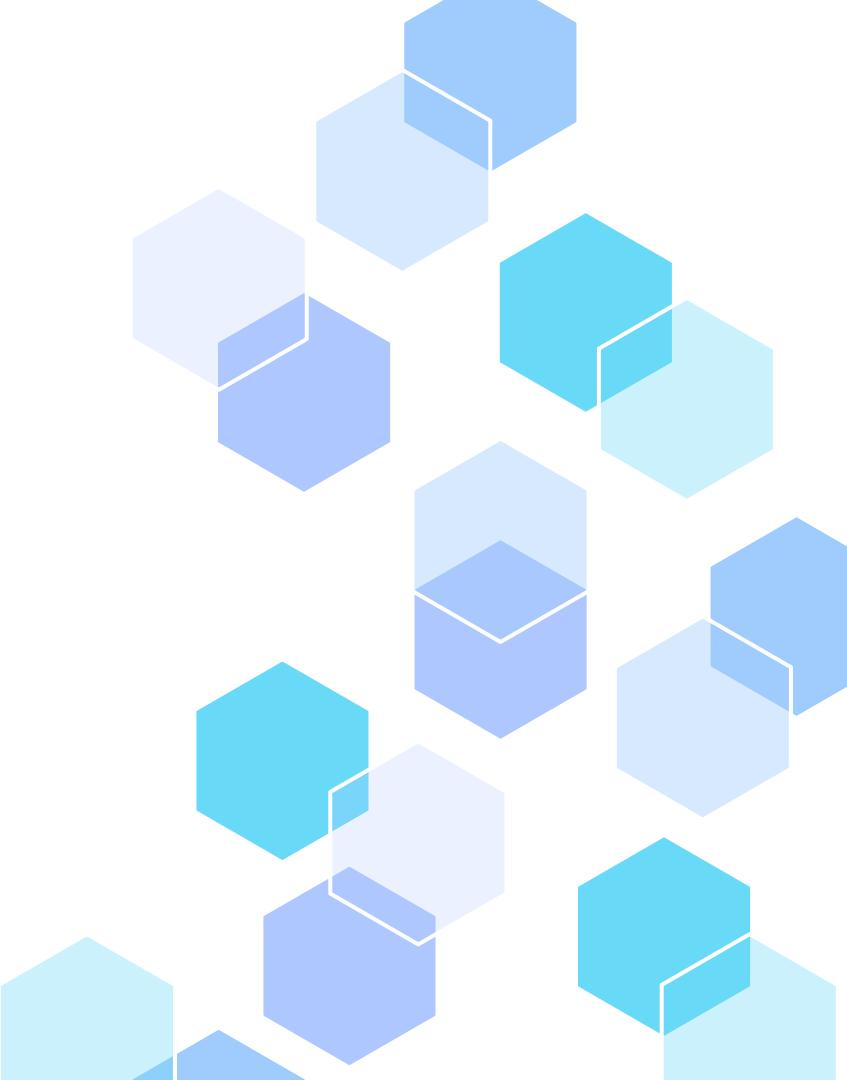


# KotVasily



# О чём будет сегодня говорить?

01

Суть задачи

02

Данные

03

Ключевые идеи

04

Пайплайн

05

Примеры  
предсказаний

06

Анализ ошибок  
и предсказаний



# Суть задачи

Предсказать AWB(Auto White Balance) на изображении.

# Данные

- 570 изображений RAW в формате PNG
- Метаданные (Exposure Time, ISO, LightValue)
- Гистограммы
- Точки белого

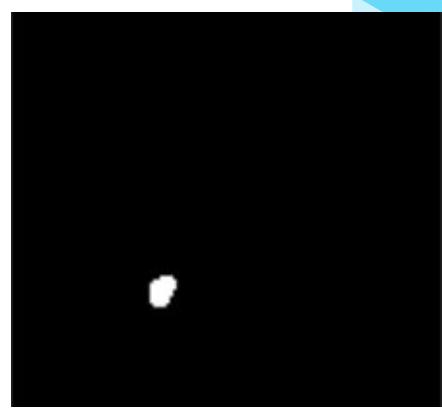


# Метаданные

Колонка	Польза для модели
Exposure Time	Помогает модели больше узнать об условиях съемки.
ISO	Чем больше ISO тем больше шума на изображении.
LightValue	Уровень яркости на изображении.

