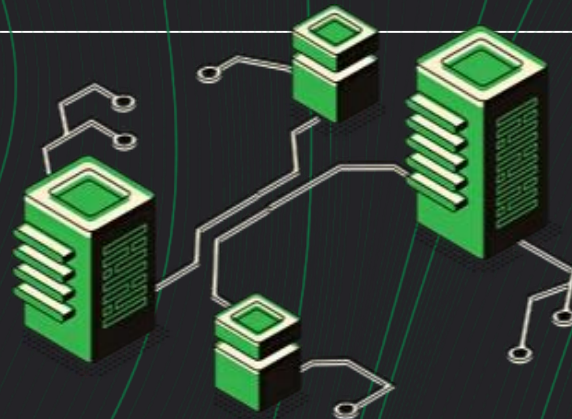




**SKILLFACTORY**

# Обзор сложных архитектур нейронных сетей



---

**Шпилевский Яромир**

Ведущий разработчик First Line Software

## Agenda

- Проблема доступности данных и подготовки датасетов.
- ImageNet. Доступность данных.
- Различные подходы в сложных архитектурах.
  - LeNet
  - AlexNet
  - VGG
  - ResNet
  - InceptionNet
  - XceptionNet
  - GoogLeNet
  - MobileNet
  - NASNet

## Проблема доступности данных и подготовки датасетов

- Нейронные сети выводят закономерности из датасетов.
- 80% работы – подготовить датасет: собрать, проверить, преобразовать к необходимому виду и т.п.
- Часто датасет собран под какую-то узкую задачу, решение получается тоже недостаточно общим.
- Мало публично доступных датасетов высокого качества.
- Нужны высококачественные эталонные датасеты для каждого класса задачи.

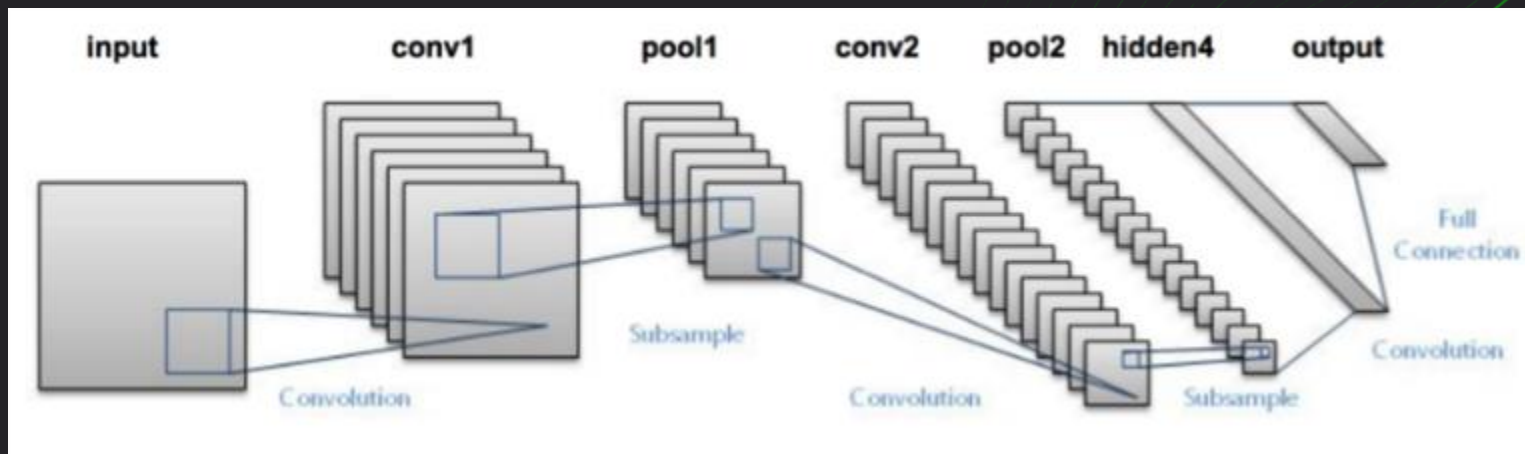
## ImageNet

- Для компьютерного зрения была предпринята попытка выработать эталонный датасет – ImageNet. [1]
- Fei Fei Li посвятила этому проекту очень много времени.
  - Работала над этим проектом в Princeton University.
  - Сейчас – профессор Stanford University.
  - И Chief Scientist Google Cloud.
- 14 000 000 изображений.
- 20 000 классов.
- С разметкой объектов.
- Таксономия классов основывается на WordNet.
- Классификация и разметка осуществлялась с помощью crowdsourcing платформы Amazon Mechanical Turk.
- Появился качественный публичный датасет для исследований.
- Создание ImageNet произвело взрывной рост разработок в области компьютерного зрения.
- Стал использоваться в качестве эталонного benchmark задач компьютерного зрения.
  - В частности, классификации.



