

ОПРЕДЕЛЕНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА С ПОМОЩЬЮ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Е. Л. Цалко

Белорусский государственный университет, г. Минск;

katrin.tsalko@gmail.com;

науч. рук. – В. С. Садов, канд. техн. наук, доц.

Уже достаточно давно существует потребность в создании автоматизированных систем распознавания эмоций. И на сегодняшний день такие технологии уже применяются во многих сферах жизни. Целью данной работы было провести анализ уже существующих систем распознавания эмоций с помощью методов машинного обучения, выявить их достоинства, а также предложить решение по устранению недостатков. Также в рамках данной работы были описаны модели классификации эмоций, выявлен общий алгоритм работы систем распознавания эмоций и предложен собственный алгоритм по распознаванию эмоций на основе сиамской нейронной сети. Данный алгоритм решает проблему больших обучающих выборок, поскольку обучается всего на нескольких изображениях.

Ключевые слова: сиамская нейронная сеть; методы машинного обучения; распознавание эмоций; модель классификации эмоций; система кодирования лицевых движений.

Среди существующих методов классификации эмоций наиболее распространенными являются три подхода: дискретный, многомерный и гибридный. Дискретный подход формирует семантическое поле с определенным набором значений для каждой эмоции в отдельности («Атлас эмоций» Пола Экмана [2], 5 базовых эмоций). Многомерный: описывает эмоции в многомерном пространстве с помощью набора значений выбранных параметров («Двумерное пространство модели эмоций Дж. Рассела», параметры: сила и знак эмоции). Гибридный метод представляет собой слияние многомерного и дискретного подходов. В нем любая эмоция может быть рассмотрена как отдельное состояние, так и как часть многомерного пространства. Как пример можно привести модель «Песочных часов эмоций». На основе дискретного метода и теории профайлинга была создана «Система кодирования лицевых движений», которая представлена в виде таблицы. Данная система позволяет закодировать конкретное выражение лица путем создания его модели из так называемых единиц действия. Также была создана «Эмоциональная система кодирования лицевых движений». Где каждой эмоции ставится в соответствие уже целая формула, а не числовое значение.

Анализ современных продуктовых решений показал, что описанные системы кодирования успешно применяются в уже существующих си-

стемах распознавания эмоций. В табл. 1 представлены основные сведения о рассматриваемых программах.

Таблица 1

Системы распознавания эмоций

Название ПО	Информация о компании	Возможности ПО	Преимущества	Недостатки
Face Reader	Компания: Noldus Information Technology (Нидерланды)	Используя видеосъёмку лиц, программа умеет распознавать базовые эмоции, пол, возраст и этническую принадлежность.	Средний процент распознавания основных эмоций – 89%; воспринимает любой наклон лица в плоскости изображения; есть возможность определять эмоции в реальном времени; хорошая визуализация (диаграммы и гистограммы эмоций)	Программа не распознает людей в очках; программа по-разному воспринимает людей с разным цветом кожи; повернутое лицо не детектируется.
EMotion Software GladOrSad	Компания – разработчик: Visual Recognition (Нидерланды)	Система создает 3D модель лица с выявлением 12 ключевых областей, таких как уголки глаза и уголки рта.	Распознает гнев, печаль, страх, удивление, отвращение и счастье. ПО не требовательно к вычислительной машине.	Неизвестны детали алгоритма реализации (недостатков не выявлено)
FaceSecurity	Компания – разработчик: Cognitec (Германия) Текущая версия: 4.6	Разработка предназначена для обработки уникальных баз и банков данных какой-либо категории людей, к примеру, сотрудников по работе. Система воплощает биометрическую идентификацию по эталону образцов из базы.	Кластерная конфигурация; сортировка списка изображений людей; доступны поисковые выборки	Не распознают плохо освещенные лица; распознавание только фронтально расположенных лиц

Помимо этого, удалось выявить общий алгоритм работы данных систем. Он состоит из четырех этапов: поиск лица на фотографии. (Наиболее популярный и эффективный алгоритм – это «Метод Виолы-Джонса»); создание 2D/3D-модели лица (контрольные точки, шаблон/деформируемая маска); выявление признаков с помощью алгоритма глубокого обучения; классификация признаков.

Больше всего внимания привлекает программа FaceReader [3], потому что она соответствует множеству критериев качества: имеет достаточно высокую скорость обнаружения признаков, базу данных, распознает лица при наклоне или повороте головы и других изменяющихся условиях.

