

MMG  
MMG  
MMG

# MMG

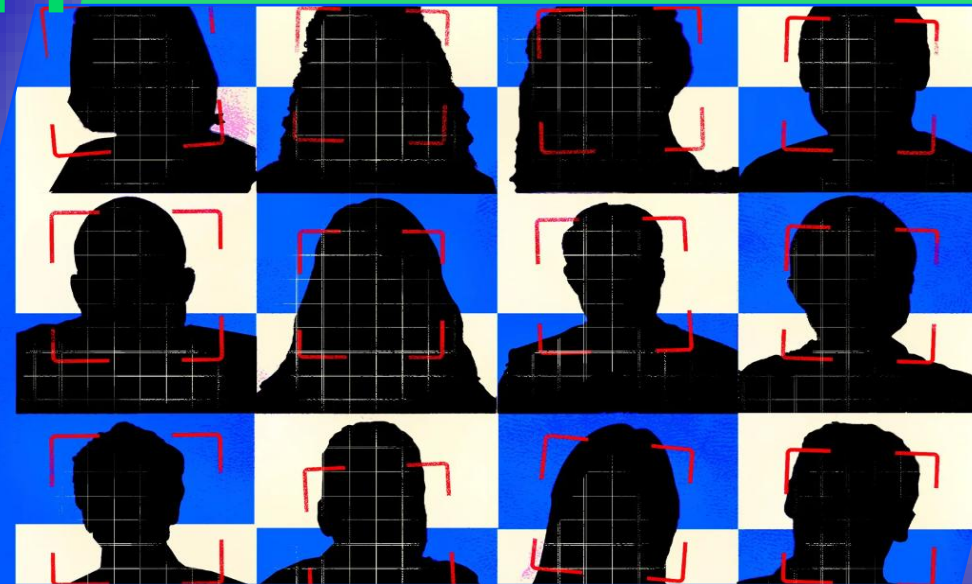
Команда Финансового Университета  
при Правительстве РФ



×



КРИПТОНИТ



## Kryptonite ML Challenge

“Чем правдоподобнее подделка,  
тем опаснее обман.”  
- MMG

XXXX XXXX XXXX XXXX XXXX XX

# О задаче

Задача обучить **энкодер**, работающий для изображений.

Он должен обладать следующими характеристиками:

**Cosine similarity :**

→ **1** для фото одного человека

→ **0** для фото разных людей

→ **0** для фото и дипфейка

Датасет

**10.000** людей  
**≈ 98.000** фото

Метрика

**EER**

# Содержание

MMG  
MMG  
MMG

**1**

Ресерч

**2**

Модель

**3**

Претрейн

**4**

Датасет

**5**

Обучение

**6**

Результат

**7**

Эмбеддинги

**8**

Атаки

**9**

Тупики

**10**

Масштаб

**11**

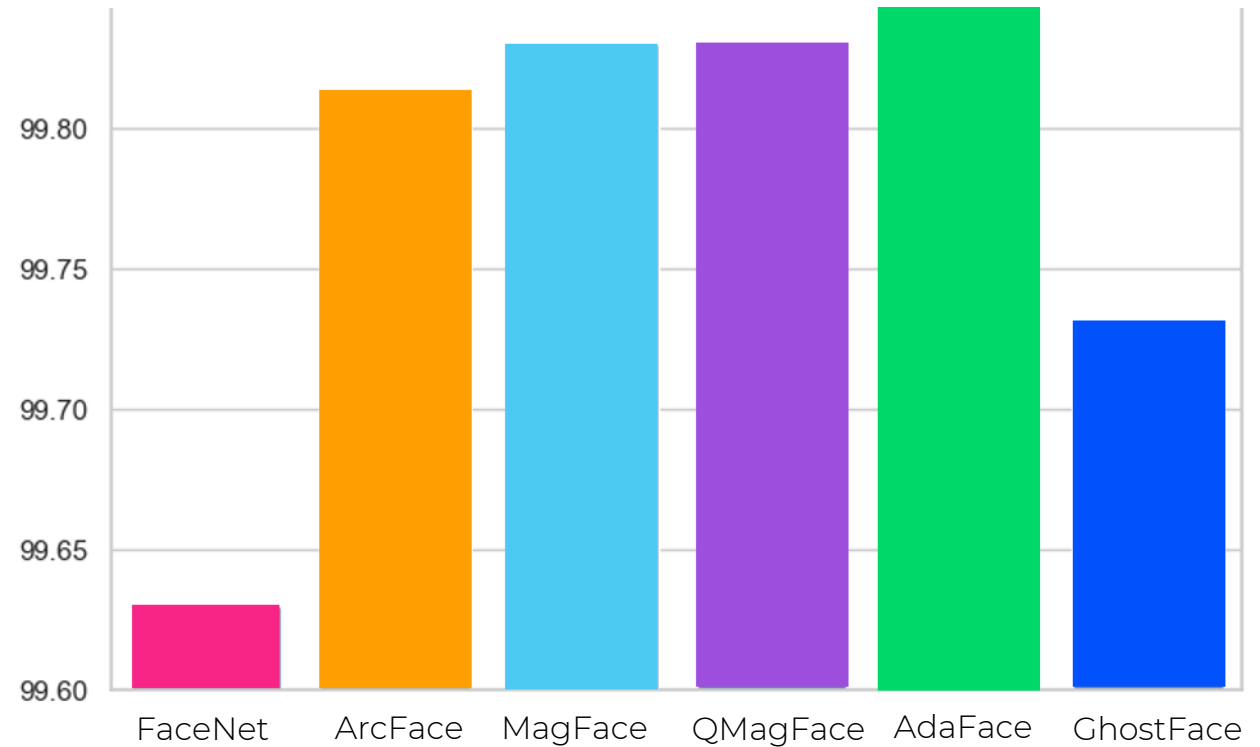
О нас

## Модели

точность, скорость, дипфейки

FaceNet	2015
ArcFace	2019
MagFace	2021
QMagFace	2021
AdaFace	2022
GhostFace	2023

## Бенчмарк LFW



## Разнообразие лосс-функций

Contrastive Loss

Triplet Loss

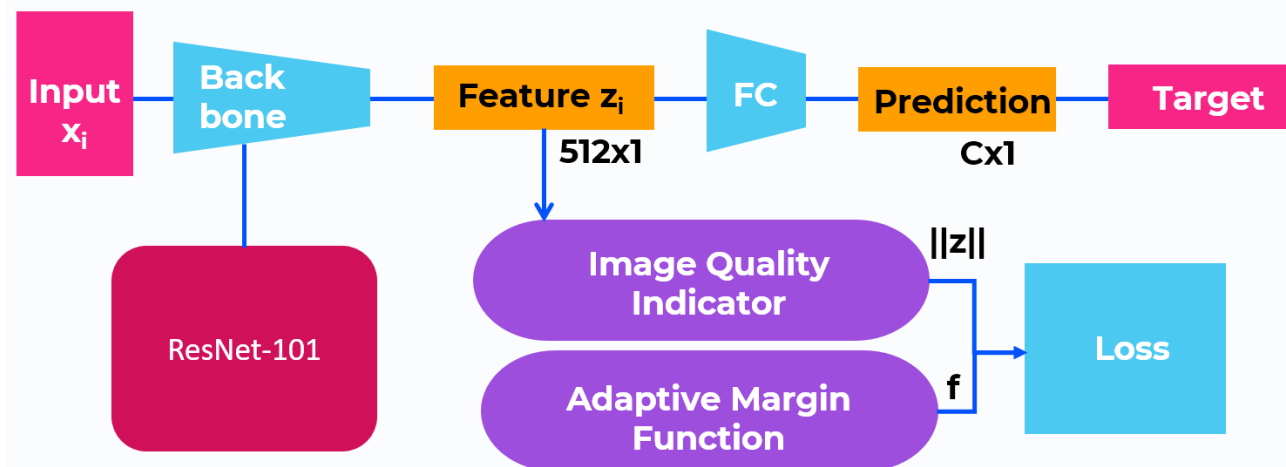
Circle Loss

Center Loss

Hard Negative Mining

Knowledge Distillation

# AdaFace



$$\mathcal{L}_{\text{AdaFace}}(x_i) = -\log \frac{\exp\left(s \left[ \cos(\theta_{y_i} + g_{\text{angle}}(\|\hat{z}_i\|)) - g_{\text{add}}(\|\hat{z}_i\|) \right]\right)}{\exp\left(s \left[ \cos(\theta_{y_i} + g_{\text{angle}}(\|\hat{z}_i\|)) - g_{\text{add}}(\|\hat{z}_i\|) \right]\right) + \sum_{j \neq y_i} \exp(s \cos \theta_j)}$$