## LeNet

- Создатель – Yann LeCun. [\[2\]](#)
- 2 свёрточных слоя.
- 2 max pooling слоя.
- 1 полносвязный.



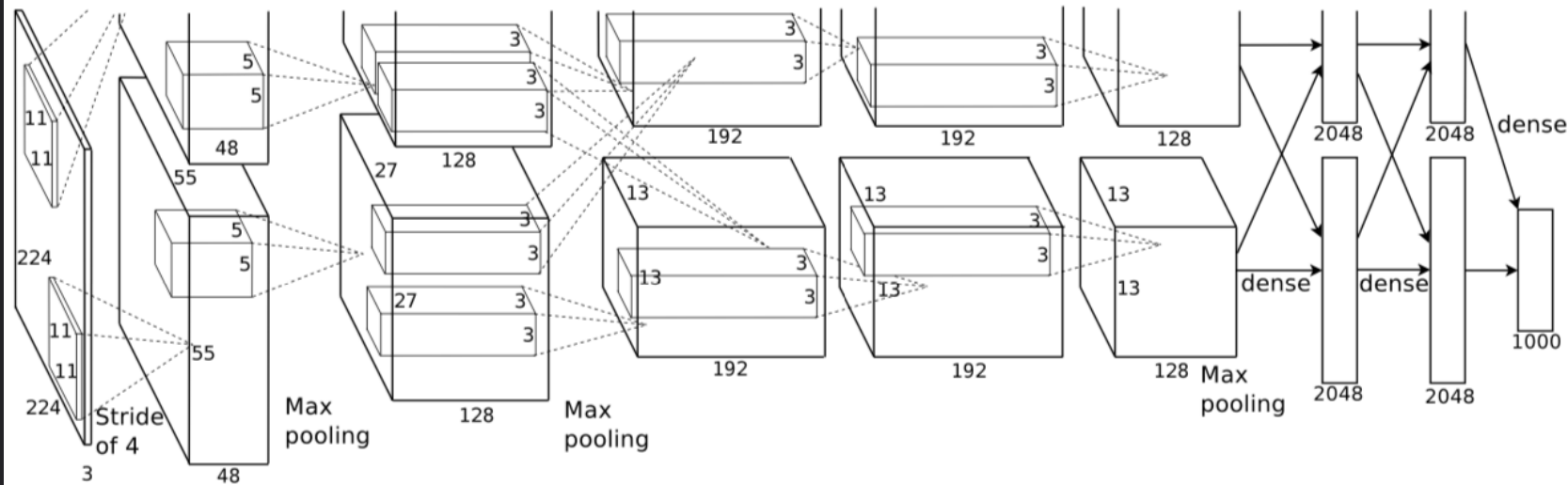
## AlexNet

- Создатель – Alex Krizhevsky. [\[3\]](#)
- Новый рекорд ImageNet.
  - Лучше предыдущего результата на 10,8 процентных пункта.
- Custom реализация – разделение на 2 видеокарты.
  - Видеопамяти одной видеокарты в 2012 году не хватало для хранения модели целиком.