# Гистограммы

Для каждой фотографии мы имеет 2D-гистограмму, показывающую, какие точки белого физически возможны для данного изображения(сцены).



# Ключевые идеи.

Какой способ предсказания  
точки белого лучший?

Какие архитектурные решения  
могут помочь?

# Какой способ предсказания точки белого лучший?

01

Регрессия точки белого

- Анализирует изображения целиком
- Тяжело понять на основе чего модель приняла решения
- Сложная задача

02

Предсказывать RAW с  
AWB и считать точку  
белого

- Сложная задача
- Высокая нестабильность

03

Unet подобная  
архитектура + Grey World.

- Высокая стабильность
- Простая задача для модели
- Наивысшие показатели метрики на CV(cross-validation)

---

# Какие архитектурные решения работают и почему?



# 1. Batch Normalization vs Instance Normalization

## Batch Normalization

- ✗ Более стабильное обучения только если мало выбросов
- ✗ Сильная зависимость от батча
- ✗ Информация о цвете привязана к батчу

## Instance Normalization

- ✓ Удаляет цветовой сдвиг.
- ✓ Низкая зависимость от батча
- ✓ Информация о цвете для каждого изображения

## 2. Использования 1x1 сверток



**Взаимосвязь каналов -**  
помогает создавать  
сложные цветовые  
комбинации



**Низкая вычислительная  
сложность -** быстро и  
эффективно

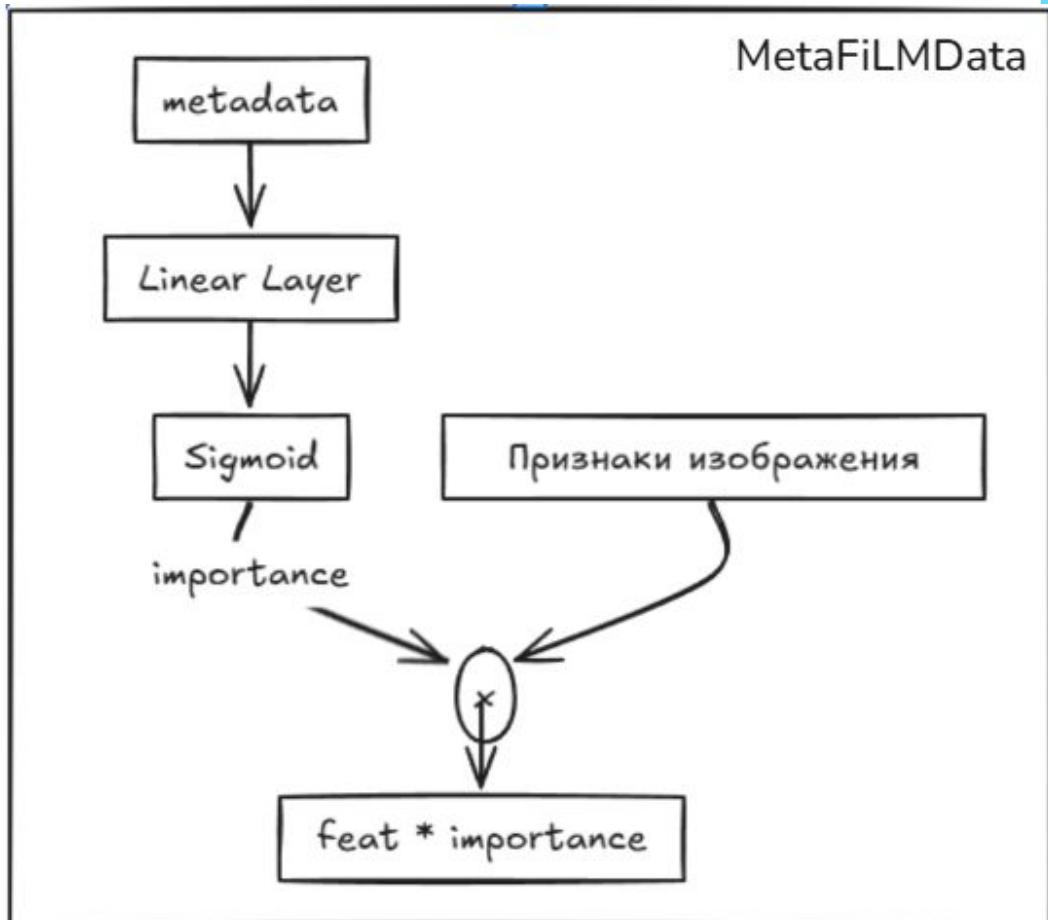
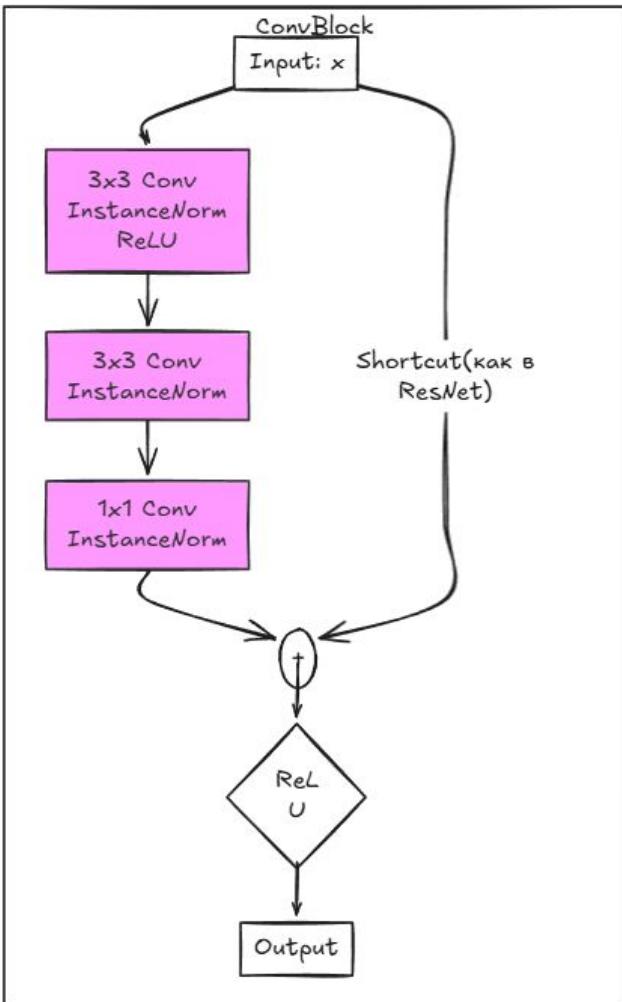
# 3. FiLM vs Моя реализация