Также компания-разработчик этого продукта не так давно выпустила программу, которая распознает эмоции детей от шести до двадцати четырех месяцев. Программа называется Baby FaceReader. Распознавание эмоций у детей, а тем более у младенцев – достаточно сложная задача.

Baby FaceReader была разработана как современная система автоматического определения выражений лица у младенцев, чтобы помочь решить вопросы психологии развития, связанные с аффективными расстройствами и расстройствами развития, такими как расстройство аутистического спектра и синдром дефицита внимания и гиперактивности [3]. Эта технология может помочь выявить данные нарушения еще в раннем возрасте.

На основе изученного материала мною был предложен собственный алгоритм распознавания эмоций. Схема алгоритма показана на рис. 1.

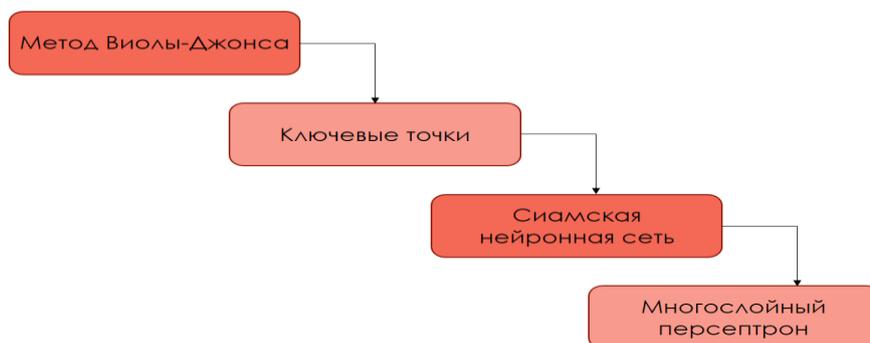


Рис. 1. Блок-схема алгоритма

Основная его особенность заключается в том, что объектом глубокого обучения является сиамская нейронная сеть, которая ранее не использовалась для решения задачи распознавания эмоций.

Использование данной сети позволит: сократить обучающую выборку, ускорить процесс обучения и сократить количество потребляемых ресурсов.

Принцип работы сиамской нейросети достаточно простой. Создаются две одинаковые нейронные сети. На входе каждая подсеть принимает одно из двух сравниваемых изображений. За входным слоем следует последовательность свёрточных слоев, они выявляют характерные признаки каждого из изображений. На выходе каждой из подсетей формируется вектор параметров (признаков), и в самом конце сеть сравнивает два выходных вектора, определяя, похожи два изображения или нет.

Предположим, что изначально данный алгоритм сможет распознавать только эмоцию радости на лице. Тогда для обучения потребуется три изображения: якорное изображение – шаблонный эталон эмоции радости, позитивное изображение, также показывающее радость и негативное изображение, но максимально похожее на эмоцию радости, в данном случае это нейтральное лицо. Далее алгоритм полностью повторяет работу, как и при распознавании лиц.

Триплет потерь минимизирует расстояние между якорным и позитивным изображениями, которые иллюстрируют одинаковую

эмоцию (эмоцию радости), а расстояние между якорным и негативным изображениями, наоборот, максимизирует (рис. 2). Подобным образом можно обучить интеллектуальную систему распознавать все семь базовых эмоций.



Рис.2. Обучение сиамской нейросети

Основные преимущества и недостатки алгоритма и дальнейшая перспектива развития представлены в табл. 2.

Таблица 2

Основные сведения об алгоритме

Преимущества	Недостатки	Перспектива развития
Небольшая обучающая выборка	Менее точный	Приложения для бизнес-компаний количеством сотрудников до 600 человек
	Сложная структура сети затрудняет анализ	
Распознавание эмоций индивидуально для каждого человека	Невозможность переобучения	Приложения для детей, страдающих расстройствами аутистического спектра
	Ограничение на количество распознаваемых лиц	

Разработка алгоритмов по распознаванию эмоций требует не только большого количества знаний в области эмоциональных вычислений, но и относительно продолжительного промежутка времени. Необходимо провести детальный анализ, чтобы учесть всевозможные особенности данного вида сетей и в дальнейшем избежать непредвиденных ситуаций.

Библиографические ссылки

1. The official web site «Atlas of Emotions». URL: <http://atlasofemotions.org/#continents/> (date of access: 19.05.2020)
2. The official web site of Noldus Information Technology. URL: <https://www.noldus.com/facereader> (date of access: 19.05.2020)