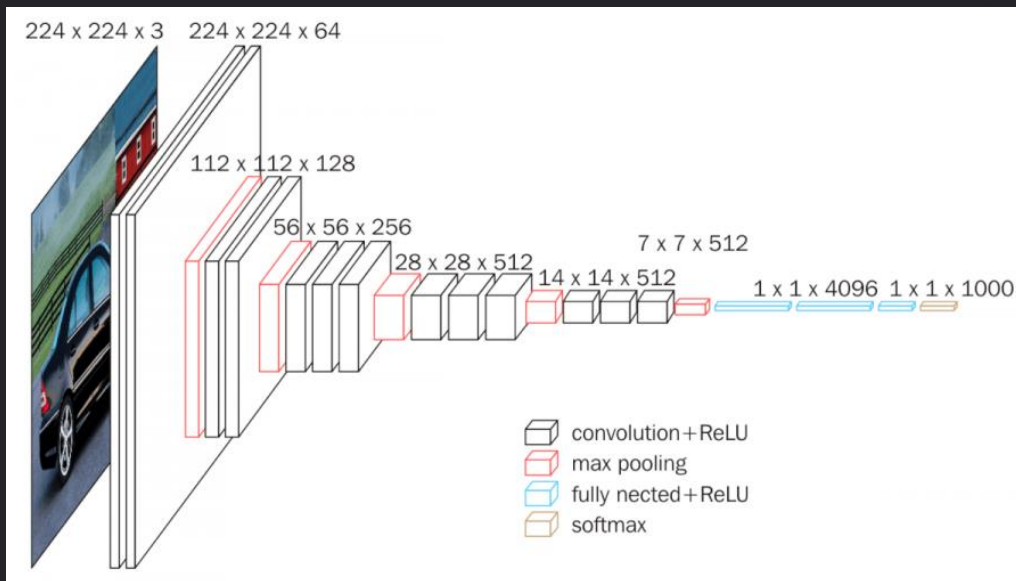
## AlexNet. Архитектура.

- 5 свёрточных слоёв.
- 3 слоя max pooling.
- 2 полносвязных слоя.
  - Dropout в полносвязных слоях.
- Впервые использована функция активации ReLU.



## VGGNet

- Используются свёртки  $1 \times 1$  после max pooling слоёв, для компенсации потерянных признаков. [5]
- Используются свёрточные слои с окнами  $3 \times 3$  – менее затратно по вычислениям.





## ResNet

- Бесконечно наращивать количество слоёв нельзя.
- Больше коэффициентов – больше возможность переобучения.
  - Но даже не это самая большая проблема.
- Возникает проблема взрыва / исчезновения градиента. [\[4\]](#)
  - С самого начала отсутствует сходимость.
- Слишком много свёрток слишком обобщают признаки, теряется исходная информация.
- Идея: после свёртки на следующем слое после взвешивания перед активацией прибавлять входной сигнал.
  - У выхода такой подсети информации не меньше, чем на входе.