Идея	Описания	Полезность для AWB
FiLM	$h' = \gamma \cdot h + \beta$ $\gamma, \beta$ - предсказываем на основе metadata	Нет
Моя	3(metadata) $\rightarrow$ 512 $\rightarrow$ sigmoid $\rightarrow$ множим на каналы	метаданные влияют на признаки по логике

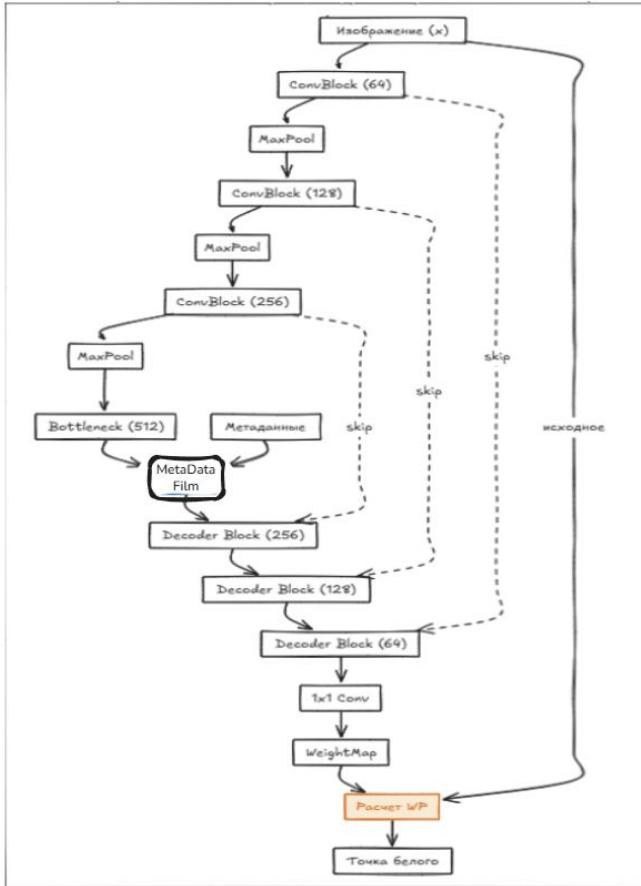


---

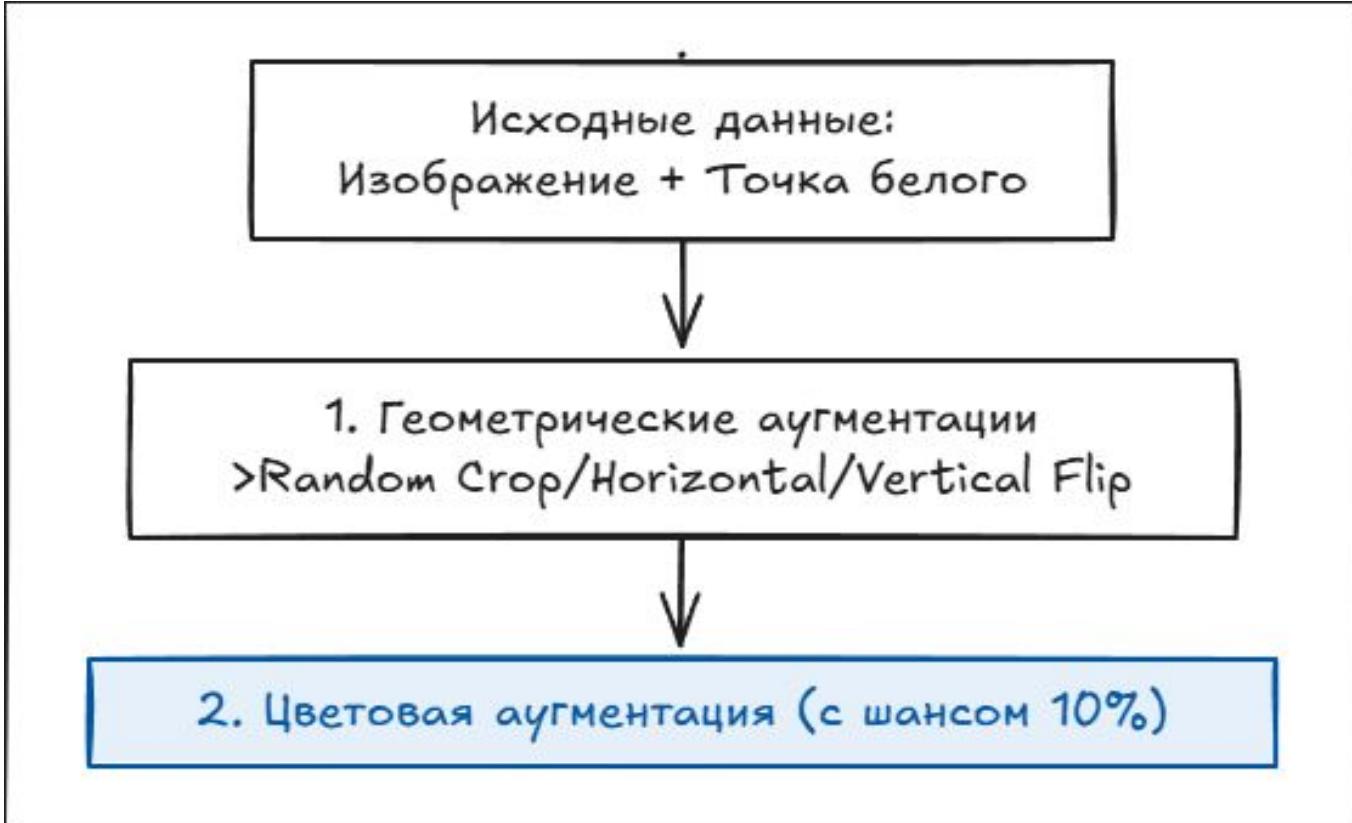
