



Летняя школа Intel – ННГУ по компьютерному зрению

# Корпорация монстров

Макаров Александр, Кулагин Павел, Кузенкова Анастасия

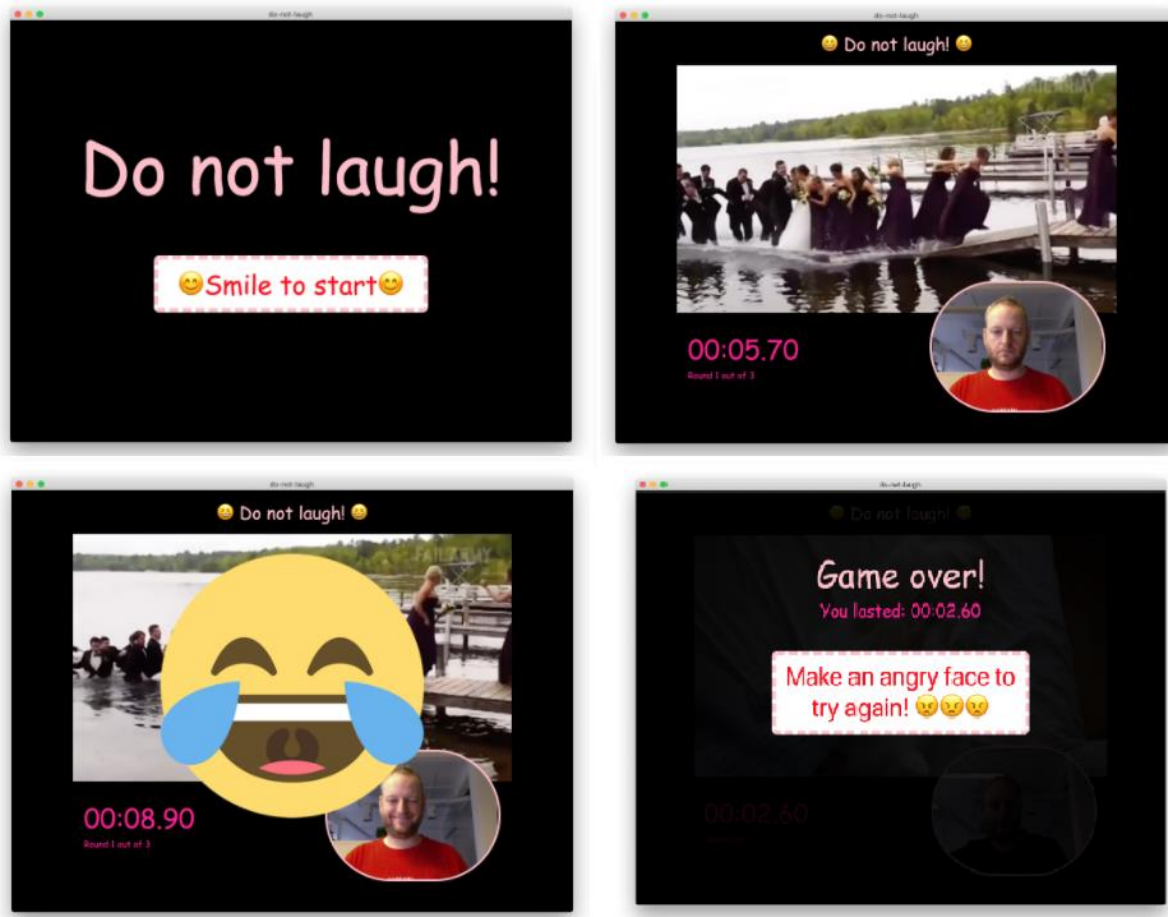
# Постановка задачи

- Создание интерактивной игры по идее мультфильма "Корпорация монстров" - уровни с накоплением энергии.
- Использование детектора эмоций по лицу человека для распознавания радости
- Добавить распознавание смеха по аудио



# Похожие приложения

## Do not laugh



Игра воспроизводит случайное видео на YouTube. Задача игрока не засмеяться. Смех распознается по мимике человека.

<https://medium.com/@andreas.schallwig/do-not-laugh-a-simple-ai-powered-game-3e22ad0f8166>

# Похожие приложения

## Детектирование смеха:

<https://github.com/jrgillick/laughter-detection>

<https://github.com/ideo/LaughDetection>

В данных репозиториях содержатся нейросети, детектирующие смех. Для их работы требуется сложная предобработка звука.

