

# Архитектура Inception (GoogLeNet)

## Эффективные глубокие сети для компьютерного зрения

Диц Даниил Денисович

28 марта 2025 г.

# Обзор архитектуры

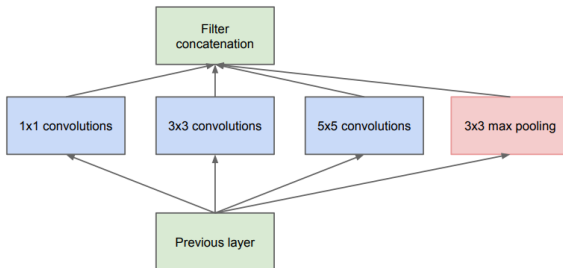
- Победитель ImageNet 2014 (Топ-5 ошибка 6.67%)
- В 9 раз меньше параметров, чем AlexNet (7 млн vs 60 млн)
- В 22 раза компактнее VGG16 (7 млн vs 138 млн параметров)
- Основные инновации:
  - Модуль Inception
  - 1x1 свёртки для уменьшения параметров
  - Замена полносвязных слоёв на глобальный пулинг
  - Вспомогательные классификаторы

- Проблемы предыдущих архитектур:
  - Большое число параметров
  - Проблема затухания градиента
  - Высокие вычислительные затраты
- Решения Inception:
  - Параллельные свёртки разного масштаба
  - Оптимизация параметров через 1x1 свёртки
  - Глобальный средний пулинг вместо FC-слоёв

# Модуль Inception

## Базовая структура

- Параллельные операции:
  - 1x1 свёртка
  - 3x3 свёртка
  - 5x5 свёртка
  - 3x3 max-pooling
- Конкатенация результатов



# Оптимизация параметров через 1x1 свёртки

## Пример вычисления параметров

- Без оптимизации (5x5 свёртка):

$$5 \times 5 \times 128 \times 256 = 819,200$$

- С 1x1 свёрткой (редукция до 64 каналов):

$$1 \times 1 \times 128 \times 64 + 5 \times 5 \times 64 \times 256 = \\ 8,192 + 409,600 = 417,792$$

## Эффект

Сокращение параметров на 49% при сохранении ёмкости модели

# Улучшенный модуль Inception

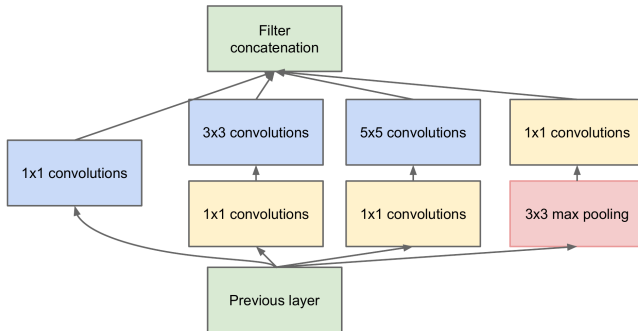


Рис.: Модуль Inception с 1x1 свёртками

- 1x1 свёртки перед 3x3 и 5x5 свёртками
- 1x1 свёртка после max-pooling
- ReLU активации после каждой операции

# Замена FC-слоёв на глобальный пулинг

- Проблема FC-слоёв:
  - 89% параметров VGG16
  - 95% параметров AlexNet
- Решение:
  - Глобальный средний пулинг
  - 1 FC-слой вместо нескольких
- Результат:
  - Только 15% параметров в FC-слоях
  - Улучшение точности на 0.6%

# Вспомогательные классификаторы

- Решают проблему затухания градиента
- Добавляются к промежуточным слоям
- Состав:
  - 5x5 average pooling
  - 1x1 свёртка (128 фильтров)
  - FC-слой (1024 нейрона)
  - Dropout (70%)
  - Softmax классификатор
- Вклад в общий loss: 0.3 веса

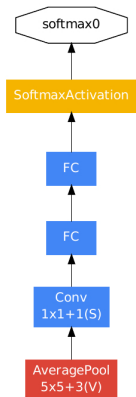


Рис.: Схема  
вспомогатель-  
ного  
классификато-



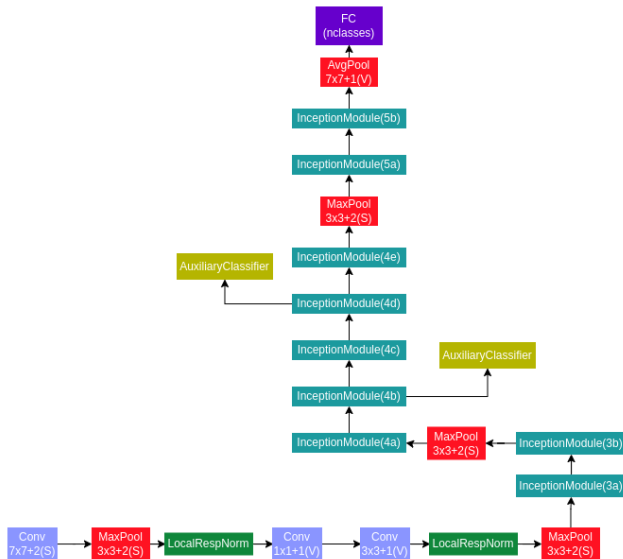


Рис.: Полная схема GoogLeNet

# Архитектура GoogLeNet

- 9 модулей Inception
- 2 вспомогательных классификатора
- Глобальный средний пулинг
- Всего 22 слоя с параметрами

## Сравнение с другими архитектурами

Модель	Параметры	Тор-1 ошибка	Тор-5 ошибка
AlexNet (2012)	60M	37.5%	17.0%
VGG16 (2014)	138M	28.5%	9.9%
GoogLeNet (2014)	7M	29.9%	10.8%

Таблица: Сравнение на ImageNet 2014

- Лучшее соотношение точности и параметров
- Эффективное использование вычислений

# Заключение

- Ключевые достижения:
  - Инновационный модуль Inception
  - Эффективное управление параметрами
  - Решение проблемы затухания градиентов
- Влияние на развитие CNN:
  - Стандарт для современных архитектур
  - Вдохновила ResNet, Inception-v2/v3/v4
- Применение:
  - Классификация изображений
  - Обнаружение объектов
  - Перенос стилей