# Пайпайн



# Модель (AWB Model)



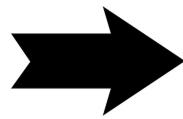
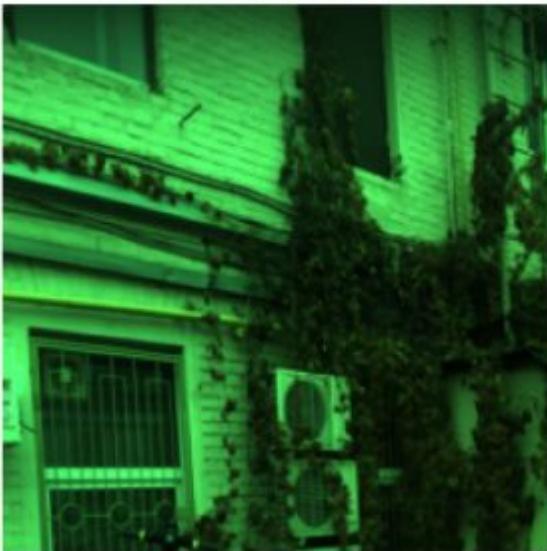
# Аугментации



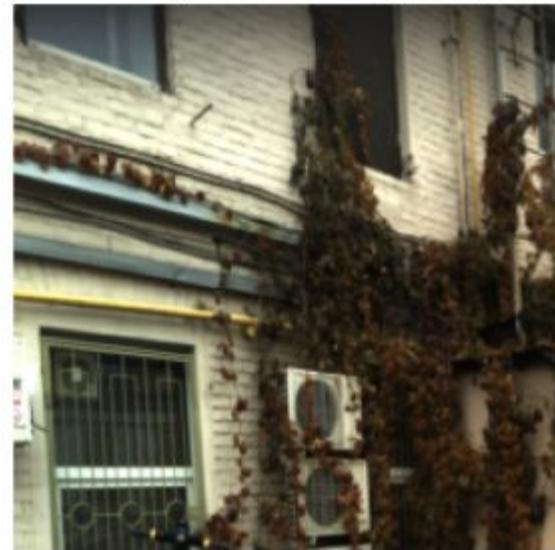
# Примеры предсказаний.



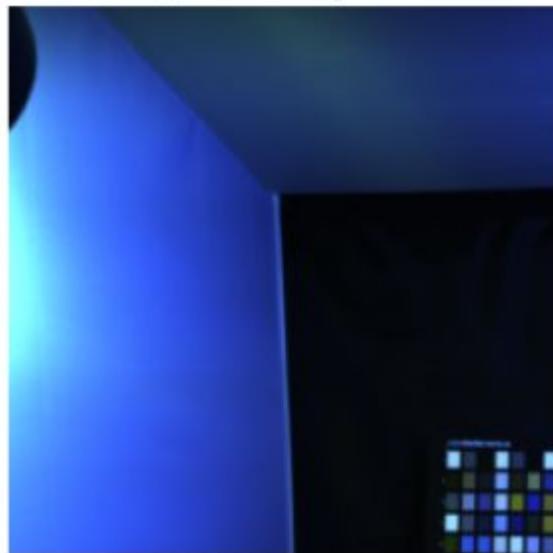
Исходное изображение



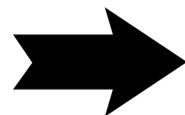
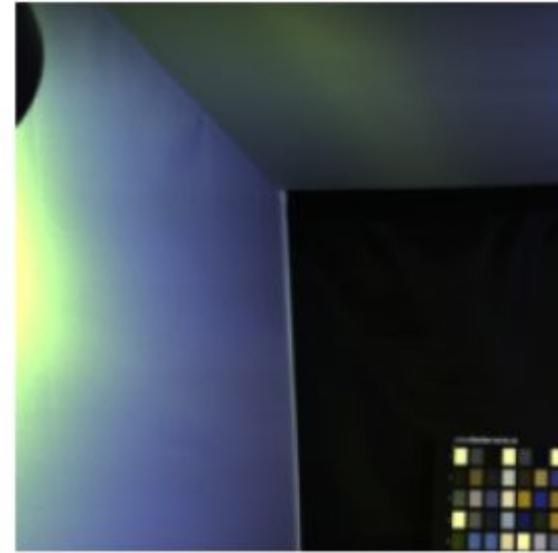
После цветовой коррекции



