

СОВРЕМЕННЫЕ МЕТОДЫ РАСПОЗНАВАНИЯ ЭМОЦИЙ

В.А. Коровкин

Томский политехнический университет

alcasar@tpu.ru

Введение

В настоящее время распознавание эмоций является одной из самых обсуждаемых тем в области машинного обучения. Результаты исследований могут быть использованы для решения различных прикладных задач: маркетинговые исследования, распознавание состояния водителя, многочисленные системы аналитики, человеко-машинное взаимодействие, робототехника и т.д.

Поэтому целью нашей работы являлось рассмотрение различных подходов для решения задачи распознавания и классификации эмоций.

Эмоции играют важную роль, как и в жизни отдельного индивида, так и при межличностном взаимодействии в социуме. Они могут выражаться различными способами: мимикой (лицом), голосом, вегетативными реакциями, позой, двигательными реакциями и т.д.

Каждый человек проявляет их по-своему. Но согласно исследованию Пола Экмана часть эмоций может быть универсальной и понятной для человека независимо от его культурной принадлежности (гнев, страх, печаль, отвращение, презрение, удивление, радость).

Сегодня широко применяются три модели к категоризации эмоций: дискретная, многомерная и гибридная.

Дискретная модель основана на категоризации эмоций на основе естественного языка. Каждая эмоция связана с семантическим параметром, то есть обладает конкретным значением (или набором значений). Теория базовых эмоций (универсальных эмоций) является ярким примером данного подхода. Таким образом, например, ностальгия – это сочетание таких базовых эмоций как радость и печаль. Стоит отметить, что сегодня существующие базовых эмоций остается спорным. Различные исследователи выделяют разное количество и типы базовых эмоций. Профессор Экман выделяет 6 первичных эмоций (гнев, страх, отвращение, удивление, печаль и радость), психоэволюционная теория Роберта Плутчика рассматривает 8 базисных эмоций (одобрение, гнев, ожидание, отвращение, радость, страх, печаль, удивление), а Дж. Грей – 3 основных эмоции. По мнению, Мауэра следует говорить, только о 2 первичных эмоциях: боль и удовольствие.

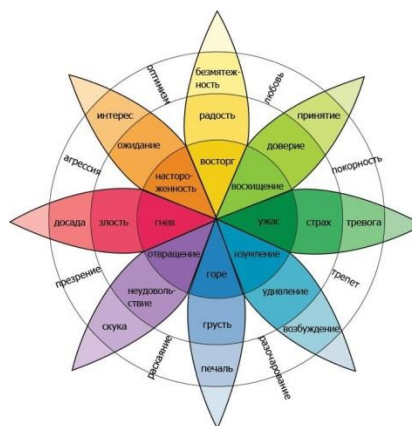


Рис. 1. Система диспозиций по Келлерману и Плутчику

Другой моделью для категоризации эмоций является многомерная. С помощью данной модели можно представить эмоции в координатном многомерном пространстве. Так как пространство является неразрывным не прерывным, то эмоции различаются по одному или нескольким параметрам (при этом они могут иметь одинаковую основу). Ярким примером описываемого подхода является модель Дж. Рассела. В ней водится двумерный базис, в котором каждая эмоция характеризуется знаком (valence) и интенсивностью (arousal). Данная модель в контексте решение задачи классификации является одной из самых востребованных.



Рис. 2. Двумерное пространство модели Дж. Рассела

Гибридные модели объединяют как дискретные, так и многомерные подходы. Отличным примером, иллюстрирующим данную модель, является «Песочные часы эмоций», которые были предложены Камбрией, Ливингстоном, Хуссейном. Каждое измерение характеризуется 6 уровнями силы, с которой выражены эмоции. Данные уровни обозначаются набором из 24 эмоций. Поэтому совершенно любая эмоция может рассматриваться как и фиксированное состояние, так и часть пространства, связанная с другими эмоциями нелинейными отношениями.

Классический подход к задаче классификации

Самый распространённый сегодня способ определения эмоций основан по классификации ключевых точек на изображении лица.

