

ML Специализация

CNN для классификации
изображений



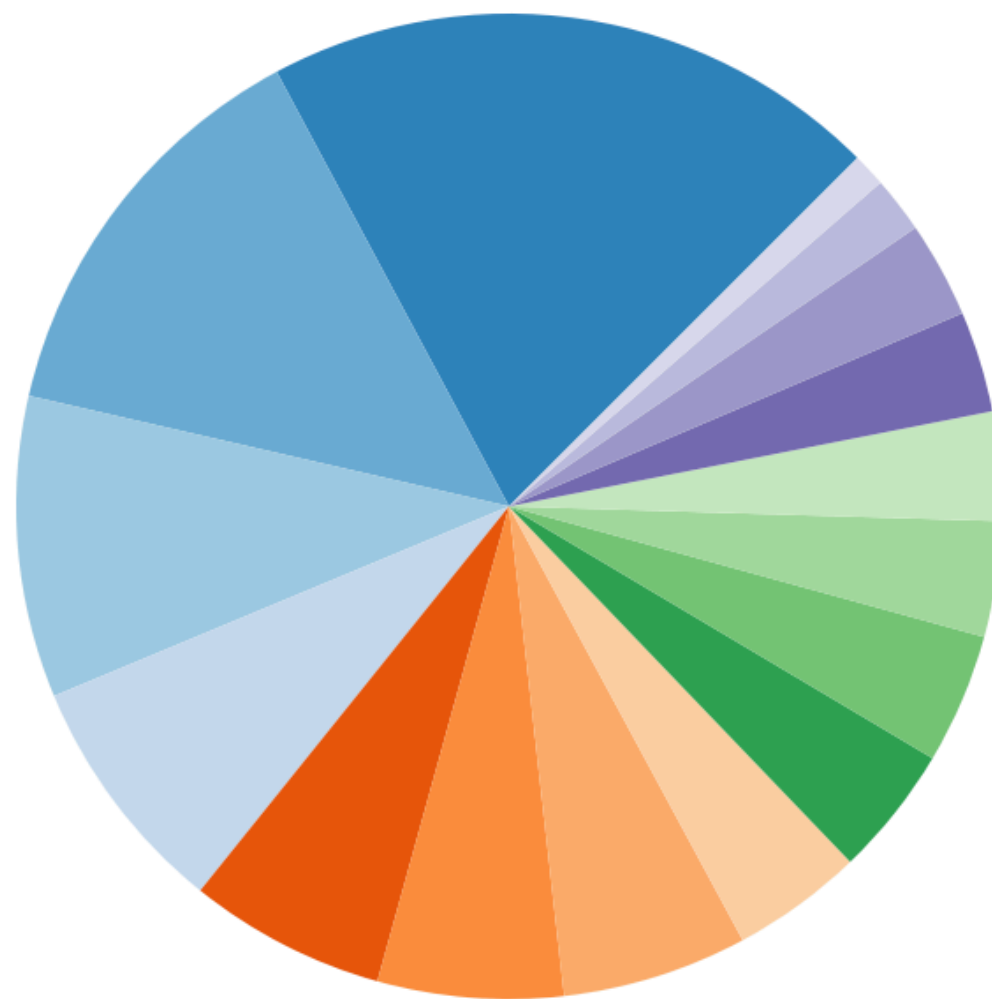
команда 21 Deep Seers

Q 4 2024

Задача

Классифицировать вид деятельности по изображению (20 категорий)