Где:

$$g_{\text{angle}}(\|\hat{z}_i\|) = -m \cdot \|\hat{z}_i\|, \quad g_{\text{add}}(\|\hat{z}_i\|) = m \cdot \|\hat{z}_i\| + m, \quad \|\hat{z}_i\| = \text{clip}\left(\frac{\|z_i\| - \mu_z}{\sigma_z/h}, -1, 1\right).$$

## Loss

$\cos \theta$  — косинусное сходство

$y \in \{0, 1\}$  — метка

$s$  — масштабный коэффициент

$m_b$  — базовый margin

$\alpha$  — кф. динамического масштабирования

1. Динамический **margin**:

$$m_d = m_b + \alpha \cdot (1 - \cos \theta)$$

2. Определение **логита**:

$$\text{logit} = s \cdot (\cos \theta - m_d \cdot y) = s \cdot (\cos \theta - (m_b + \alpha \cdot (1 - \cos \theta))y)$$

3. Функция потерь **Adaptive Margin Loss** :

$$\mathcal{L} = -\left[y \cdot \log(\sigma(\text{logit})) + (1 - y) \cdot \log(1 - \sigma(\text{logit}))\right]$$

## Aligner

1

1. MTCNN
2. RetinaFace
3. **CVLFace DFA**

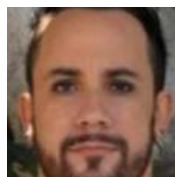
## Сбор и подготовка данных

2

Датасет: **CelebA**

10.000 персон  
200.000 фото

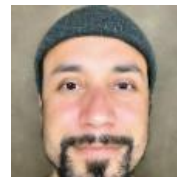
Генерация **дипфейков**:



Original



Arc2Face



InstantID



Ghost



Roop

## Пары

3

**Real + Real** = 1  
(50%)

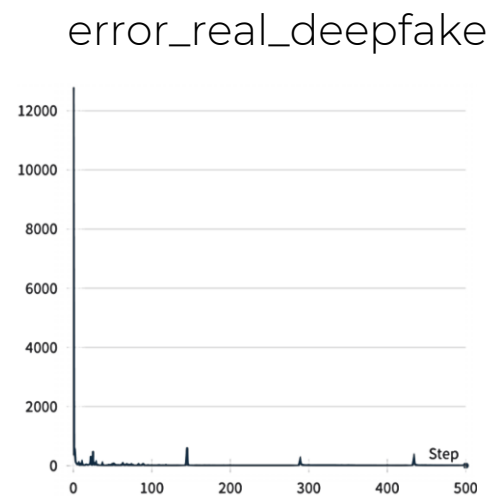
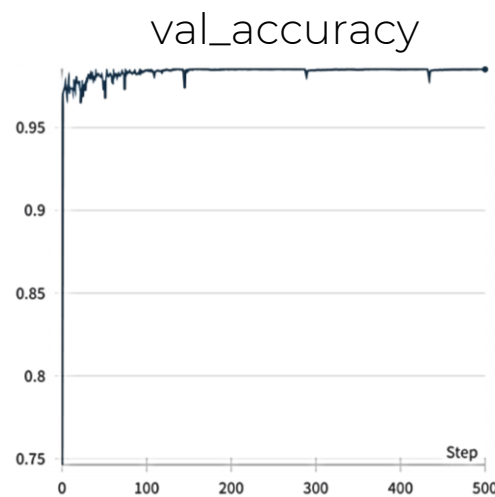
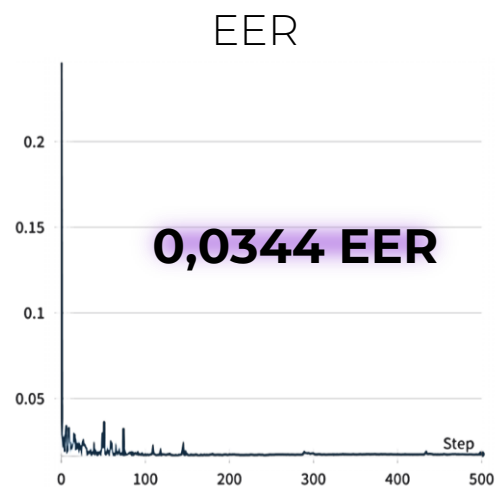
**Real + Fake** = 0  
(25%)

