



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

Факультет информатики, математики и  
компьютерных наук  
Программа подготовки бакалавров по  
направлению Бизнес-информатика

## ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

«Технологии глубокого обучения для  
мониторинга эмоций групп пользователей для  
систем видеоаналитики»

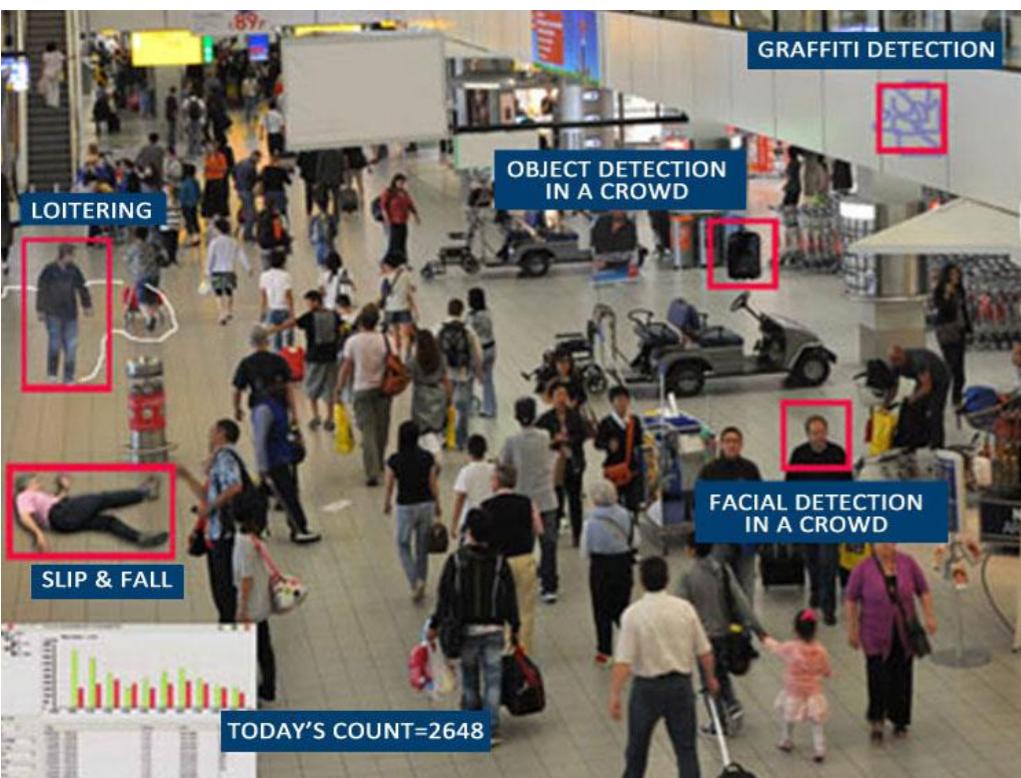
Выполнил студент группы 14БИ-2

Тарасов Александр Вячеславович

Научный руководитель: Доктор технических  
наук, профессор кафедры ИСиТ  
Савченко Андрей Владимирович

Нижний Новгород, 2018

# АКТУАЛЬНОСТЬ





# ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

**Цель работы:** исследовать применимость технологий глубокого обучения для мониторинга эмоций групп пользователей в системах видеоаналитики.