Распределение категорий



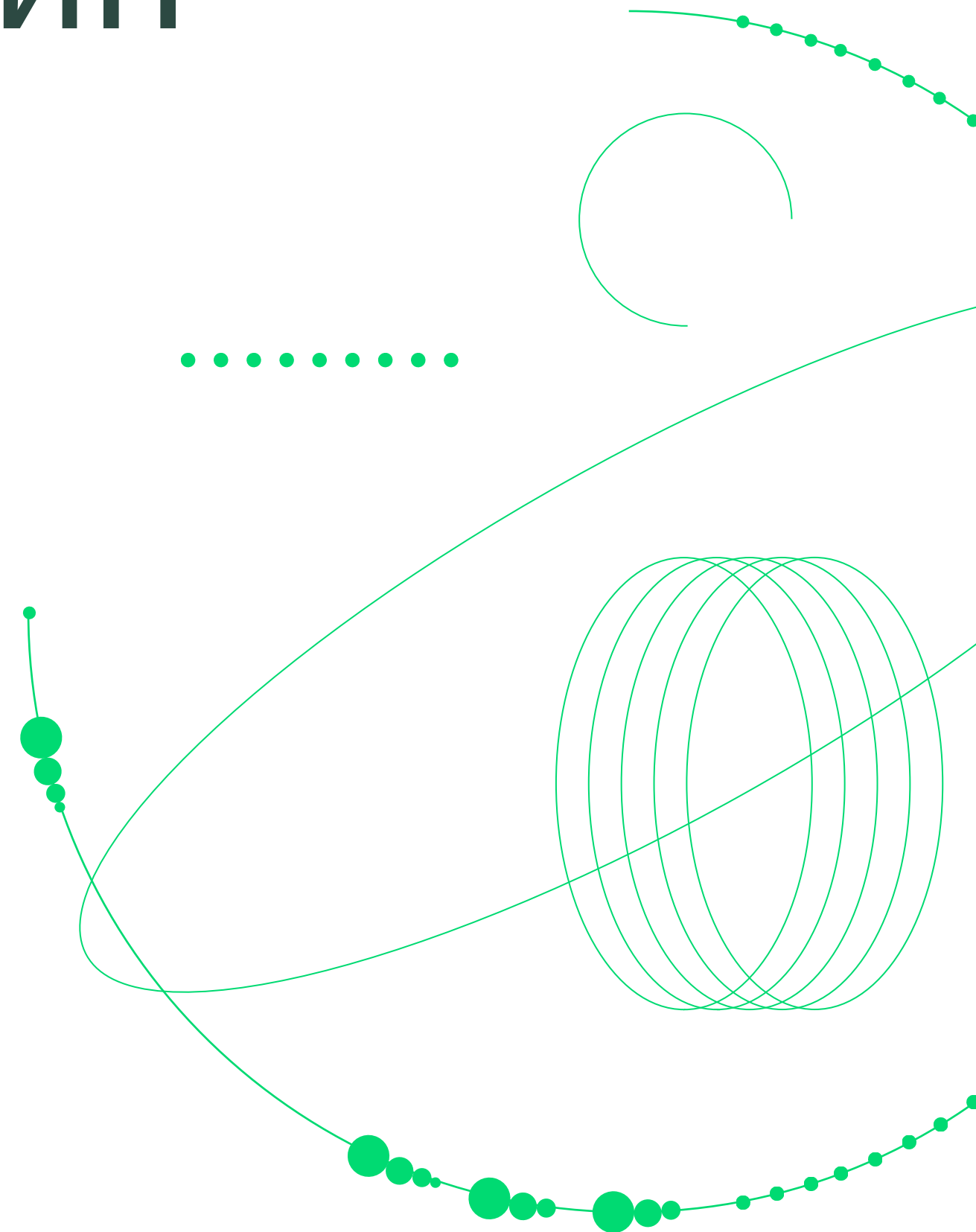
Категории	
■	sports
■	occupation
■	conditioning exercise
■	home activities
■	lawn and garden
■	home repair
■	water activities
■	winter activities
■	miscellaneous
■	fishing and hunting
■	dancing
■	walking
■	music playing
■	bicycling
■	running
■	inactivity quiet/light



Пару слов про бейзлайн

Один из наиболее влияющих на продуктивность на дистанции пунктов - складно написанный бейзлайн

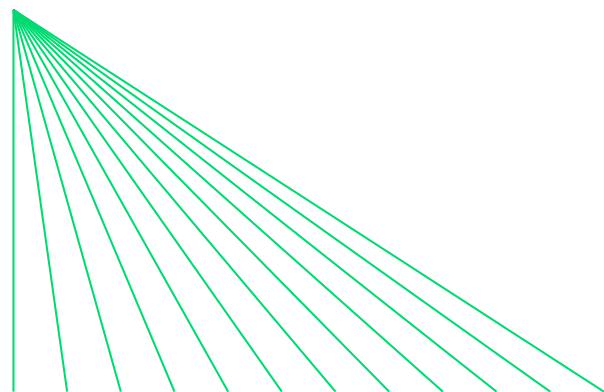
- 01 Удобная системой чекпоинтов
- 02 Автоматизированный процесс обучения
- 03 Структурированный код
- 04 Визуализированные итоги обучения



Подход к решению

Аугментация данных

Провести исследование методов аугментации данных, таких как повороты, отражения, обрезка и изменения яркости, для повышения устойчивости модели.