## Статья о детекции смеха.

Kimiko Ryokai, Elena Durán López, Noura Howell, Jon Gillick, and David Bamman (2018), "Capturing, Representing, and Interacting with Laughter," CHI 2018

<https://dl.acm.org/doi/10.1145/3173574.3173932>

## Challenges in Laughter Recognition

Ha Ha or Uh Huh?

Hee Hee or Hey Hey?

**Acoustic Model**

Does this sound  
like a laugh?

**Language Model**

?

**Prediction**

# Преимущество нашего подхода

- Детекция смеха и по лицу с веб камеры, и по аудио с микрофона
- Примеры программ с аудио детекцией требуют либо довольно сложной предобработки, либо выделения из звука признаков, по которым проводится предсказание. Благодаря интерфейсу OpenVINO можно избежать предобработки, сразу подав выбранной нами модели аудио с микрофона.



# MemCheck



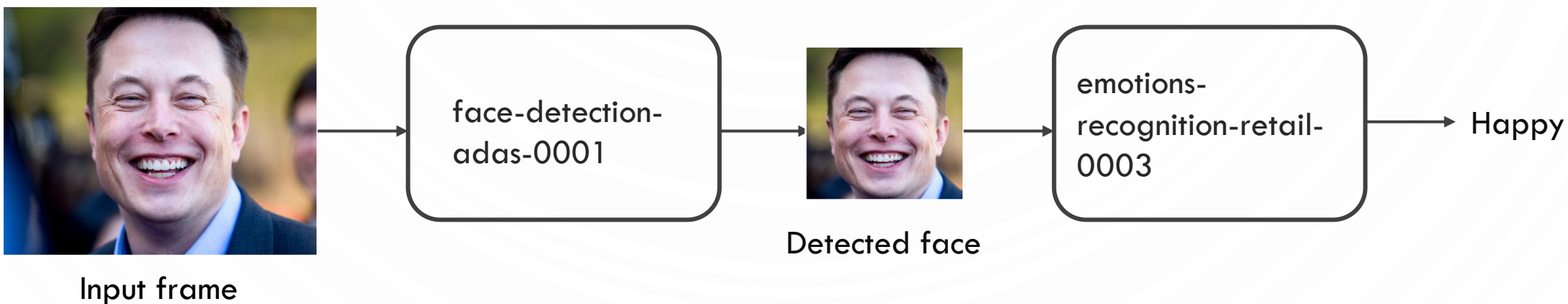
“MemCheck” – мини-игра, которая позволяет оценить уровень мемов той или иной площадки. С помощью изображения с веб-камеры и звука с микрофона определяется уровень счастья участника игры, и заполняется шкала радости. Мемы считаются годными, если за заданный промежуток времени пройден установленный «порог счастья».



# Реализация проекта

## Распознавание радости:

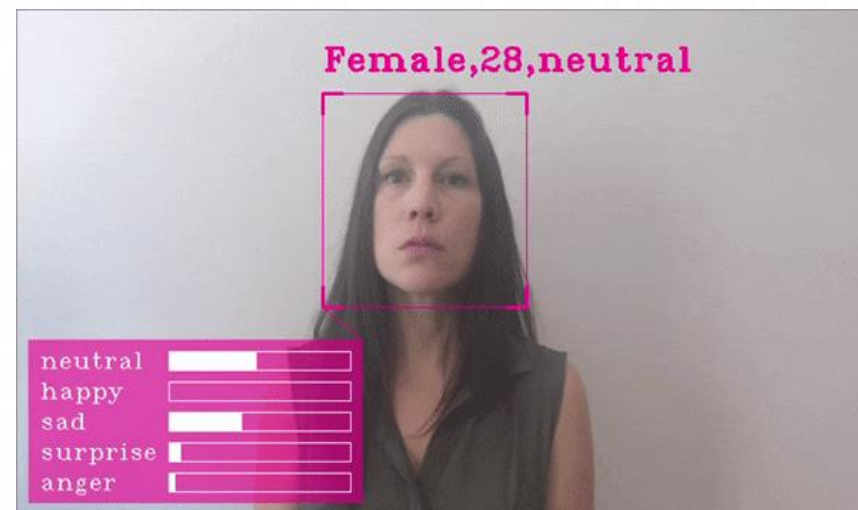
- face-detection-adas-0001 – сеть, детектирующая лицо
- emotions-recognition-retail-0003 – модель, распознающая 5 видов эмоций