**Real + Other** = 0  
(25%)

# Пре трейн!

## Обучение AdaFace

4



# Датасет

MMG  
MMG  
MMG

## Обнаруженная проблема:

~10% реальных фото неправильно размечены

000010



0.jpg



1.jpg

000204



1.jpg



5.jpg

## Попытки очистить датасет:

1. Использование модели для построения матриц сходств
2. Графовая кластеризация по cosine similarity
3. Удаление выбивающихся изображений

Папка: **000513**



## Результаты:

1. Проблема некорректной разметки частично **решена**
2. Качество модели без привязки к лидерборду **выросло**

XXXX  
XXXX  
XXXX  
XXXX  
XXXX  
XXXX  
XXXX

# Обучение

## Гиперпараметры

## Оптимизатор

**AdamW:**

LR: **5e-6**

Weight decay: **1e-2**

## Scheduler

**ReduceLROnPlateau:**

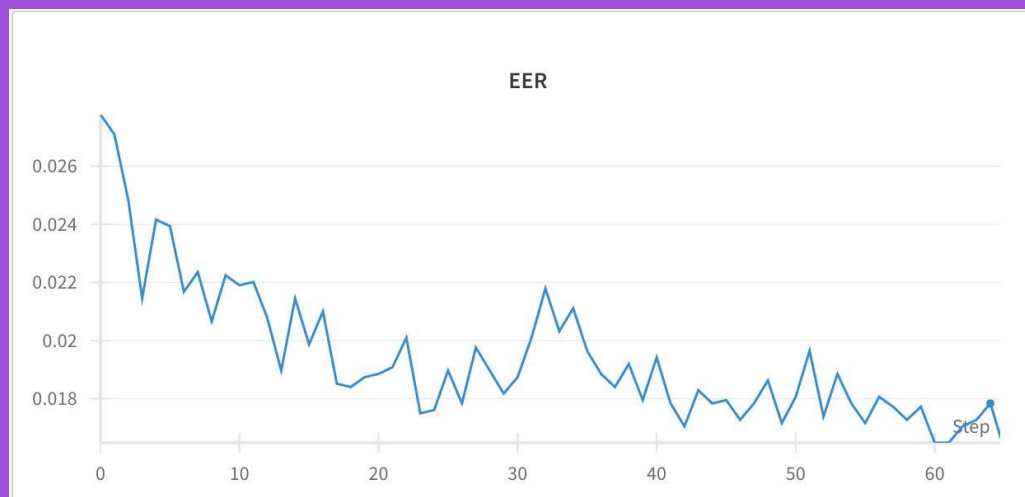
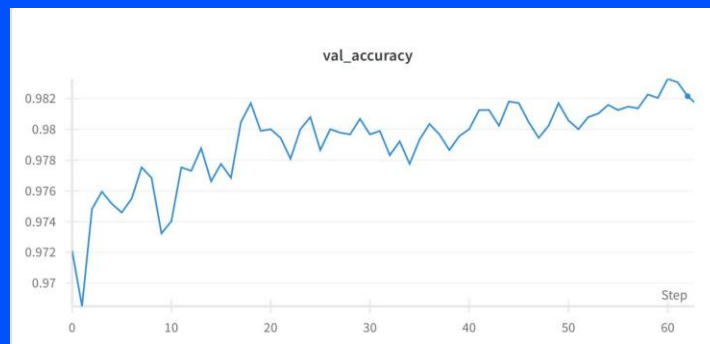
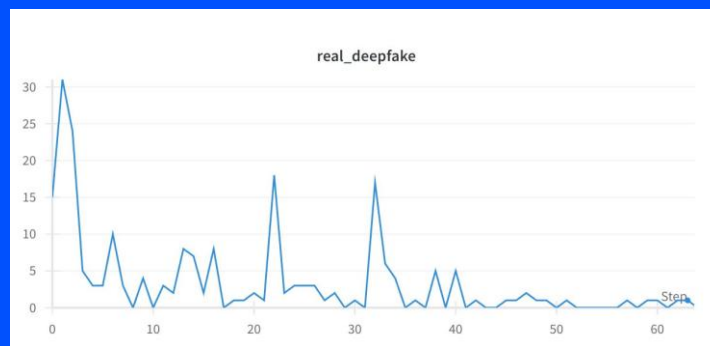
Mode: **min**