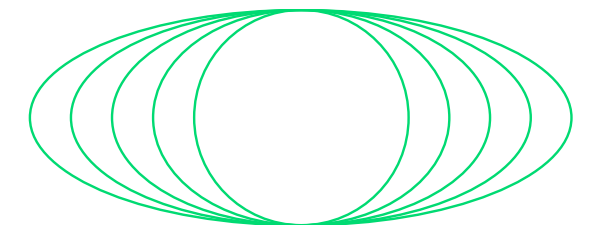
Тестирование различных архитектур

Провести исследование различных архитектур моделей, включая базовые CNN и более сложные сети, чтобы определить наиболее подходящую структуру для задачи.

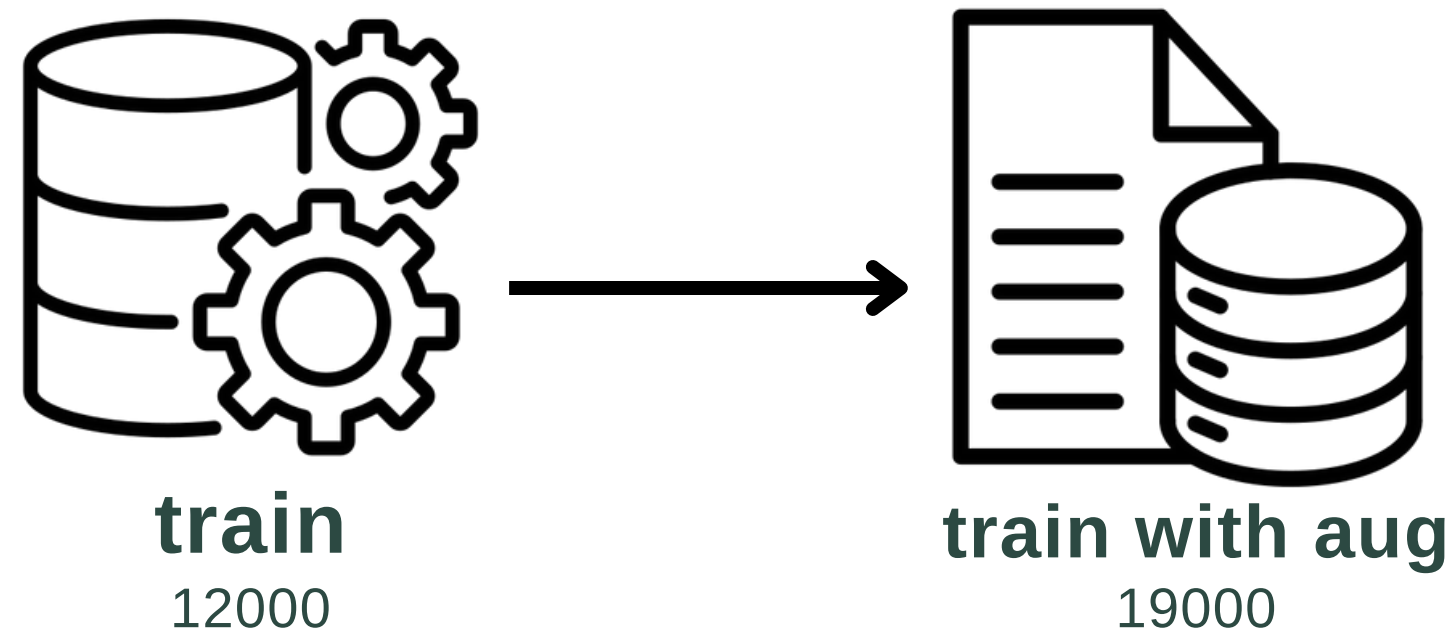


Оптимизация обучения

Провести исследование подходов к оптимизации обучения, включая настройку гиперпараметров, использование различных оптимизаторов и внедрение методов регуляризации.