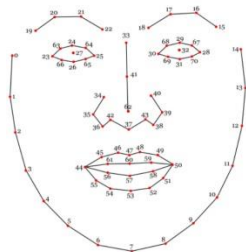


Рис. 3. Ключевые точки (facial landmarks)

Задача поиска координат ключевых точек (facial landmarks) сегодня отлично решается многочисленными алгоритмами – PDM, AAM, DPM, CNN, SML. Обычно размечают от 5 до 68 точек, которые привязываются к положению бровей, глаз, губ, носа, челюсти. Данная привязка позволяет захватить мимику объекта. Полученные координаты точек нормализуются и передаются непосредственно в классификатор. В роли классификатора может выступать метод опорных векторов (SVM) или, например, Random Forest.

Для улучшения качества получаемых результатов к координатам ключевых точек можно добавить визуальную информацию. С помощью алгоритмов и методов LBP, HOG, SIFT, LATCN и т.п. На рисунке 4 показан классический подход в классификации эмоций, где 1 – входные изображения, 2 – определение лица (например, методом Виолы-Джонса) и поиск координат ключевых точек, 3 – извлечение признаков, 4 – работа классификатора.

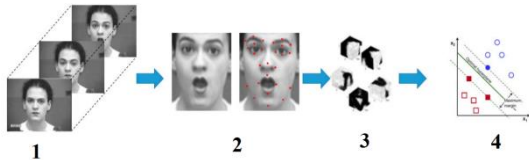


Рис. 3. Классический подход в распознавании эмоций

Классификация эмоций с применением глубокого обучения

При анализе визуальных данных лучшим выбором является использование глубоких сверточных нейронных сетей. Стоит отметить, что мгновенные снимки не совсем точно отражают текущую эмоцию. Поэтому для анализа необходимо использовать последовательность кадров. В качестве нейросетевого классификатора можно использовать сеть для распознавания лиц.

Решение задачи анализа последовательности кадров можно решить двумя способами:

1. использование рекуррентной сети (LSTM);
2. использование архитектуры 3D-CNN.

В первом варианте входные данные подаются на сверточную сеть. Высокоуровневые признаки,

полученные от CNN, которая классифицирует каждый отдельный кадр, затем передаются в рекуррентную сеть. В ней происходит обработка временной составляющей.

Во втором случае непосредственная подача последовательности кадров с некоторым шагом выполняется прямо на вход сверточной нейронной сети 3D-CNN. Данная сеть использует свертки с тремя степенями свободы, которые преобразует четырехмерный вход в трехмерные карты признаков.

В общем случае можно объединить оба метода. Результат представлен на рисунке 5.

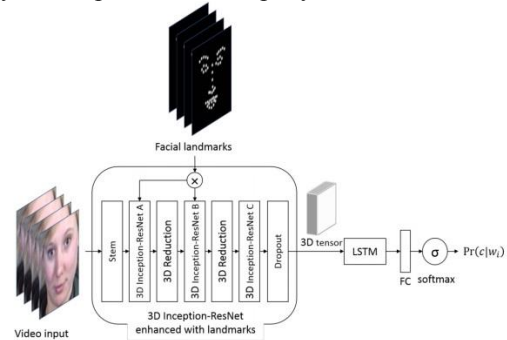


Рис. 4. Распознавание эмоций с применением deep learning

Заключение

Были рассмотрены различные подходы для классификации эмоций. Наиболее лучшие результаты показывают методы с применением сетей глубокого обучения (CNN). Стоит отметить, что для достижения наиболее точного результата необходимо использовать (комбинировать) несколько видов входных данных: например, аудио (речь) и видео (лицо и мимика).

Список использованных источников

1. Convolutional neural networks pertained on large face recognition dataset for emotion classification from video. [Электронный ресурс]. – URL: <https://arxiv.org/pdf/1711.04598.pdf> (дата обращения 22.11.2018).
2. A brief review of facial emotion recognition based on visual information. [Электронный ресурс]. – URL: <http://www.mdpi.com/1424-8220/18/2/401/pdf> (дата обращения 22.11.2018).
3. Audio-visual emotion recognition using deep transfer learning and multiple temporal models. [Электронный ресурс]. – URL: <https://dl.acm.org/citation.cfm?id=3143012> (дата обращения 22.11.2018).
4. Facial expression recognition using enhanced deep 3D convolutional neural networks. [Электронный ресурс]. – URL: <https://arxiv.org/pdf/1705.07871.pdf> (дата обращения 22.11.2018).

Работа поддержана грантом
РФФИ №18-08-00977 А.