**Задачи работы:**

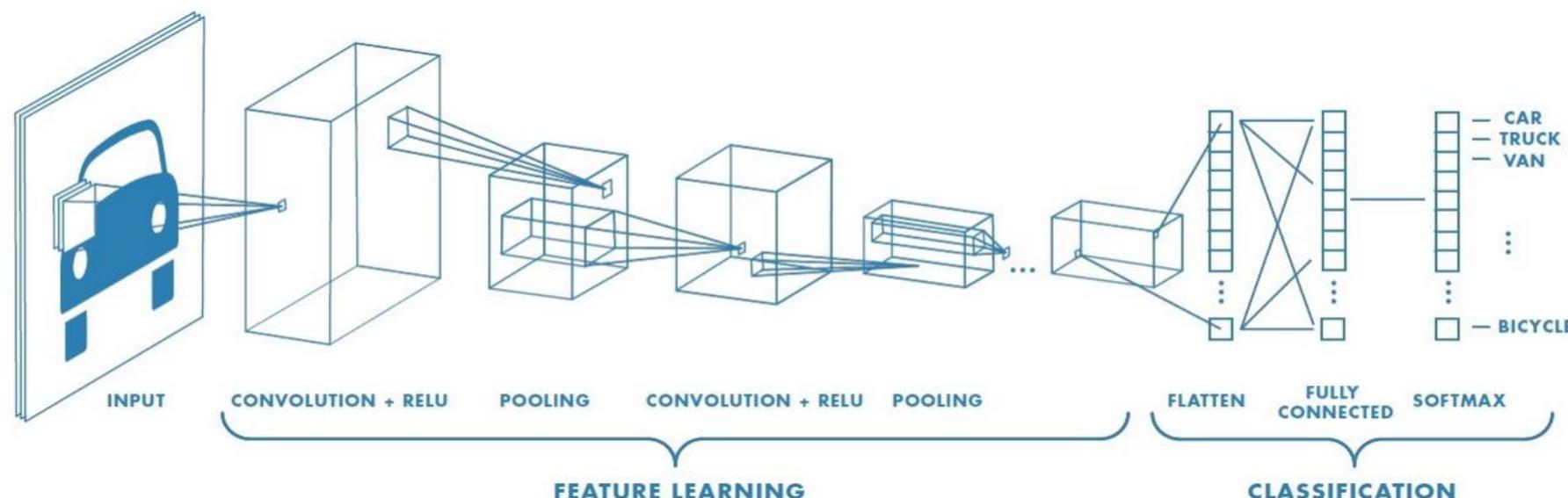
1. провести аналитический обзор современной литературы по распознаванию эмоций в системах видеоаналитики;
2. выявить требования к системам видеоаналитики, установленным в торговых центрах, сетевых магазинах и т.п., целью которых является анализ реакции пользователей на маркетинговые кампании и процесс обслуживания;
3. определить эффективный подход к реализации прототипа программного продукта, соответствующего выявленным требованиям;
4. провести экспериментальные исследования предлагаемого подхода;
5. расширить предлагаемый алгоритм для задачи распознавания эмоций групп людей по видео-контенту;
6. разработать прототип программного продукта согласно требованиям.