Factor: **0.5**

Patience: **10**

Threshold: **1e-4**

Validation threshold: **0.2**



**val\_accuracy**

**error\_real\_deepfake**

**EER:**  
**0.0264**

**MMG**  
**MMG**  
**MMG**



# Результаты



## Скорость:

RTX 4090 — ~900 изображений в секунду

## Стабильность:

Быстрый энкодер, стабильно работает даже на среднем «железе»

ТАБЛИЦА РЕЗУЛЬТАТОВ

						
	NULL	0.62	0.66	0.00	0.00	0.00
	0.62	NULL	0.60	0.00	0.00	0.00
	0.66	0.60	NULL	0.00	0.00	0.00
	0.00	0.00	0.00	NULL	0.96	0.95
	0.00	0.00	0.00	0.96	NULL	0.97
	0.00	0.00	0.00	0.95	0.97	NULL





# MMG

КОМАНДА ФИНАНСОВОГО УНИВЕРСИТЕТА ИТИАБД

НАША КОМАНДА

ТЕСТИРОВАНИЕ

ЗАГРУЗКА ФОТОГРАФИЙ

ПЕРЕТАЩИТЕ ФАЙЛЫ  
ОТЗДА

ЗАГРУЗИТЕ ОТ 2 ДО 10 ФОТОГРАФИЙ

ОТПРАВИТЬ

РЕЗУЛЬТАТ



ВЕРИФИЦИРОВАНО: 0.75

MMG  
MMG  
MMG

ТАБЛИЦА РЕЗУЛЬТАТОВ

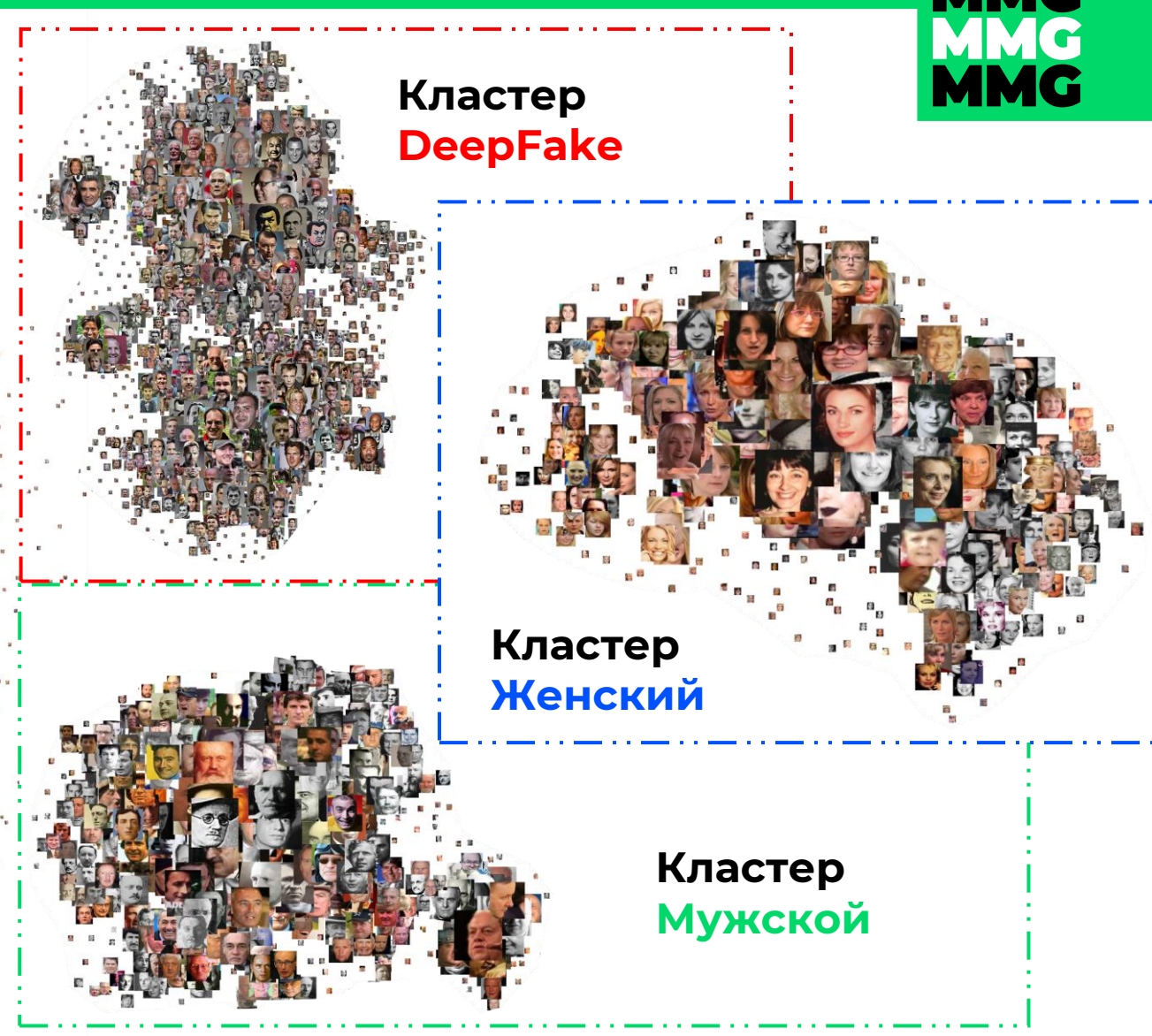
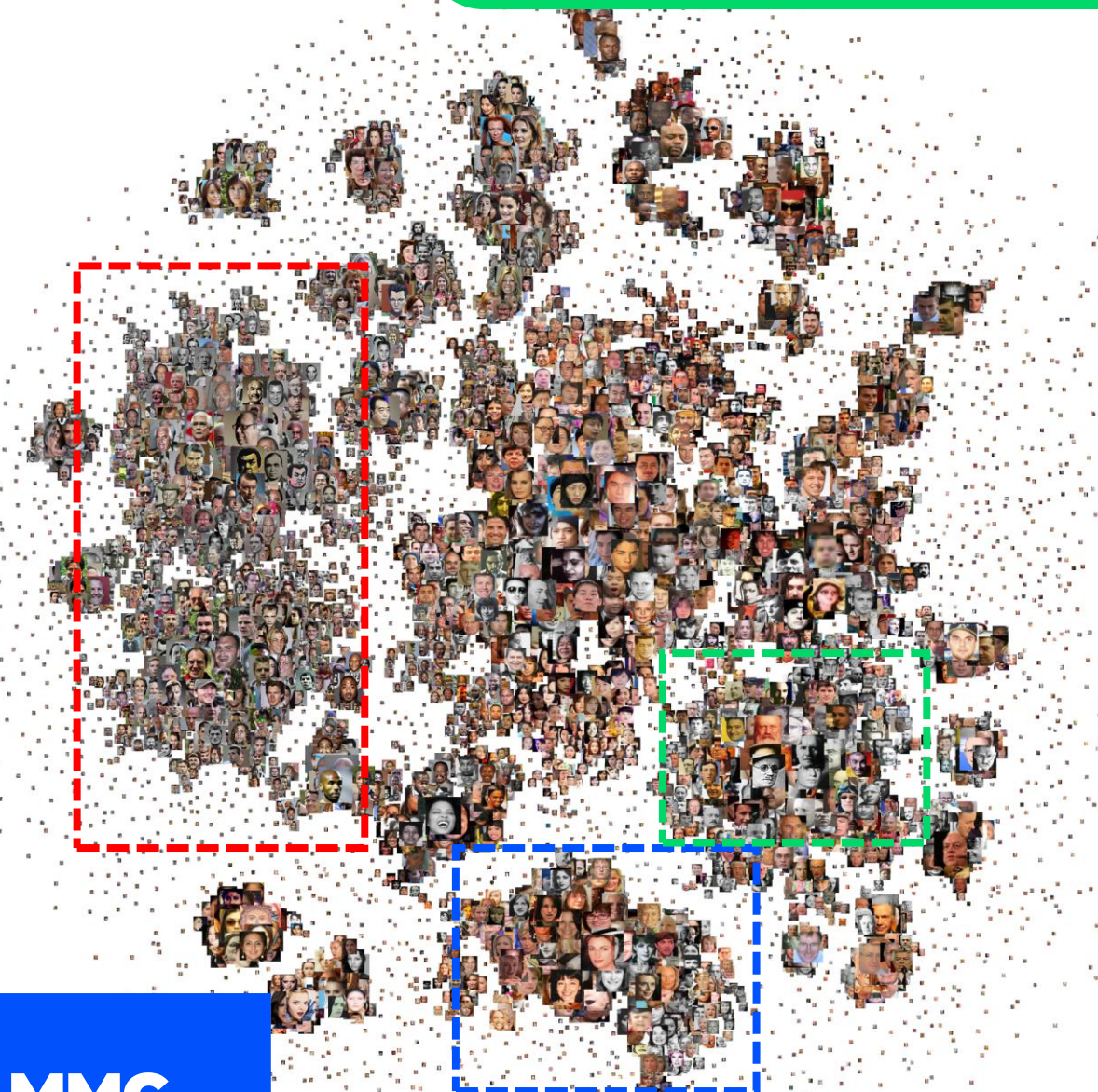
	NULL	0.99	0.65	0.66
	0.99	NULL	0.64	0.66
	0.65	0.64	NULL	0.66
	0.66	0.66	0.66	NULL

