

АЛГОРИТМ РАСПОЗНОВАНИЯ ЭМОЦИЙ ЛЮДЕЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ

*Т.Е. Мамонова, доцент ОАР.
Е.А. Шергин, студент гр. 8Е11.
Томский политехнический университет
E-mail: eas141@tpu.ru*

Введение

На сегодняшний день проводится достаточно много мероприятий, от небольших уличных выставок, до крупных праздников, где присутствует большое количество людей. Для поддержания порядка необходимо постоянно мониторить настроенность людей, их поведение и соответствие окружающей атмосфере происходящих вокруг событий, и, при обнаружении несоответствия, как можно быстрее принимать соответствующие действия.

Поэтому целью работы является создание автоматизированной системы мониторинга общественных мест, которая будет детектировать лицо и идентифицировать эмоцию человека, что поможет оперативно определить несвойственное поведение и настроение человека или группы людей при проведении общественных мероприятий. В рамках указанной цели в данной работе выполняется задача распознавания эмоций с применением нейросетевого подхода.

Описание алгоритма

Для решения задачи распознавания эмоций был написан код на языке Python, в котором задействованы две составляющие: модель искусственной нейронной сети и модель HAAR детектора лица из библиотеки OpenCV. Детектирование лица было решено реализовать при помощи HAAR каскадного детектора из библиотеки OpenCV т. к. скорость обработки изображения выше, чем при использовании нейросети. Идентификация эмоции, в свою очередь, реализована при помощи модели искусственной нейронной сети т. к. это даст наибольшую точность распознавания.

Алгоритм состоит из следующих действий:

- 1) при помощи камеры получаем изображение;
- 2) преобразуем изображение в трёхмерный массив с закодированными пикселями в формате BGR для дальнейшей обработки;
- 3) при помощи HAAR каскадного детектора из библиотеки OpenCV находим рамки лица;
- 4) применяя модель нейросети на изображение лица, получаем предсказание каждой эмоции;
- 5) обводим лицо на изображении цветом, соответствующим эмоции;
- 6) выводим в консоль список эмоций, которые присутствуют на изображении.

Архитектура нейронной сети показана на рисунке 1. Нейронная сеть состоит из 12 слоёв, из которых – 8 свёрточные, 4 линейные. Каждый слой, кроме выходного, использует функцию активации «ReLU» т. к. он имеет меньшую подверженность к затуханию функции градиентного спуска, что ускоряет сходимость стохастического градиента. Последний слой, в свою очередь, использует функцию активации “softmax”.

Размер тренировочной выборки 28709 изображений, 7 классов предсказания. Angry: 3995; Disgust: 436; Fear: 4097; Happy: 7215; Neutral 4965; Sad: 4830; Surprise: 3171 (шт).

Размер проверочной выборки 3589 изображений, 7 классов предсказания.

Обучение нейронной сети происходило при параметрах:

- Графический ускоритель GPU с ядрами CUDA
- Также была применена аугментация тренировочной выборки для улучшения качества предсказаний модели.
- Размер подаваемых данных при обучении за одну итерацию - 64
- Оптимизатор – Adam
- Learning-rate – 0.0015
- Количество эпох – 50
- Функция потерь – Перекрёстная энтропия
- Функция сохранения модели при повышении точности на валидационных данных.

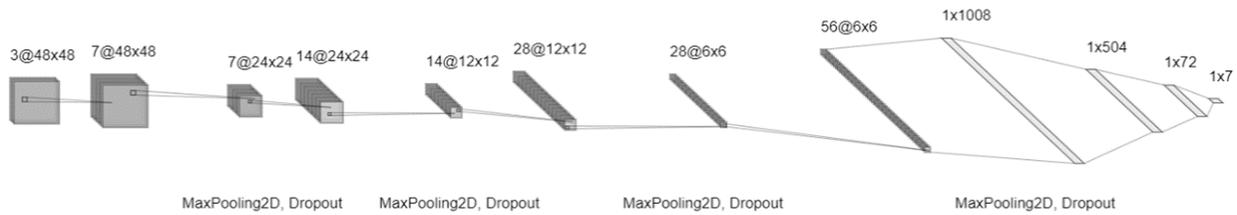


Рис.1. Архитектура нейронной сети

На 43 эпохе получилась наибольшая точность на валидационной выборке, поэтому было решено использовать модель с контрольной точки, а не с полных 50 эпох обучения.