# АНАЛИТИЧЕСКИЙ ОБЗОР ЛИТЕРАТУРЫ (1)

## ТЕХНОЛОГИИ ГЛУБОКОГО ОБУЧЕНИЯ

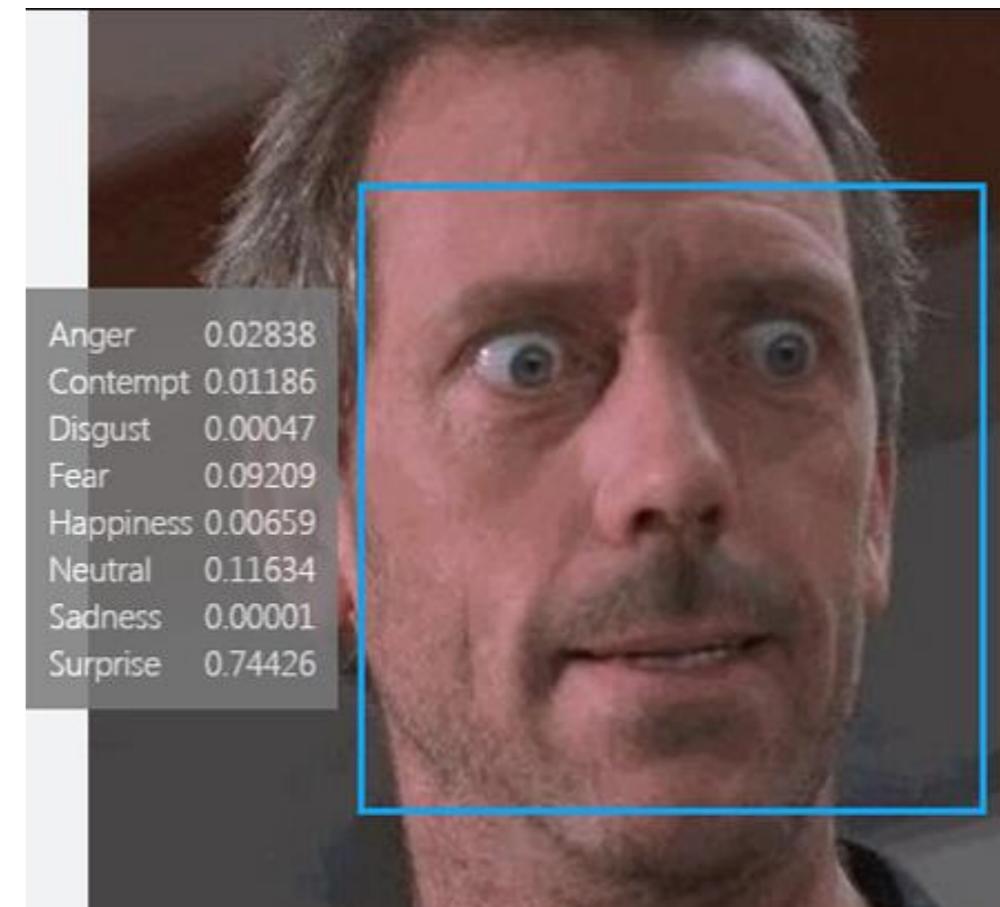
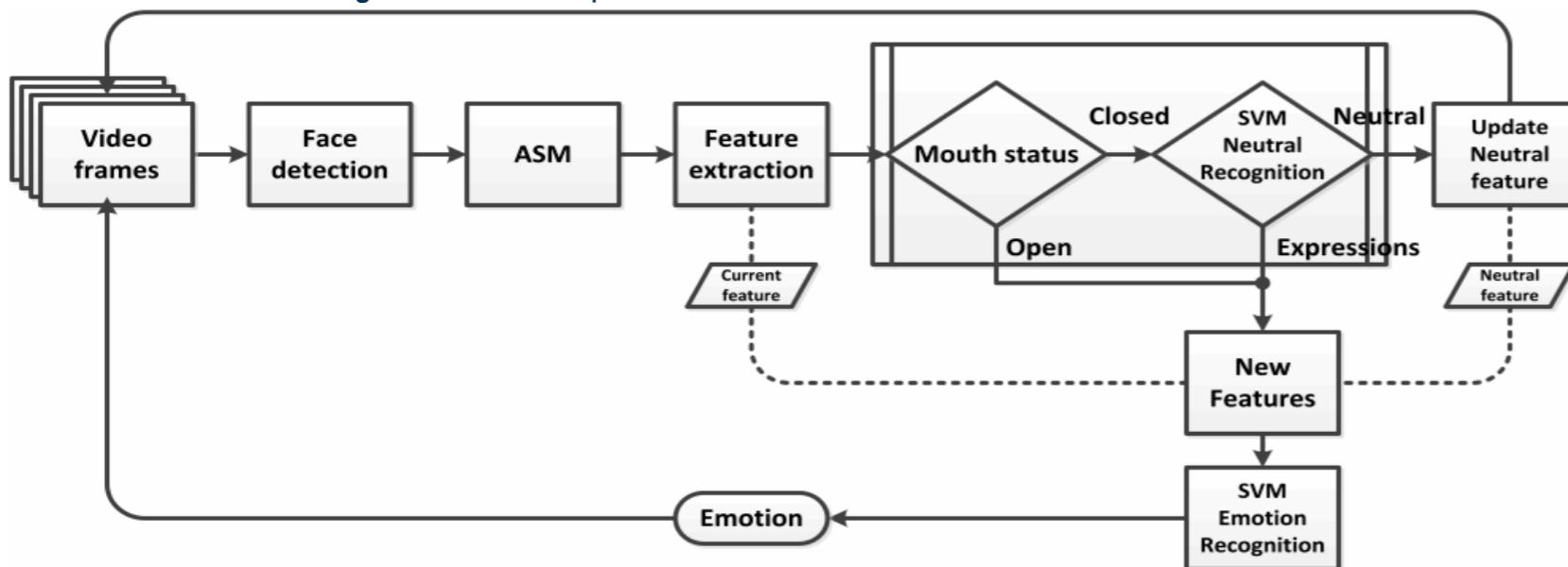
- Litjens G., Sánchez C. I., Timofeeva N., Hermsen M., Nagtegaal I., Kovacs I., Hulsbergen C., Bult P. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis // Nature, 2016.
- Goodfellow I., Bengio Y. Deep Learning / – MIT Press, 2016.
- Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift // Proceedings of the 32nd International Conference on International Conference on Machine Learning, 2015
- He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proceedings of the 2016 IEEE Computer Society Conference on Computer Vision and Pattern Recognition
- Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // arXiv preprint arXiv: 1409.1556, 2014.
- Large Scale Visual Recognition Challenge 2012 (ILSVRC2012) // URL: <http://www.image-net.org/challenges/LSVRC/2012/>
- Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012