# Реализация проекта

## Сравнение моделей по детекции лица:

Название модели	AP	GFlops	MParams
face-detection-0200	86.74%	0.786	1.828
face-detection-0204	92.89%	2.406	1.851
face-detection-adas-0001	94.1%	2.835	1.053
face-detection-retail-0004	83.00%	1.067	0.588
ultra-lightweight-face-detection-slim-320	83.32% (mAP)	0.1724	0.2844



## emotions-recognition-retail-0003:

Точность – 70.20% , GFlops – 0.126, MParams – 2.483



# Реализация проекта

AcNet – сеть, классифицирующая звуки (53 класса).

Точность – 86.3%, GFlops – 1.42, MParams – 2.71

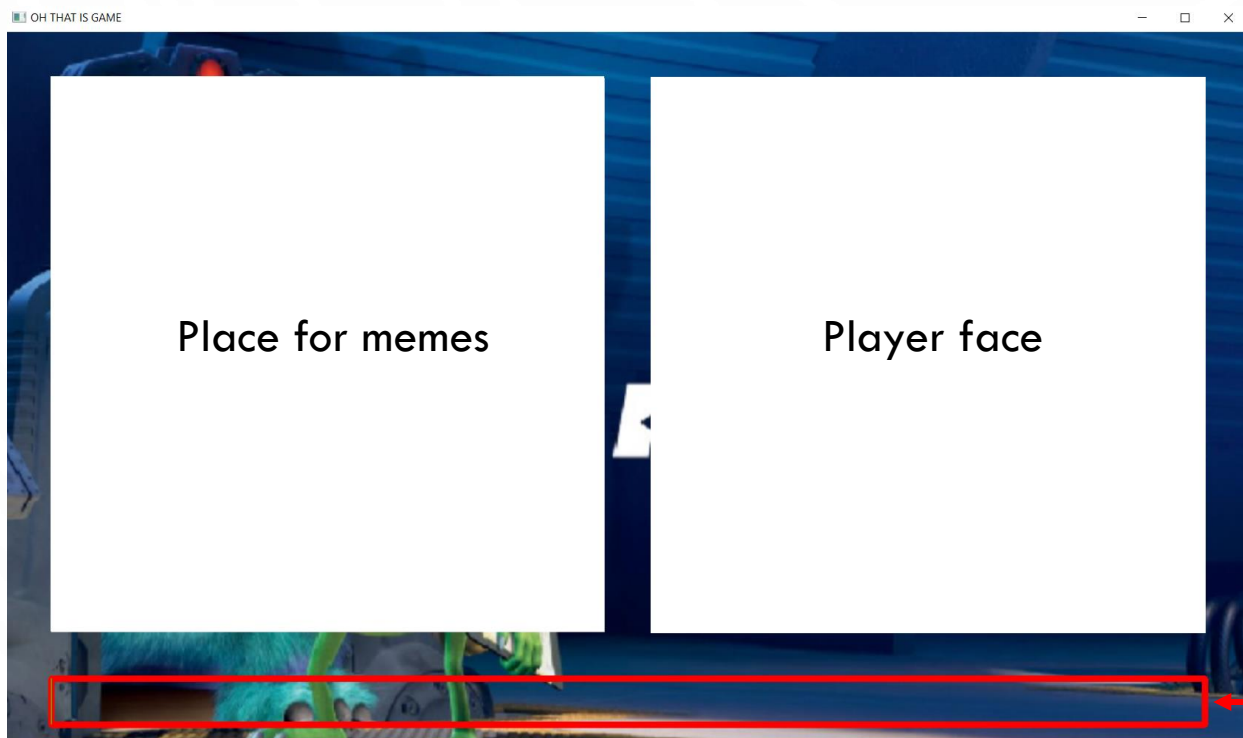


dog - 5-231762-A-0.wav



# Реализация проекта

Интерфейс приложения состоит из одного окна, где отображается мем и видео с веб-камеры. Следующий мем открывается при нажатии на клавишу N.



Neutral



Smile



Laugh



Scale of happiness



Демо

# Итоги

- Реализовали приложение распознающее улыбку и смех человека в реальном времени
- Приложение можно использовать для интеграции в платформу, занимающуюся рейтингом развлекательных интернет платформ

# Q&A



Спасибо за внимание!