## Пути развития

Начинает проявляться проблема исчезновения градиента

LeNet

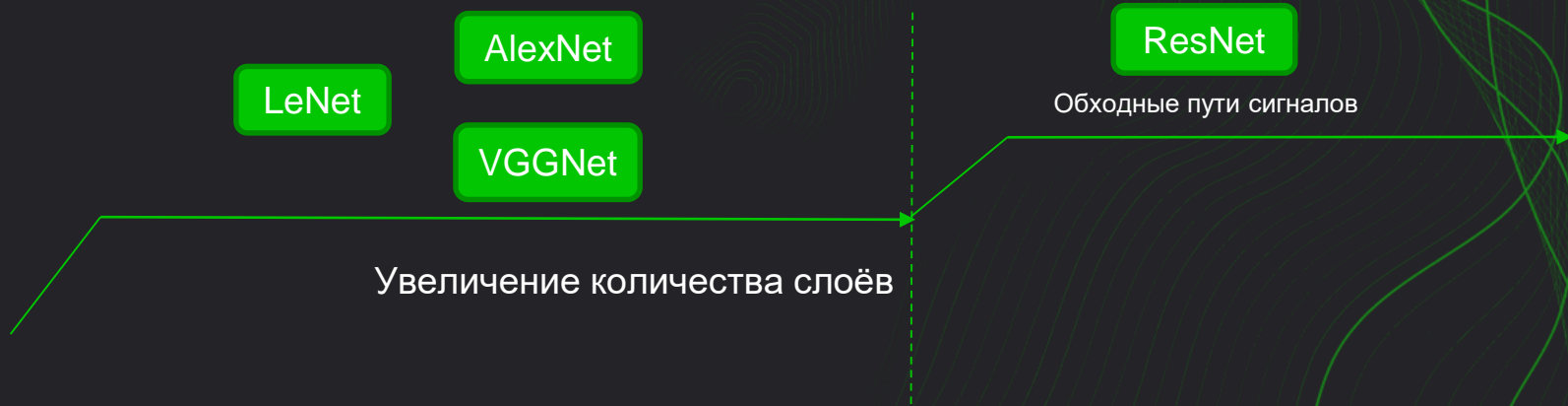
AlexNet

VGGNet

ResNet

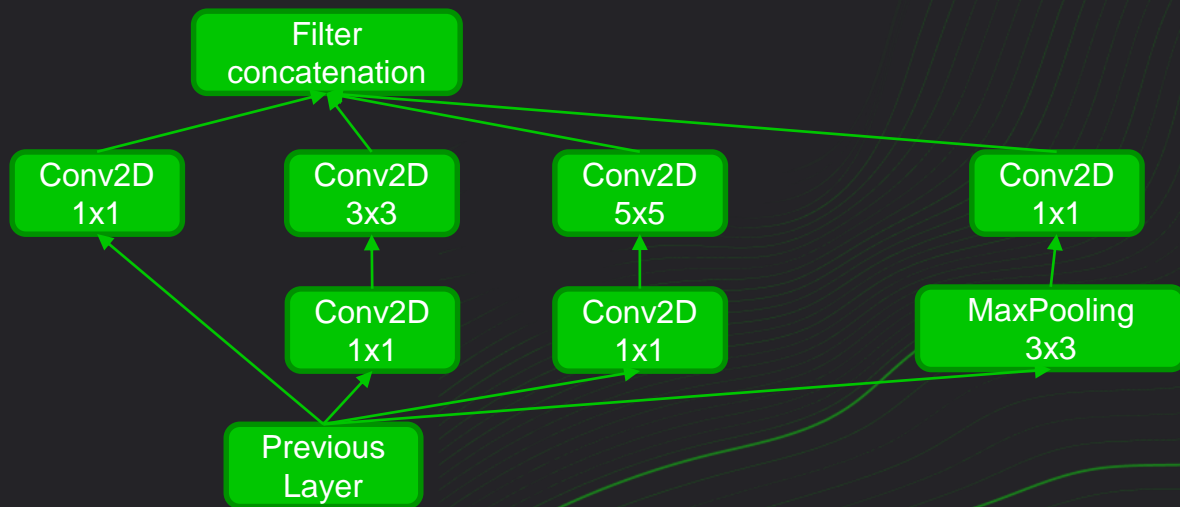
Обходные пути сигналов

Увеличение количества слоёв



## InceptionNet. Основная идея

- Увеличивать вглубь нельзя. [6]
- А что, если попробовать увеличить нейронную сеть вширь?
  - Может ли это дать какие-нибудь преимущества?
- Нет идеального универсального размера окна свёртки.
  - Один и тот же объект может быть чуть больше, может быть чуть меньше.
- Сеть состоит из нескольких Inception блоков:



## Пути развития

Начинает проявляться проблема исчезновения градиента

LeNet

AlexNet

VGGNet

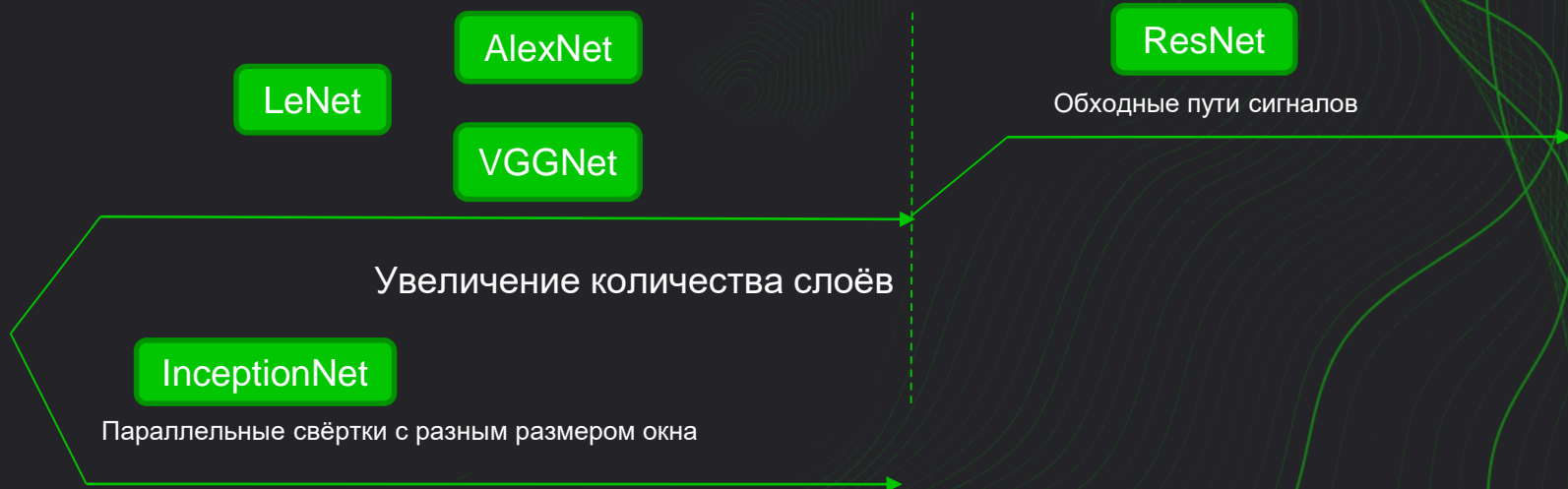
ResNet

Обходные пути сигналов

Увеличение количества слоёв

InceptionNet

Параллельные свёртки с разным размером окна



## XceptionNet

- Xtreme Inception. [\[7\]](#)
- Использует похожую на Inception архитектуру.
- Основан на более сильной гипотезе:
  - Кросс-канальную корреляцию и корреляцию областей можно вычислять независимо без потери признаков, вычисляемой свёрткой.