# Сайт для инференса



Визуализация эмбедингов / Интерпретация модели







# Атаки

## ШУМ



### Вывод:

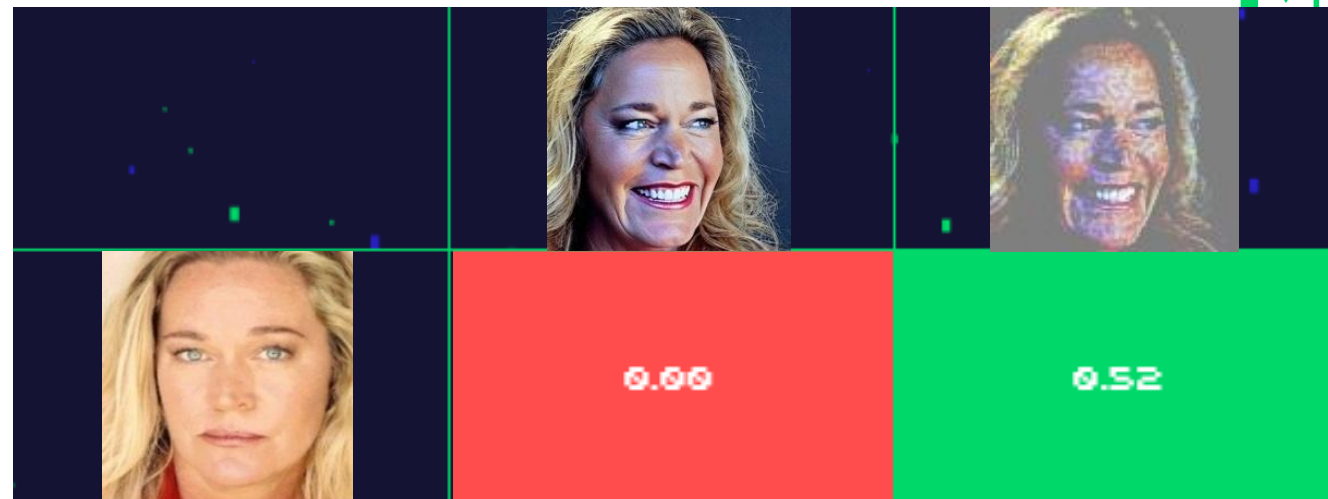
Шум может увеличить схожесть с другой фотографией, но это не оказывает критического влияния.