# АНАЛИТИЧЕСКИЙ ОБЗОР ЛИТЕРАТУРЫ (2)

## МЕТОДЫ РАСПОЗНАВАНИЯ ЭМОЦИЙ ПО ИЗОБРАЖЕНИЯМ ЛИЦ

1. Rassadin A., Gruzdev A., Savchenko A. Group-level Emotion Recognition using Transfer Learning from Face Identification // Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI), 2017
2. Surace L., Patacchiola M., Sönmez E. B., Spataro W., Cangelosi A. Emotion Recognition in the Wild using Deep Neural Networks and Bayesian Classifiers // Proceedings of the 19th ACM ICMI, 2017
3. Tan L., Zhang K., Zeng X., Peng X., Qiao Y. Group Emotion Recognition with Individual Facial Emotion CNNs and Global Image Based CNNs // Proceedings of the 19th ACM ICMI, 2017
4. Ghimire D., Lee J. Geometric feature-based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines // Sensors. 2013
5. Happy S.L., George A., Routray A. A real time facial expression classification system using local binary patterns // In Proceedings of the 4th International Conference on Intelligent Human Computer Interaction, 2012
6. Suk M., Prabhakaran B. Real-time mobile facial expression recognition system—A case study // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014



Пример метода распознавания эмоций [6]



# ВЫЯВЛЕННЫЕ ТРЕБОВАНИЯ

Программный продукт должен:

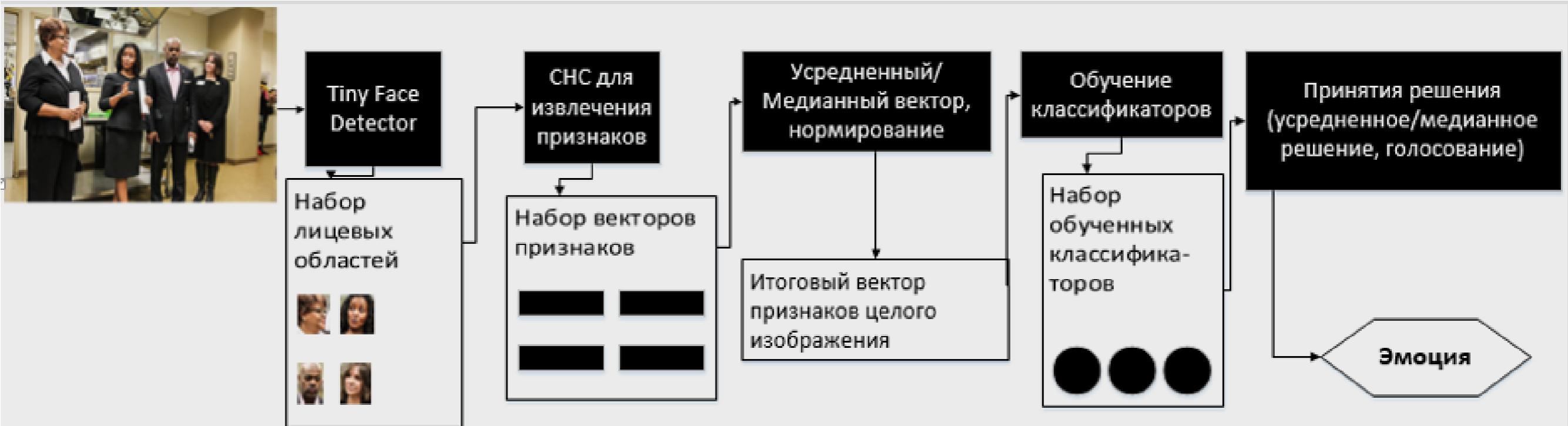
- Иметь возможность распознать количество лиц, превышающее число 10, на изображениях групп людей.

- быть способным классифицировать эмоции выделенных лиц, разрешение которых менее 128 на 128 пикселей.