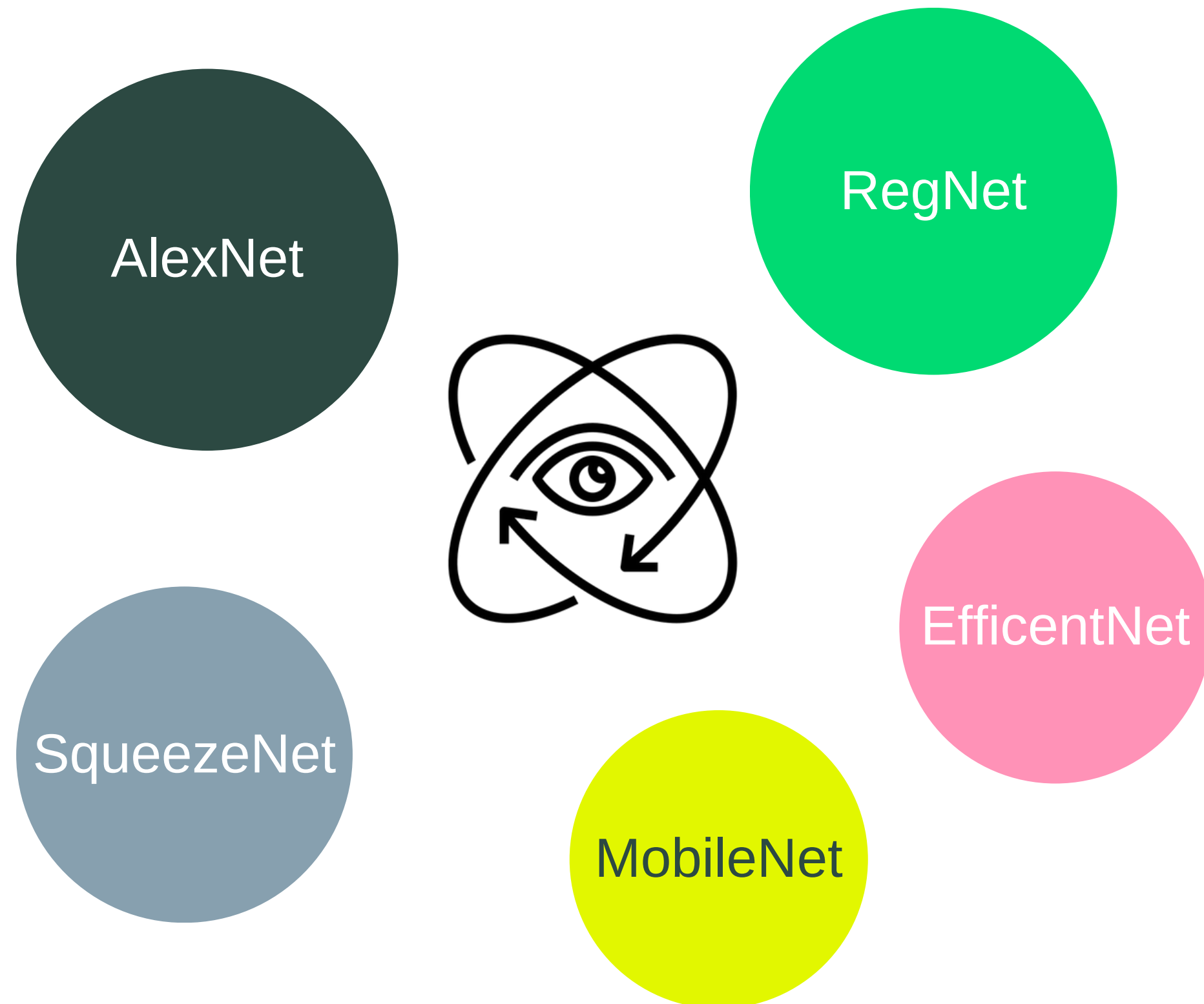


Аугментация данных

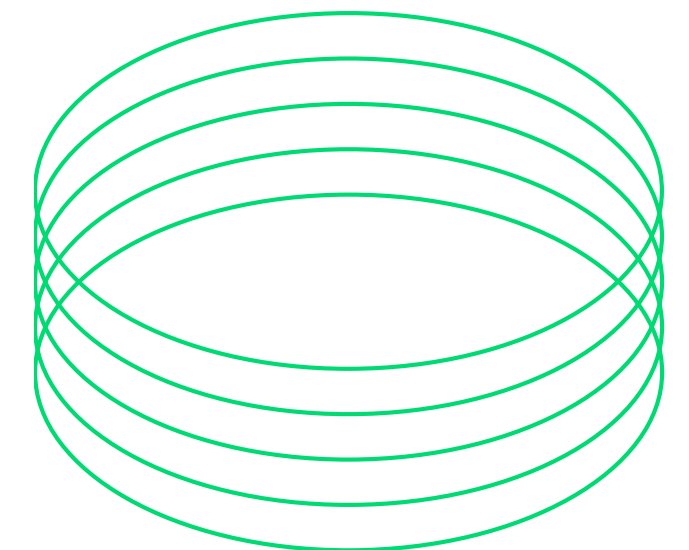


Аугментация улучшает обобщающие способности модели, увеличивая разнообразие данных. Для этого применены трансформации, такие как изменение размера, повороты, изменения цвета и добавление шума.