## Пути развития

Начинает проявляться проблема исчезновения градиента

LeNet

AlexNet

VGGNet

ResNet

Обходные пути сигналов

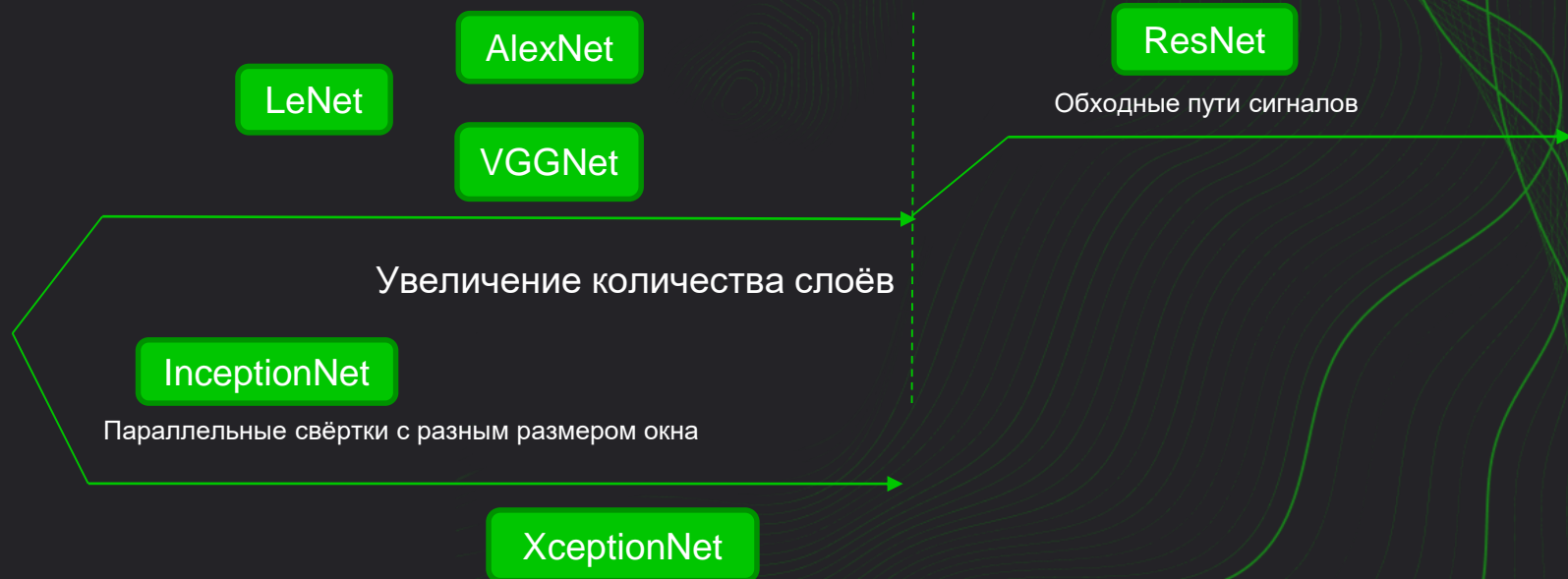
Увеличение количества слоёв

InceptionNet

Параллельные свёртки с разным размером окна

XceptionNet

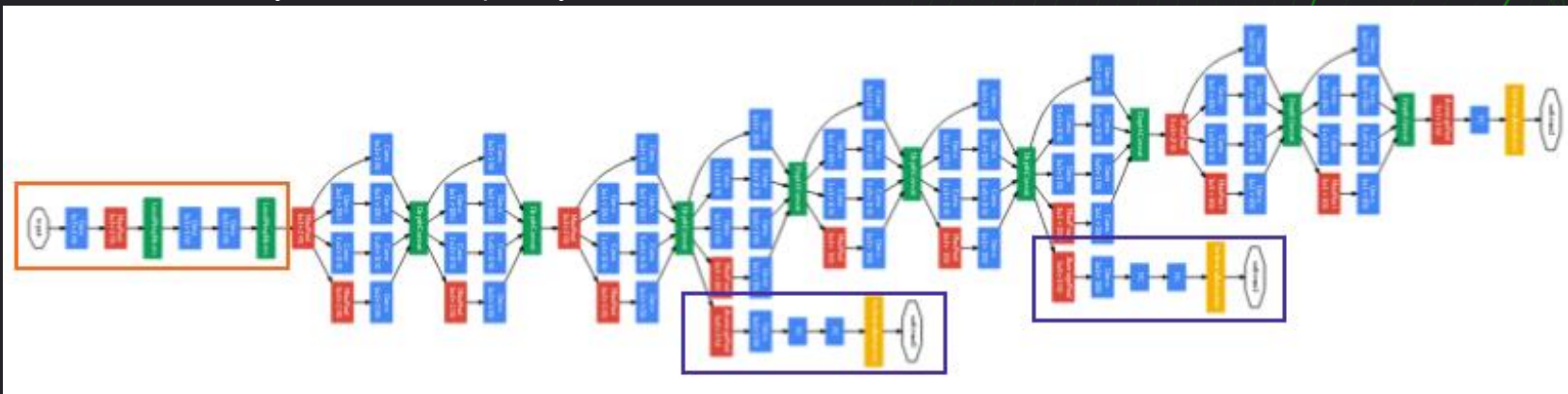
Кросс-канальная корреляция отдельно от свёртки области





## GoogLeNet

- 9 слоёв Inception блоков + max pooling. [6]
- 22 свёрточных слоя.
- 27 слоёв всего, с учетом max pooling слоёв.
- Подвержена проблеме исчезновения градиента.
- Трюк:
  - 2 вспомогательных классификатора.
  - Их loss с небольшим весом прибавляется к основному loss.
  - Используются только при обучении.