MMG  
MMG  
MMG

	NULL	0.00	0.00	0.00	0.00	0.13	0.01	0.00	0.00	0.05

# FGSM

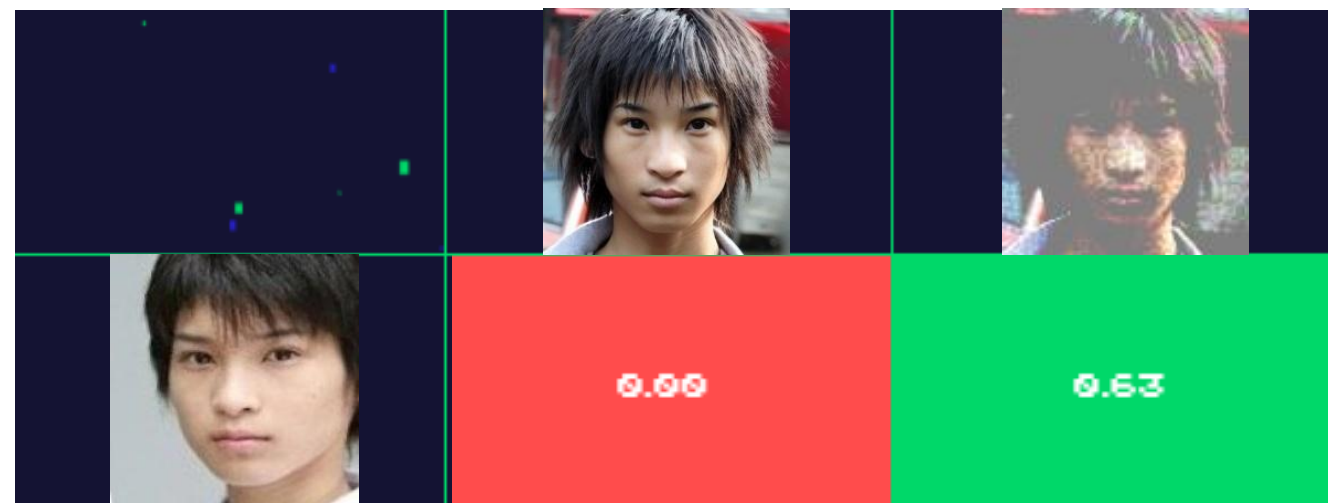
Epsilon = 0.1



Real

DeepFake

FGSM



Real

DeepFake

FGSM

# Тупики

Попробовали, но не использовали в финальной модели

## Выводы:

1. Глубокий ресерч – важно
2. Важность предобучения
3. Сложность  $\neq$  качество
4. Важность предобучения
5. ~~Агрессивная аугментация~~

1

### Triplet loss

Не дало прироста качества

2

### Обучение без претрейна

Чекпоинт **Кима Минчула** не дал прироста по метрике

### Заморозка модели

**Полная заморозка** — недообучение  
Дополнительные ветки с **attention** - усложнение без улучшения

3

4

### Гиперпараметры и оптимизаторы

Перепробовали множество конфигураций – не все давали улучшения

5

### Аугментация Автолейблинг

Ухудшило качество модели

# Масштаб



**1** Использование более продвинутых и современных DeepFake и FaceSwap моделей

**2** Увеличение количества обучающих данных

**3** Борьба с FGSM-подобными атаками





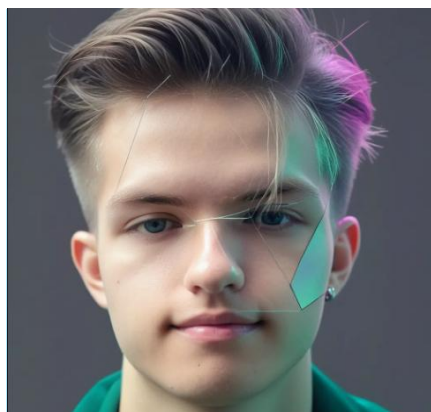
MMG  
MMG  
MMG

Команда Финансового Университета  
при Правительстве РФ

# MMG



**Артем Таратин**



**Денис Маликов**



**Даниил Аль-Натор**



**Илья Обухов**

Команда **MMG** благодарит **Криптонит** за возможность участия в хакатоне!



КРИПТОНИТ

XXXX XXXX XXXX XXXX XXXX XX



**MMG**  
**MMG**  
**MMG**

**MMG**  
**MMG**  
**MMG**