

**ПЛЕХОВ П. В., ДВИНЯНИНОВ А. А.**  
**ОБНАРУЖЕНИЕ ЛИЦ В ВИДЕОПОТОКЕ ПЛОХОГО КАЧЕСТВА ДЛЯ**  
**ВЫДЕЛЕНИЯ ЭМОЦИОНАЛЬНЫХ ТОЧЕК**

*УДК 004.93, ВАК 1.2.2, ГРНТИ 28.23.37*

Обнаружение лиц в видеопотоке  
плохого качества для выделения  
эмоциональных точек

Detection of faces in a video stream of  
poor quality to highlight emotional  
points

**П. В. Плехов, А. А. Двинянинов**

**P. V. Plekhov, A. A. Dvinyaninov**

Березниковский филиал ФГАОУ ВО  
«Пермский национальный  
исследовательский политехнический  
университет», г. Березники

Berezniki branch of FSAEI HE "Perm  
National Research Polytechnic  
University", Berezniki

*В статье описаны  
результативные действия,  
позволяющие локализовать лица на  
кадрах низкокачественного  
видеоряда и выделять на них  
эмоциональные точки для  
последующего анализа.*

*The article is about effective  
actions that allow you to localize faces  
on low-quality video stream and  
highlight emotional points on them for  
further analysis.*

**Ключевые слова:** видео,  
распознавание лиц, локализация,  
эмоциональные точки

**Keywords:** video, face recognition,  
localization, emotional points

### **Введение**

Задача обнаружения лица на изображении часто является первым шагом в процессе решения задачи более высокого уровня — распознавания лица, деталей лица или его мимики. Кроме того, информация о присутствии и количестве лиц на изображении может быть полезна в системах автоматического учета числа посетителей; системах пропускного контроля в учреждениях, аэропортах и метро; автоматических системах предотвращения несчастных случаев; в фототехнике для автоматической фокусировки на лице человека и во многих других областях. Отдельным направлением деятельности является использование мимических точек для идентификации психологического состояния человека [1].

Распознавание лиц в плохом качестве в режиме реального времени усложняет задачу. Положение лица к камере, недостаточное освещение, фон, масштаб изображения лица – все это может препятствовать обнаружению лица.

**Целью** является разработка приложения для обнаружения лиц в режиме реального времени и выделения на них эмоциональных точек.

**Основными задачами** являются:

- нахождение наиболее подходящего алгоритма обнаружения лиц в реальном времени;
- нахождение алгоритма расставления эмоциональных точек;
- разработка приложения.

Существует много эффективных алгоритмов обнаружения лиц со своими плюсами и минусами. На наборе фотографий разного качества произведено определение того, насколько справляются с определением лиц разные алгоритмы (Таблица 1). Критериями сравнения были: скорость работы, точность определения и особенности каждого из них.

Таблица 1. Сравнение методов

Метод определения	Скорость работы алгоритма (с)	Точность определения (всего 12 лиц на фотографии)	Особенности
HOG (Histogram of Oriented Gridients) + SVM	1,57	6/12	Определяет лицо, если точно смотреть в камеру, при хорошем освещении и качестве
CNN (Сверточная нейронная сеть)	157,76	10/12	Хорошо определяет лица, долго работает. непригоден для определения лиц в реальном времени. Для ускорения нахождения может использовать GPU
Face_recogtiton (api python) нейронная сеть	1,29	6/12	Те же что и у HOG
MT-CNN (Многозадачная каскадная сверточная нейронная сеть)	1,6	12/12	С хорошим освещением обнаруживает всех
OpenCV Haar Cascade	0,2	8/12	Не работает на лица, которые не смотрят в камеру. Может неправильно находить лица(можно исправить параметрами)
YOLOv3	0,12	12/12	Быстро и наиболее точно обнаруживает лица с возможностью использования GPU

В итоге, наиболее подходящим для плохого качества показал себя YOLOv3 [2]. Он хорошо показывает себя при недостаточном освещении, повернутых лицах.

Главная особенность нейронной сети YOLO по сравнению с другими состоит в том, что большинство систем применяют CNN несколько раз к разным

регионам изображения, а YOLO применяет CNN один раз ко всему изображению сразу [3, 4]. Благодаря этому сеть обладает высокой скоростью и точностью обнаружения объектов.