- определять общий эмоциональный окрас изображения и короткого видеопотока



# ПРЕДЛОЖЕННЫЙ ПОДХОД К РЕАЛИЗАЦИИ ПРОТОТИПА(1)

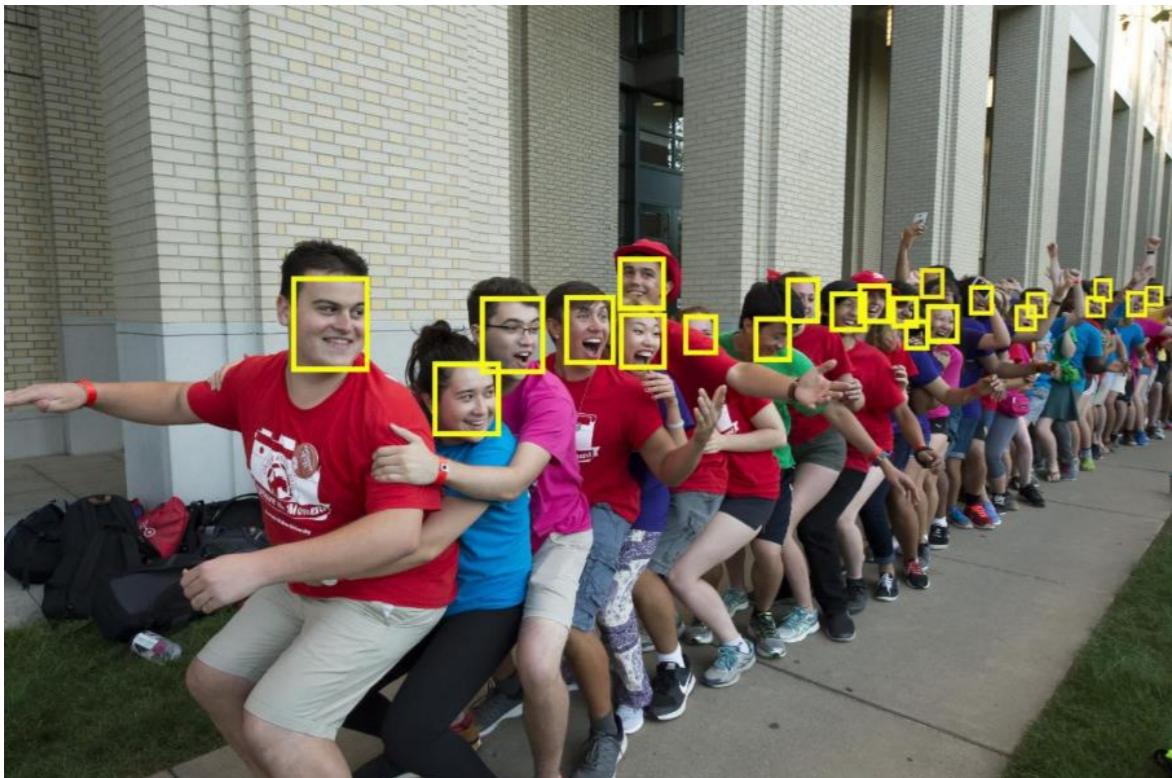




# ПРЕДЛОЖЕННЫЙ ПОДХОД К РЕАЛИЗАЦИИ ПРОТОТИПА(2) ДЕТЕКТОР ЛИЦЕВЫХ ОБЛАСТЕЙ

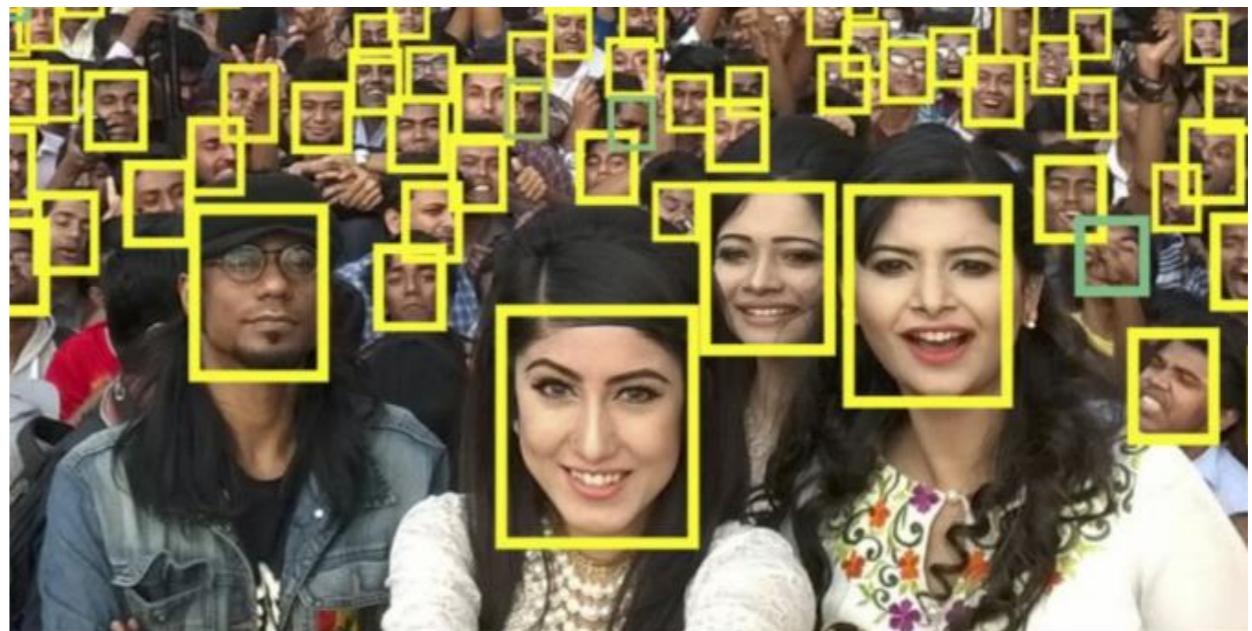
## Tiny Face Detector

Hu P., Ramanan D. Finding Tiny Faces.  
*2017 IEEE Conference on Computer  
Vision and Pattern Recognition (CVPR)*



## MTCNN

Zhang K., Zhang Z., Li Z., Qiao Y.  
Joint face detection and alignment  
using multitask cascaded  
convolutional networks.  
2016 IEEE Signal Processing Letters





# ПРЕДЛОЖЕННЫЙ ПОДХОД К РЕАЛИЗАЦИИ ПРОТОТИПА(3)

## ИЗВЛЕЧЕНИЕ ПРИЗНАКОВ ЛИЦ

Для извлечения признаков лиц  
использована сверточная нейронная  
сеть, обученная на наборе данных  
**FER-2013**.



Примеры изображений  
из набора FER-2013

# РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ (1)

## ОПИСАНИЕ ЭКСПЕРИМЕНТОВ

Характеристики ПК, на котором проводилось экспериментальное исследование:

- Intel Core i5-5200U CPU
- 64-bit ОС
- NVIDIA GeForce 940M

### Описание экспериментов:

- Набор данных с EmotiW 2017  
<https://sites.google.com/site/emotiwchallenge/>