Архитектуры:



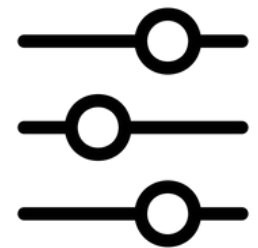
Модели с меньшим количеством параметров (MobileNet, SqueezeNet) показали лучшие результаты на небольшом объёме данных. EfficientNet обошёл всех по ключевым метрикам благодаря оптимальной архитектуре.



Оптимизация обучения



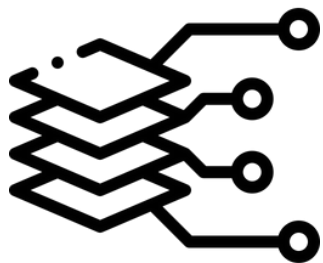
Loss функции



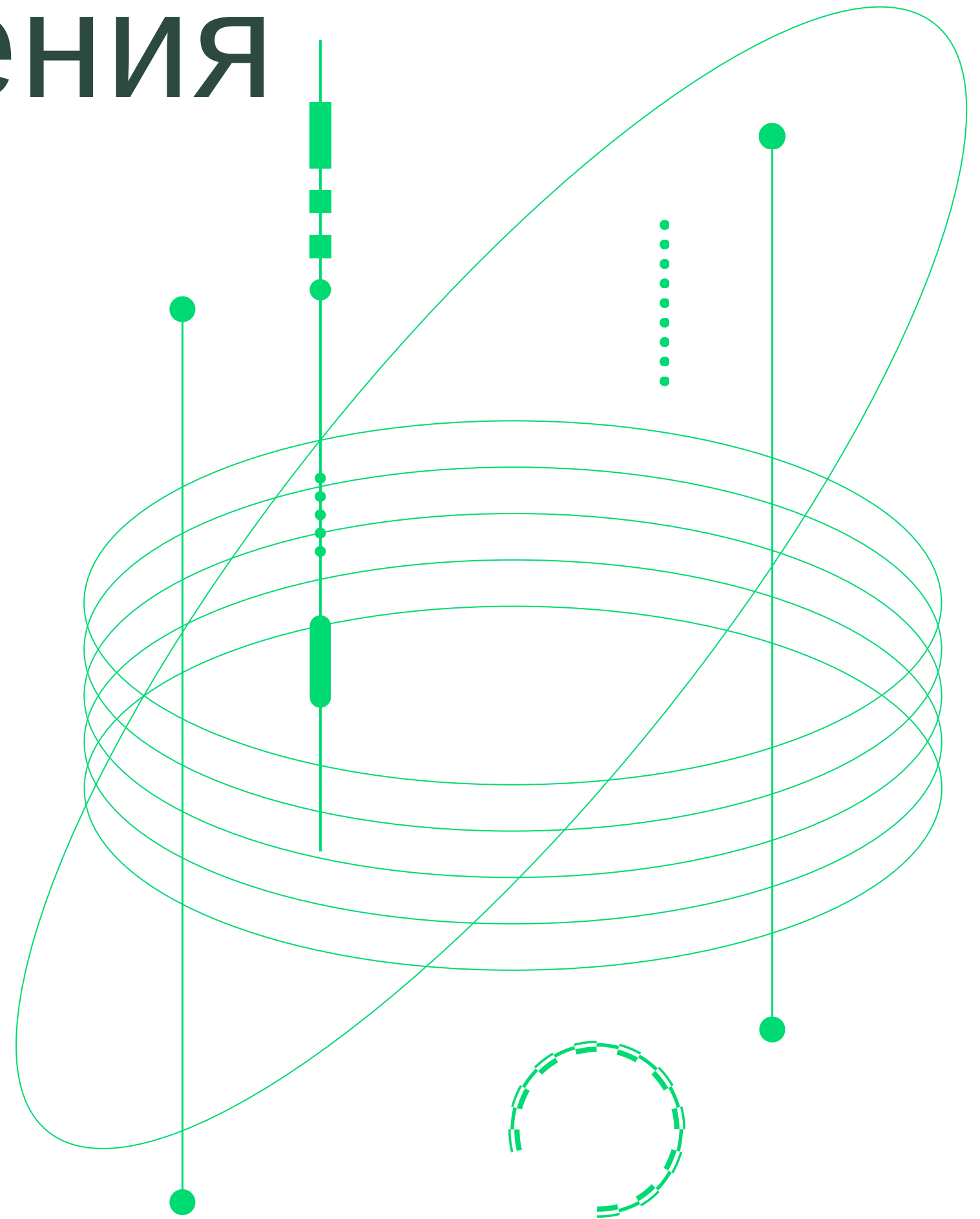
Гиперпараметры



Оптимизаторы



Архитектуры моделей



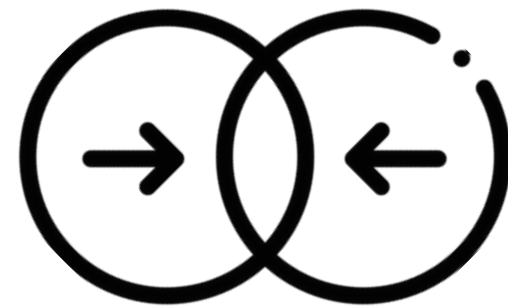
Итоговое решение

Ансамбль

18 и 40 эпохи

Модель на основе ResNet18

- 11,273,812 параметров
- Mish активация
- Group Normalization
- SEBlock
- Dropout
- Свой ResidualBlock