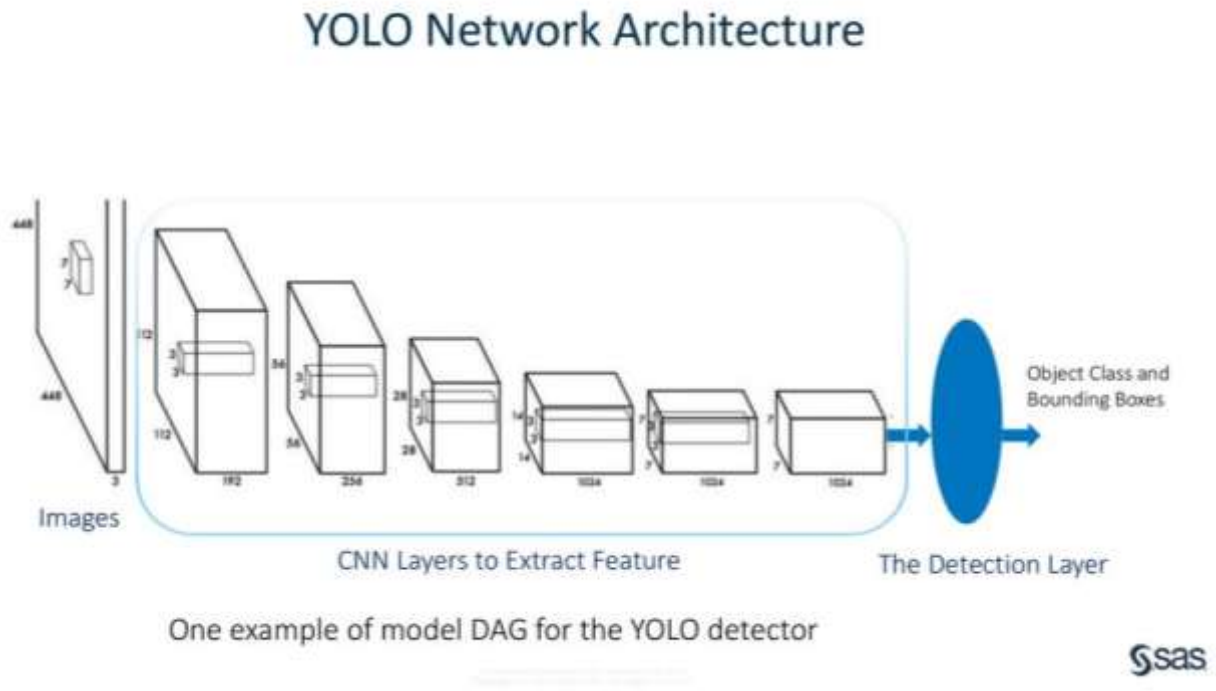


Рисунок 1. Архитектура YOLOv3

Поскольку YOLO «смотрит» на изображение только один раз, то метод с плавающим окном не подходит для данной сети. Вместо этого, все изображение с помощью сетки разбивается на ячейки размером  $S \times S$ . После этого для каждой ячейки предсказываются два показателя: ограничительные рамки (bounding boxes) и вероятности нахождения в них нужного объекта (confidence) [4].

После того как были получены все ограничительные рамки и соответствующие вероятности для всех искомых объектов необходимо избавиться от дубликатов bounding box для каждого объекта. Для этого применяется алгоритм «подавления не-максимумов» (Non-maximum suppression), который ищет ограничительные рамки, содержащие объект распознавания.

У лица есть несколько черт, которые можно идентифицировать, например, наши глаза, рот, нос и т. д. Для выделения эмоциональных точек, определяющих их положение и динамику, была выбрана библиотека Dlib [5]. Когда мы используем алгоритмы DLib для обнаружения этих черт, мы фактически получаем карту точек, окружающих каждую черту.

На этой карте, состоящей из 67 точек (называемых ориентирами) [6], можно определить следующие особенности:

- Точки челюсти = 0–16
- Точки правой брови = 17–21
- Левая надбровная дуга = 22–26
- Носовые точки = 27–35
- Точки правого глаза = 36–41

- Точки левого глаза = 42–47
- Точки рта = 48–60
- Точки губ = 61–67

С помощью данной библиотеки и алгоритма YOLOv3, которое это лицо и обнаруживает [7], выделяются 67 контрольных точек. Ниже представлен пример работы (Рисунок 2).