- Обучающая выборка - 3630 фотографий
- Тестовый набор – 2065 фотографий
- Классификация групповых изображений по трем эмоциональным категориям

“positive”



“neutral”



“negative”





# РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ (2)

## ТОЧНОСТЬ КЛАССИФИКАЦИИ

Классификатор	Точность	Набор классификаторов в ансамбле	Точность		
			Усредненное решение	Медианное решение	Решение голосованием
RBF SVM	0.693		0.702	0.702	0.713
GradientBoosting	0.701		0.712	0.724	0.723
Extra Trees	0.702		0.732	0.731	0.746
Linear SVM	0.703		0.718	0.743	0.746
Bagging	0.707		0.727	0.732	0.751
RandomForest	0.709		0.696	0.696	0.753
Ada Boost	0.709		0.723	0.753	0.755
<b>Используемые методы повышения точности:</b> Многопараметрическая оптимизация grid search (повышение точности на 2-8%) Метод главных компонент (повышение точности на 2%)					

# РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ (3)

## ВРЕМЯ ОБРАБОТКИ

Классификатор	Среднее время обработки одного изображения в миллисекундах	
	CNN-5	FCNN
Random Forest	35	28
Extra Trees	36	29
Linear SVC	36	30
Gradient Boosting	37	30
AdaBoost	38	30
SVC	41	30
Bagging	43	33
Ансамбль (Random Forest, SVC и Linear SVC)	49	40
Ансамбль всех классификаторов	76	69