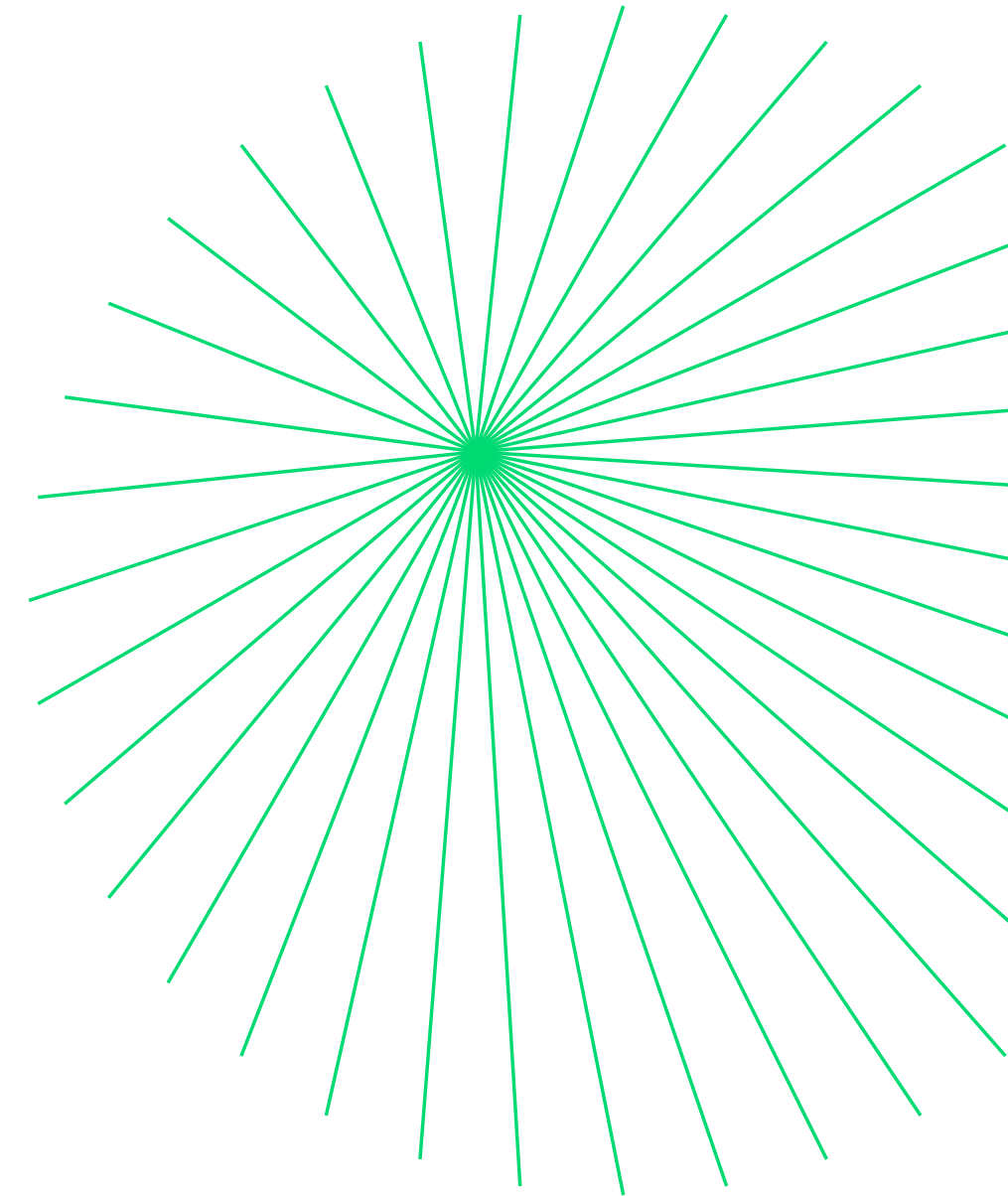


40, 142 и 251 эпохи

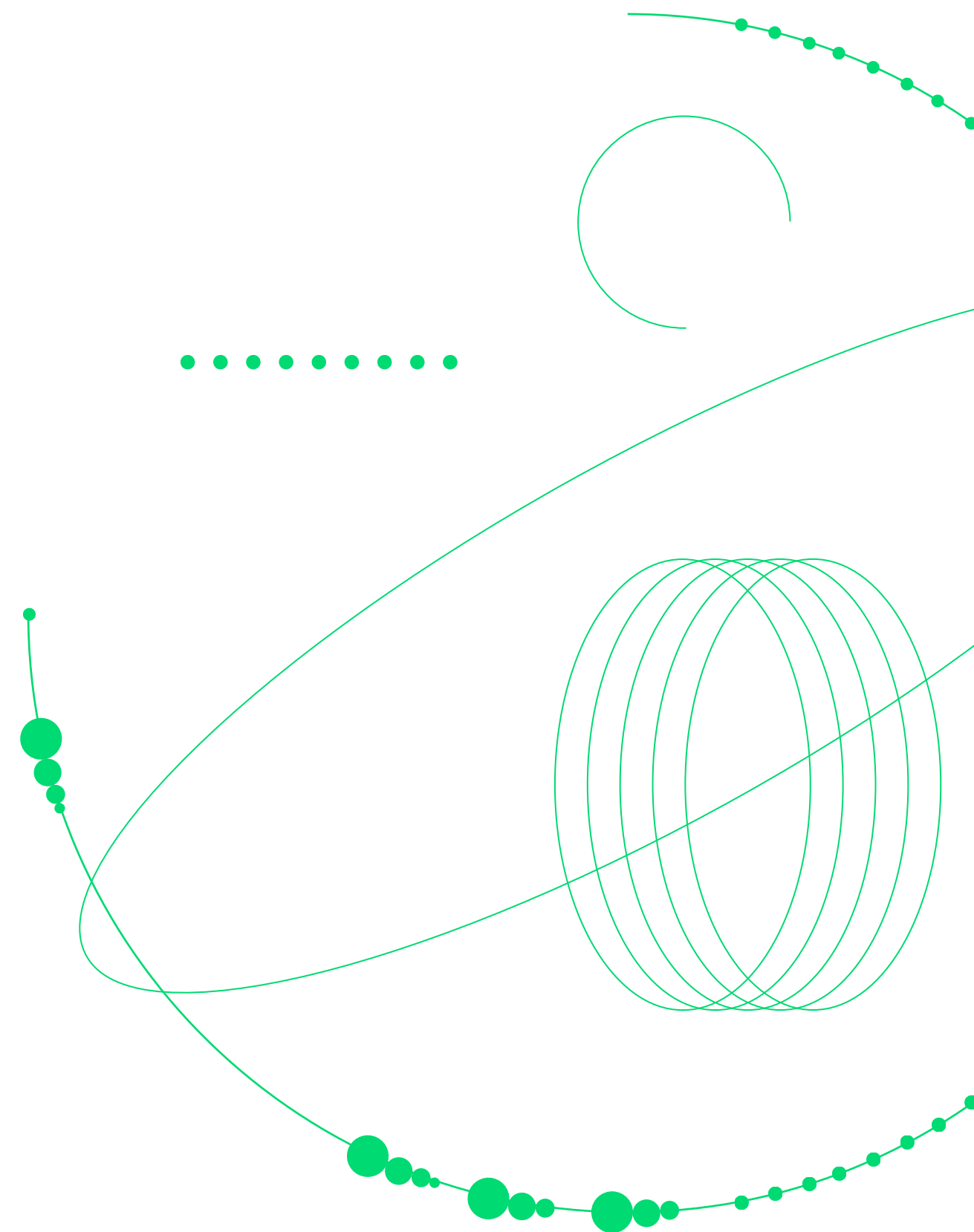
Модель на основе EfficientNet-B0

- 4,074,512 параметров
- Эффективная свертка
- Identity mapping
- Dropout

F1-score: 0.831



Почему это работает?