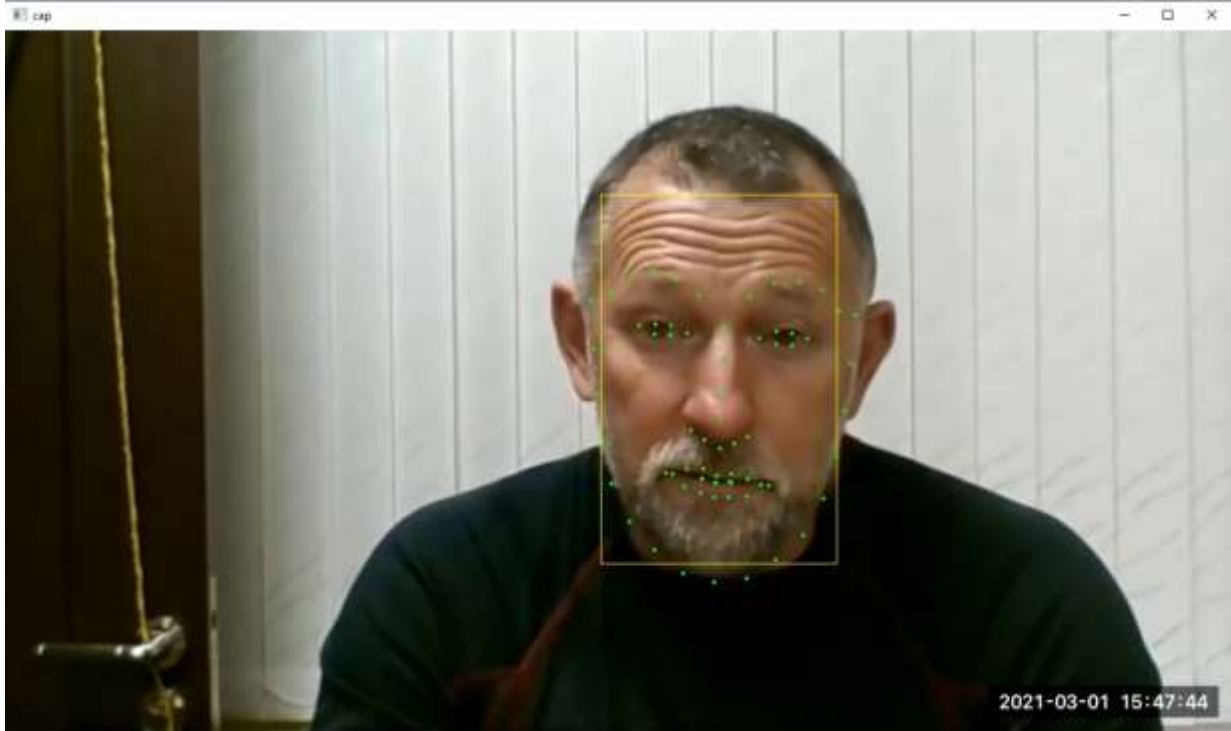


Рисунок 2. Результат работы программы

### Заключение

В соответствии с целью работы разработана программа распознавания лиц и выделения эмоциональных точек на видео на языке программирования Python для сверточной нейронной сети YOLOv3. Использованы обученные модели от YOLOv3 и Dlib для распознавания лиц и выделения эмоциональных точек.

### Список использованных источников и литературы:

1. Батенькина О. В., Иноземцева К. В. Метод распознавания эмоций детей дошкольного возраста с использованием мимических выражений // Омский научный вестник, 2017. №. 6 (156). С. 146-150.
2. Архитектура YOLOv3 [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1804.02767>, (дата обращения: 02.04.2023).
3. Болховитина Е. И. Исследование моделей сверточных нейронных сетей YOLOV3 и RETINANET для задачи детектирования лица человека на изображении // StudNet, 2022. № 5 (6), С. 5439-5448.
4. Затонский А. В., Варламова С. А., Малышева А. В., Мясников А. А. Использование видеографической информации для уточнения динамической стохастической модели процесса флотации калийной руды // Интернет-журнал Науковедение. 2017. Т. 9. № 2. С. 87.

5. Затонский А. В., Малышева А. В. Совершенствование алгоритмов распознавания видеопотока для идентификации переходных процессов флотационной машины калийной руды // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2019. № 1. С. 26-39.

6. Асташенкова Л. К., Рощупкина С. Н., Кудринская К. В. Распознавание ключевых точек лица на изображении человека // Молодой ученый. 2019. С. 23-25

7. Визильтер Ю.В., Желтов С.Ю., Осоков М.В. Система распознавания и визуализации характерных черт человеческого лица в реальном времени на персональной ЭВМ с использованием web-камеры // Тр. Междунар. Конф. По компьютерной графике и машинному зрению «Графикон 2002». Нижний Новгород, 2002. Н. Новгород: Графикон, 2002. С. 251-254.

### **List of references:**

1. Batenkina O. V., Inozemtseva K. V. A method for recognizing the emotions of preschool children using facial expressions // Omsk Scientific Bulletin, 2017. No. 6 (156). pp. 146-150.

2. YOLOv3 architecture, <https://arxiv.org/abs/1804.02767>, (date of access: 05/02/2023).

3. Bolkhovitina E. I. Research of models of convolutional neural networks YOLOV3 and RETINANET for the problem of detecting a person's face in an image // StudNet, 2022. No. 5 (6), P. 5439-5448.

4. Zatonsky A. V., Varlamova S. A., Malysheva A. V., Myasnikov A. A. Use of videographic information to refine the dynamic stochastic model of the potash ore flotation process. 2017. V. 9. No. 2. S. 87.

5. Zatonsky A.V., Malysheva A.V. Improving video stream recognition algorithms for identifying transient processes of a potash ore flotation machine. Bulletin of the Astrakhan State Technical University. Series: Management, Computer Engineering and Informatics. 2019. No. 1. S. 26-39.

6. Astashenkova L. K., Roshchupkina S. N., Kudrinskaya K. V. Recognition of facial key points in a human image // Young scientist. 2019. S. 23-25.

7. Vizilter Yu.V., Zheltov S.Yu., Osokov M.V. The system of recognition and visualization of the characteristic features of a human face in real time on a personal computer using a web-camera // Tr. International Conf. On computer graphics and machine vision "Grafikon 2002". Nizhny Novgorod, 2002. Nizhny Novgorod: Grafikon, 2002, pp. 251-254.