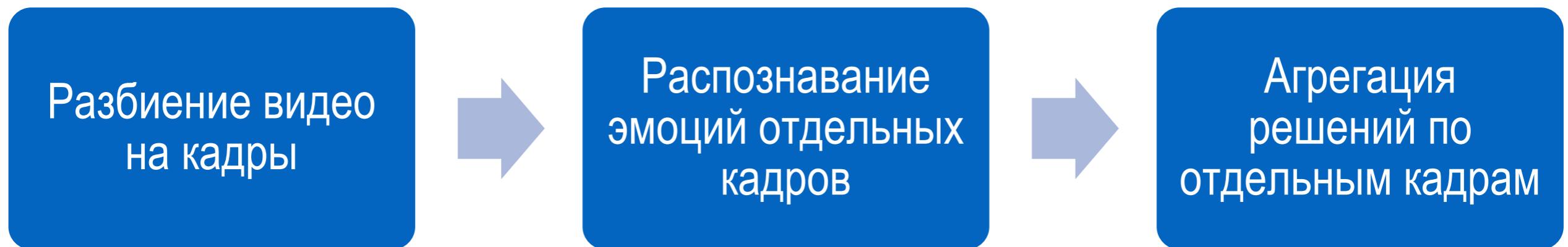
# РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ (4)

## СРАВНЕНИЕ С СУЩЕСТВУЮЩИМИ ПОДХОДАМИ

Точность предложенного подхода

- 1) на **23% выше** по сравнению с базовым методом классификации специально подобранных признаков изображений [Dhall A., Goecke R., Ghosh S., Joshi J., Hoey J., Gedeon T. From individual to group-level emotion recognition: EmotiW 5.0. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017]
- 2) на **20% выше** по сравнению с методом применения СНС на тепловых картах [Shamsi S., Rawat B. P. S., Wadhwa M. Group Affect Prediction Using Multimodal Distributions. Available at: <https://arxiv.org/pdf/1710.01216.pdf>]
- 3) на **9% выше** по сравнению с точностью метода на основе Tiny Face Detector и передачи знаний [Rassadin A., Gruzdev A., Savchenko A. Group-level Emotion Recognition using Transfer Learning from Face Identification. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017]
- 4) на **8% выше** по сравнению с комбинацией глубоких СНС с байесовскими классификаторами [Surace L., Patacchiola M., Sönmez E. B., Spataro W., Cangelosi A. Emotion Recognition in the Wild using Deep Neural Networks and Bayesian Classifiers. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017]
- 5) на **8% ниже** по сравнению с методом победителя EmotiW 2017, основанного на использовании двух СНС для определения эмоций отдельных лиц и извлечения признаков из остальной части изображения. Но **вычислительная сложность предложенного подхода значительно ниже** [Tan L., Zhang K., Wang K., Zeng X., Peng X., Qiao Y. Group Emotion Recognition with Individual Facial Emotion Convolutional Neural Networks and Global Image Based Convolutional Neural Networks. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017]

# АЛГОРИТМ РАБОТЫ С ВИДЕО-КОНТЕНТОМ



# ПРОТОТИП ПРОГРАММНОГО ПРОДУКТА





# АПРОБАЦИЯ РАБОТЫ

- 1) IV Междисциплинарная студенческая конференция «Интеллектуальный город: наука как искусство» в НИУ ВШЭ – Нижний Новгород [<https://nnov.hse.ru/studentconf/>]
- 2) XXI Международная конференция **SCM-2018** в Санкт-Петербургском государственном электротехническом университете «ЛЭТИ» [<https://scm.eltech.ru/2018/en/>]
- 3) 7-я конференция по Data Science и приложениям **AIST-2018** в Московском Политехническом Университете [<https://aistconf.org/>]



# ЗАКЛЮЧЕНИЕ

При выполнении работы были получены **следующие результаты**:

1. По результатам аналитического обзора литературы были выбраны наиболее эффективные технологии глубокого обучения для применения при решении поставленной задачи
2. Выявлены требования к системам видеоаналитики, целью которых является анализ реакции покупателей на маркетинговые кампании и процесс обслуживания.
3. Определен эффективный подход к реализации прототипа программного продукта
4. Проведены экспериментальные исследования предлагаемого подхода.
5. Работа предлагаемого алгоритма расширена для задачи распознавания эмоций групп людей по видео-контенту.
6. Разработан прототип программного продукта согласно требованиям.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

Спасибо за внимание!