

# A ConvNet for the 2020s

Review

Пилипенко Сергей  
Кульпин Павел  
Петров Олег

# Контекст

- Появление ViT снизило популярность сверточных сетей.
- Непонятно, почему трансформеры так хороши.
- Авторы хотят воскресить ConvNet, показав их эффективность.
- Постепенно модифицируют ResNext фишками из Swin Transformer.

# О статье

- Опубликовано: 01.2022
- Конференция: CVPR 06.2022
- Facebook AI Research
- Авторы: Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie

# Базовые статьи

- **[ResNext]** Aggregated Residual Transformations for Deep Neural Networks. CVPR, 2017
- **[Swin Transformer]** Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. 2021
- **[DeiT Training]** Training data-efficient image transformers & distillation through attention. 2020

# Цитирования и продолжения

- 297 цитирований за 3 месяца.
- Scaling Up Your Kernels to 31x31: Revisiting Large Kernel Design in CNNs. 2022
- More ConvNets in the 2020s: Scaling up Kernels Beyond 51x51 using Sparsity. 2022

# Результаты и сильные стороны

- Достижение/превосхождения результатов ViT для разных задач CV
- Достижение поставленной цели.
- Закрытие большого пробела в сфере CV, так как никто до них не пытался этого делать.
- Рассеивает мифы о трансформерах.
- Хорошо рассказывает про обучения сверточных сетей.

# Минусы и улучшения

- Ничего нового не придумали.
- По сути их успех - некоторая случайность.
- Строго однозначная последовательность экспериментов.