Наша команда



Языкин Артем
@I_D_DQD

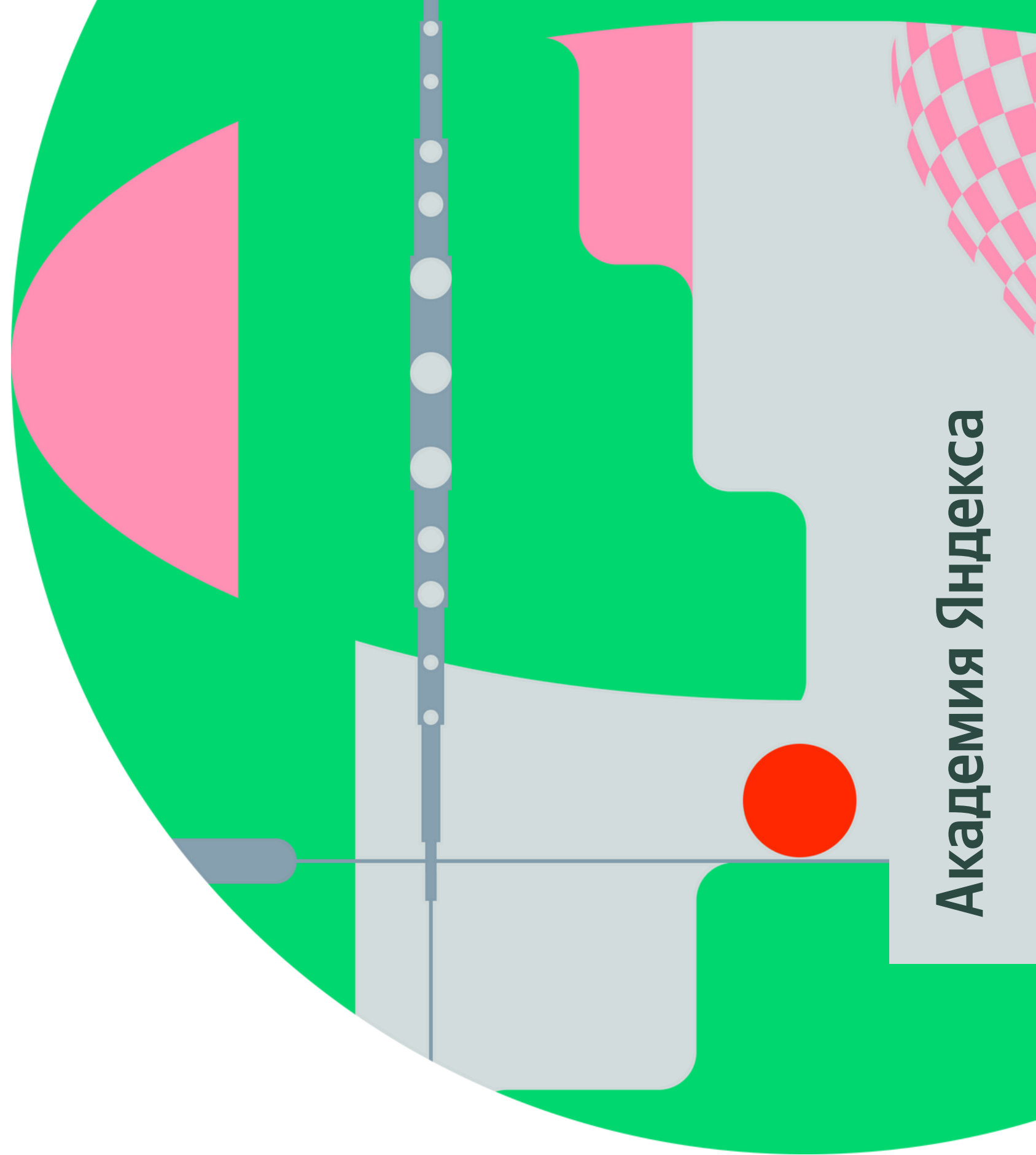


Владислав Сунгуров
@VladG00dman



Пудовкин Ярослав
@polkopolk

Спасибо за
внимание!



Академия Яндекс