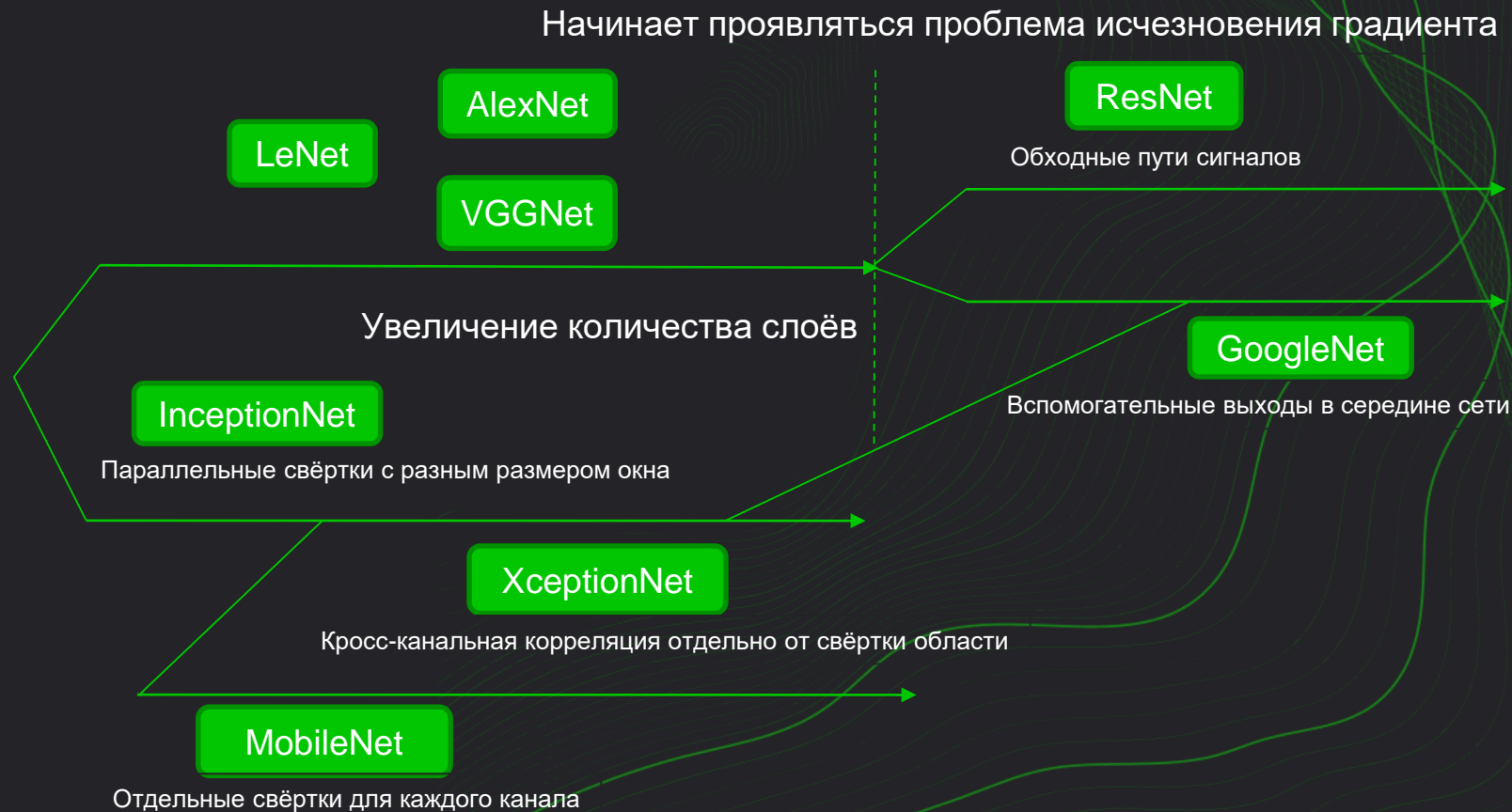
## Пути развития



## MobileNet

- Адаптация InceptionNet для мобильных устройств.
- Использует свёртки с отдельными фильтрами для каждого канала.
- Благодаря этому сильно уменьшается количество параметров сети.
- С незначительным ухудшением точности.

## Пути развития



## NASNet

- Новый рекорд ImageNet.
- Neural Architecture Search.
- Reinforcement learning для поиска гиперпараметров. [\[8\]](#)
- Требуется очень больших вычислительных мощностей.
  - 500 GPU
  - 4 дня