Исходное изображение



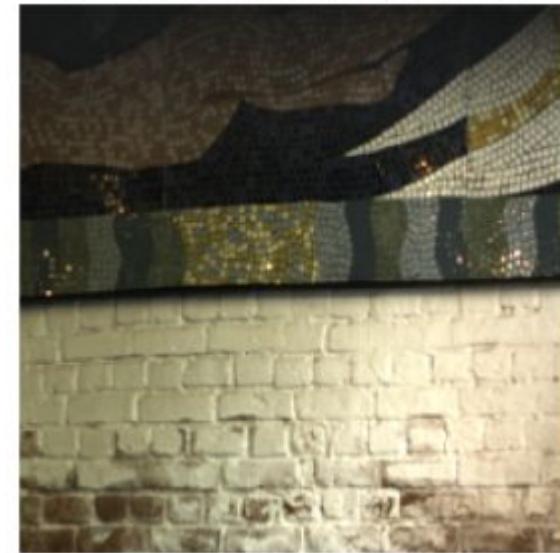
После цветовой коррекции



Исходное изображение



После цветовой коррекции



Исходное изображение



После цветовой коррекции





---

# **Анализ ошибок и предсказаний**

Исходное изображение



Важность пикселей R



Важность пикселей G



Важность пикселей B



После цветовой коррекции



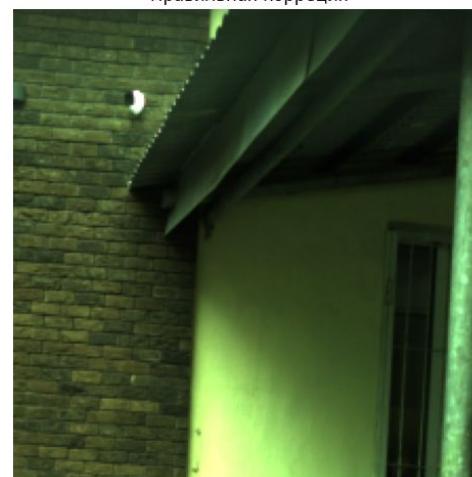
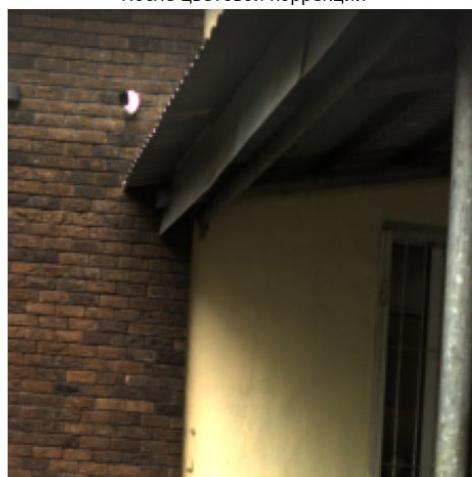
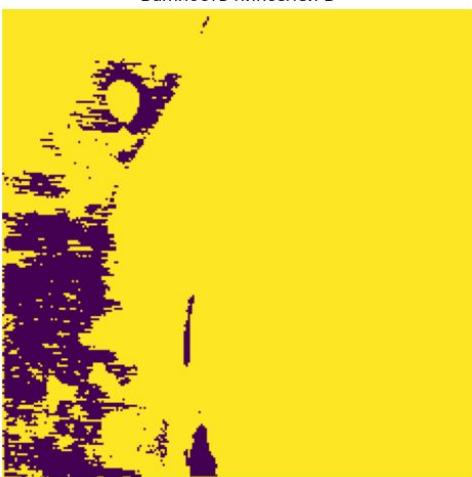
Правильная коррекция





---

# **Анализ ошибок**



Исходное изображение



Важность пикселей R



Важность пикселей G



Важность пикселей B



После цветовой коррекции



Правильная коррекция



# Анализ ошибок:

1 - на первом изображении видно что правильная разметка немного зеленовато.

**ГИПОТЕЗА: неправильная разметка**

2 - на втором изображении, мы видим одну из проблем модели иногда свет получается фиолетовым.