Используемый детектор – HAAR cascade frontal face alt2 из библиотеки OpenCV, т.к. он имеет достаточные для поставленной задачи параметры. Существуют и ограничения его использования, такие как: 1) при наклоне головы, лицо детектору найти невозможно из-за особенности алгоритма; 2) проблема детекции лица человека, если его не видно полностью, например, из-за медицинской маски.

Собрав детектор и модель в единую систему, получаем функцию определения лиц на изображении с обводкой их цветами, соответствующими эмоциям. Блок-схема алгоритма представлена на рисунке 2. При написании кода программы была использована информация из источников [2–4].

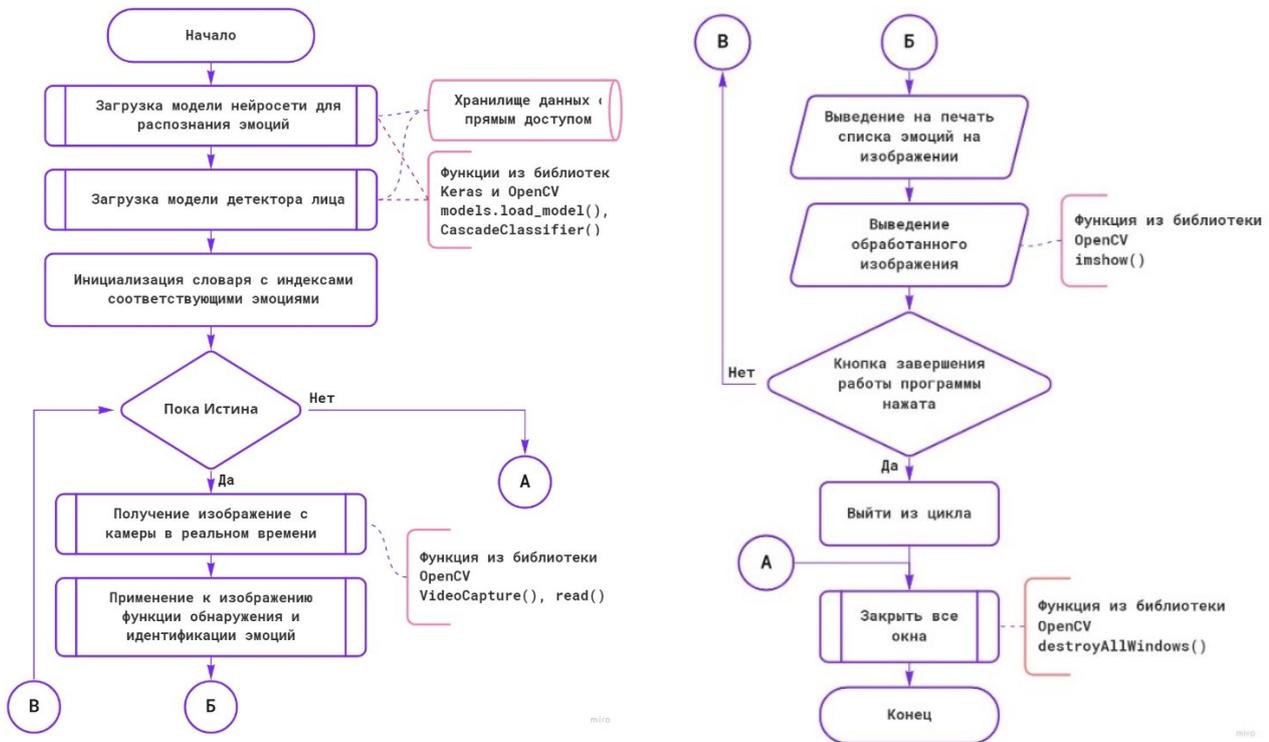


Рис. 2. Блок-схема алгоритма

Размеры изображений лиц при ддетектировании могут быть разного размера, для использования сети их необходимо привести к виду входного слоя сети, а именно к размерности $(1 \times 48 \times 48 \times 3)$, где первый параметр – количество изображений, 2й и 3й – размер в пикселях, 4й – количество каналов. Изначально при прочтении изображения значения закодированных пикселей лежат в отрезке $[0; 255]$, поэтому их также нужно привести к нормальному виду, а именно, чтобы значения были в отрезке $[0; 1]$

При каждой итерации алгоритма вокруг найденных лиц рисуются прямоугольники с цветом соответствующим эмоции.

Красному цвету соответствуют: испуг (fear), злость (angry);

Жёлтому цвету соответствуют: отвращение (disgust), грусть (sad), удивление (surprise);

Зелёному цвету соответствуют: нейтральное (neutral), радость (happy).

Тестирование алгоритма распознавания эмоций людей