## Пути развития

NASNet

Обучение с подкреплением для поиска гиперпараметров

Начинает проявляться проблема исчезновения градиента

LeNet

AlexNet

VGGNet

ResNet

Обходные пути сигналов

Увеличение количества слоёв

InceptionNet

Параллельные свёртки с разным размером окна

GoogleNet

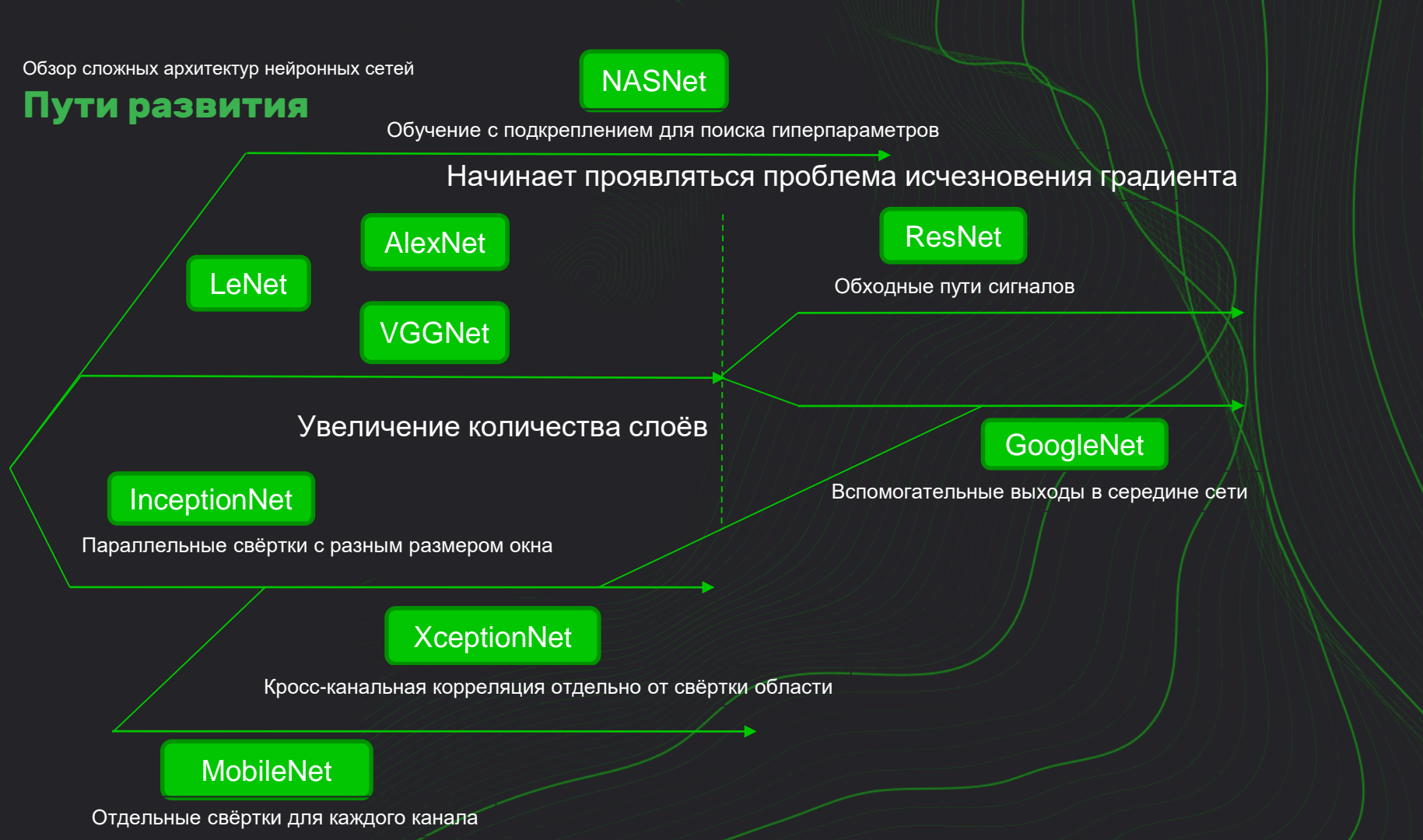
Вспомогательные выходы в середине сети

XceptionNet

Кросс-канальная корреляция отдельно от свёртки области

MobileNet

Отдельные свёртки для каждого канала





## Ссылки

1. ImageNet (<https://www.image-net.org/>)
2. Gradient-Based Learning Applied to Document Recognition (<http://yann.lecun.com/exdb/publis/pdf/lecun-98.pdf>)
3. ImageNet Classification with Deep Convolutional Neural Networks  
(<https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>)
4. Deep Residual Learning for Image Recognition (<https://arxiv.org/pdf/1512.03385.pdf>)
5. VGG16 – Convolutional Network for Classification and Detection (<https://neurohive.io/en/popular-networks/vgg16/>)
6. Going Deeper with Convolutions  
(<https://static.googleusercontent.com/media/research.google.com/ru//pubs/archive/43022.pdf>)
7. Xception: Deep Learning with Depthwise Separable Convolutions  
([https://openaccess.thecvf.com/content\\_cvpr\\_2017/papers/Chollet\\_Xception\\_Deep\\_Learning\\_CVPR\\_2017\\_paper.pdf](https://openaccess.thecvf.com/content_cvpr_2017/papers/Chollet_Xception_Deep_Learning_CVPR_2017_paper.pdf))
8. NASNet – A brief overview (<https://iq.opengenus.org/nasnet/>)

## Резюме

- Рассмотрена проблема доступности данных и подготовки датасетов.
- Рассмотрены различные сложные архитектуры и их подходы:
  - LeNet
  - AlexNet
  - VGG
  - ResNet
  - InceptionNet
  - XceptionNet
  - GoogLeNet
  - MobileNet
  - NASNet

## Вопросы для самоконтроля

- В чём заключается проблема доступности данных и подготовки датасетов?
- Какие сложные архитектуры вы можете назвать?
- Какие проблемы ограничивают сложность архитектур глубоких нейронных сетей?
- Опишите эти проблемы.

# Спасибо!