**ПРОБЕЛМА: модель делает слишком большой wp\_g из-за этого зеленый подавляется и кажется изображения фиолетовым.**

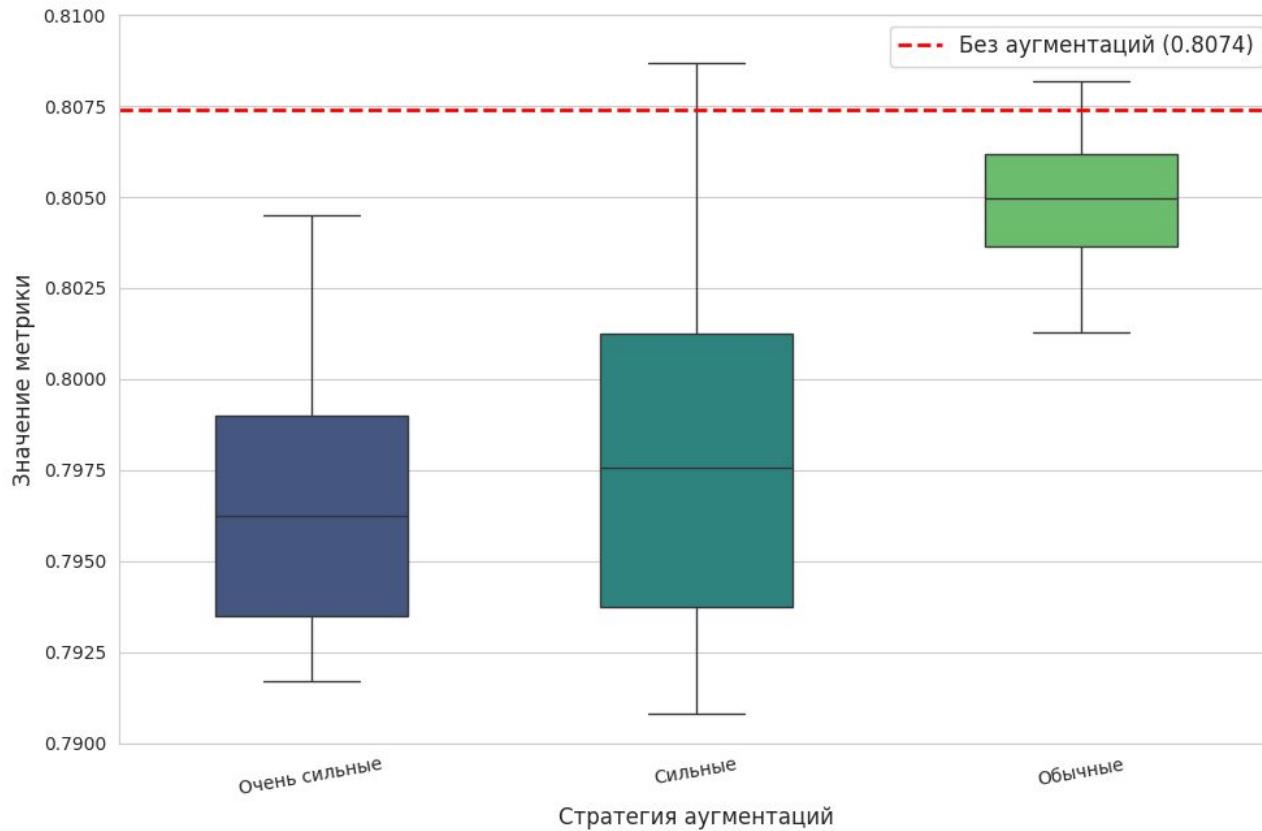
**ГИПОТЕЗА: В данных очень много зеленых изображений и мало где доминируют другие цвета.**

# Результаты

Fold 1	0.8074
Fold 2	0.8316
Fold 3	0.8200
Fold 4	0.8032
Fold 5	0.8063
Mean average folds	0.8137

# Оценка стабильности модели

Влияние силы аугментаций на итоговую метрику



---

# Вопросы?

Материалы которые использовались:

- FC4: Fully Convolutional Color Constancy with Confidence-weighted Pooling
- “Physically-plausible illumination distribution estimation” Egor Ershov ICCV 2023.
- Deep White-Balance Editing, CVPR 2020 (Oral)