Для оценки работоспособности системы были проведены тесты на изображениях с различным количеством людей, высоким и низким качеством, ближним и дальним расположением людей от камеры (рисунок 3).

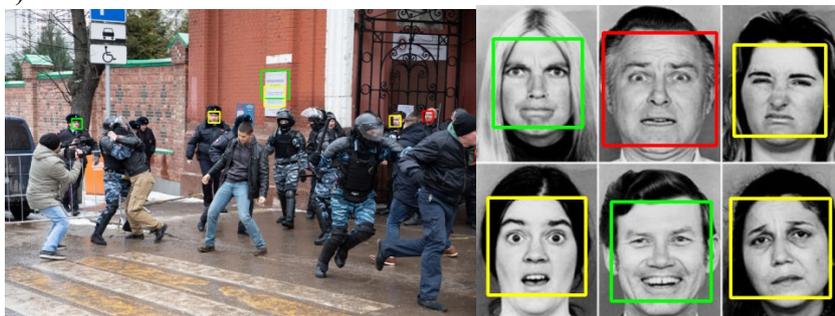


Рис. 3. Результаты оценки работоспособности сети

Заключение

В результате тестирования можно сделать вывод о том, что разработанный алгоритм распознавания эмоции по лицу человека является работоспособным. Точность была определена аналитически, путём проверки сети на тренировочной выборке, и составила 55 %. Результат в реальных условиях значительно превысил значение точности на проверочной выборке, расхождение связано с качеством валидационных данных, а именно – несоответствие некоторых изображений своим классам. С учётом ряда проведённых экспериментов по идентификации эмоций, точность алгоритма составила 70 %. При проведении экспериментов были выявлены проблемы, связанные с детектированием лица и определением эмоции из-за следующих причин:

- 1) наклон лица;
- 2) наличие маски на лице человека;
- 3) низкое качество изображения;
- 4) ложные срабатывания.

Список использованных источников

1. Датасет для обучения нейросети. [Электронный ресурс]. – URL: <https://www.kaggle.com/juniorbueno/rating-opencv-emotion-images> (дата обращения: 13.02.2022)
2. Документация к библиотеке OpenCV. [Электронный ресурс]. – URL: <https://docs.opencv.org/4.x/index.html> (дата обращения: 13.02.2022)
3. Документация к библиотеке Keras. [Электронный ресурс]. – URL: <https://keras.io/api/> (дата обращения: 13.02.2022)
4. Видеоурок о принципе работы и обучении нейросети. [Электронный ресурс]. – URL: <https://avt.global/neuralnets> (дата обращения: 11.02.2022)