваем количество каналов;



Хотим повысить количество каналов в каждой свёртке, чтобы выделять больше признаков. Проблема: слишком много параметров, решение: уменьшить каналы

- Max Pooling — берёт максимум из области
- Average Pooling — усреднение значений
- Снижает вычислительные затраты и переобучение

- ReLU:  $f(x) = \max(0, x)$  — ускоряет обучение
- Batch Normalization стабилизирует распределение активаций



# Задача классификации

- Вход: изображение
- Выход: метка класса
- Пример: распознавание животных, предметов, сцен

- LeNet (1998)
- AlexNet (2012)
- VGGNet (2014)
- ResNet (2015)

- Использует остаточные связи (skip-connections)
- Позволяет строить очень глубокие сети
- Решает проблему затухающего градиента

- Для классификации: Cross-Entropy Loss
- Оптимизаторы: SGD, Adam
- Используется backpropagation

# Что такое сегментация

- Пиксельная классификация изображения
- Цель: выделить объекты на уровне пикселей
- Применения: медицина, автономные автомобили, спутниковые снимки

- Семантическая сегментация
- Инстанс-сегментация
- Паноптическая сегментация

- U-Net
- SegNet
- DeepLab

- Симметричная encoder-decoder архитектура
- Skip-соединения между слоями
- Широко применяется в медицине



- Использует dilated convolutions
- Применяет CRF для уточнения границ объектов
- Высокая точность на сложных сценах

- Для классификации: Accuracy, Precision, Recall, F1
- Для сегментации: IoU, Dice coefficient

- Распознавание лиц, жестов, объектов
- Медицинская диагностика
- Обработка видео и спутниковых изображений

- CNN — ключевой инструмент компьютерного зрения
- Эффективны для классификации и сегментации
- Продолжают активно развиваться