

РАСПОЗНАВАНИЕ ЛЖИ ПРИ КОММУНИКАЦИЯХ ЛЮДЕЙ

Студент гр. 5ПИ Котова К. А.

Кафедра интеллектуальных систем

Научный руководитель – кандидат технических наук Садов В. С.

Белорусский государственный университет

Минск, Беларусь

Введение

Современные технологии машинного обучения и анализа данных оказывают значительное влияние на различные сферы жизни. Важным направлением исследований является разработка алгоритмов для определения эмоций человека на основе анализа мимики лица. Эти исследования важны для понимания его поведенческих мотиваций в различных областях деятельности, включая психологию, медицину и маркетинг и др..

В данной работе основное внимание уделяется разработке алгоритма решения задачи определения эмоций при анализе микровыражений на лице человека, во время произнесения лживой или правдивой информации, и исследованию его эффективности. Для решения этой задачи проведен анализ существующих методов и подходов к определению эмоций человека по его физиологическим данным, полученным по изображению лица.

Цель работы: разработать алгоритм распознавания эмоций и найти связь между мимикой лица и истинным отношением человека к сути излагаемой им информации, а также проанализировать полученные результаты.

Задачи: создать математическую модель, описывающую проявление лжи человека по эмоциям на его лице; обучить нейронную сеть для классификации эмоций на основе анализа микровыражений по ключевым точкам на изображении лица.

Общие понятия эмоций

Эмоции человека - комплексное явление, выражающееся в изменении настроения, поведения и мимики лица.

Согласно Полу Экману основными (базовыми) эмоциями человека, которые существенно изменяют выражение лица, являются: удивление, страх, отвращение, гнев, радость, печаль и нейтральное состояние [1].

Математическая модель проявления эмоций

Эмоции и проявление лжи - это сложные психофизиологические концепции, и их взаимосвязь может быть нелинейной и зависеть от множества факторов [2]. Каждая эмоция имеет свое процентное соотношение в общем эмоциональном состоянии человека. Обозначим условно базовые эмоции, следующим образом: A (злость), H (счастье), D (отвращение), F (страх), S (грусть), W (удивление), N (нейтральное состояние).

Тогда математическая модель, которая описывает проявление лжи по этим базовым эмоциям, может быть выражена следующим образом:

$$L = A * w_1 + H * w_2 + D * w_3 + F * w_4 + S * w_5 + W * w_6 + N * w_7, \quad (1)$$

где L – проявление лжи, $w_1, w_2, w_3, w_4, w_5, w_6, w_7$ – веса которые определяют вклад каждой эмоции в выражение лжи. Эти веса могут быть определены через эмпирические исследования с помощью нейронной сети.

Обучение нейронной сети

Для обучения нейронной сети с целью распознавания эмоций использовался готовый набор данных FER2013+balanced [3]. Датасет вручную очищен и отфильтрован от «неинформативных» изображений, а также сбалансирован, то есть для всех категорий эмоций имеется примерно одинаковое количество изображений.

Для распознавания микровыражений использовался датасет, который был составлен из видео с YouTube, на которых различные люди шоу-бизнеса проходят полиграф и отвечают на вопросы. Моменты, где полиграф показал ложь на ответы, были вырезаны и добавлены в набор данных. Таким образом,

было собрано 65 видео разных людей, где полиграф показал ложь, и 156 видео – правду. Далее вручную были отобраны изображения с микровыражениями лица, эмоции на которых отличаются от эмоций на кадрах ранее или после.

Датасет FER2013+balanced был разделен на тестовую и тренировочную выборки: для тестовой выборки отобраны 5772 экземпляра, а для тренировочной - 23251.

Распознавание эмоций по ключевым точкам лица осуществлялось с использованием сверточной нейронной сети. Для извлечения ключевых точек лица использовалась Python библиотека MediaPipe.

В следствие обучения нейронной сети на основе ключевых точек лица получены следующие результаты: точность обучения для тренировочной выборки – 70.15%, точность обучения для тестовой выборки – 69.43%, потери на тренировочной выборке – 0.819, потери на тестовой выборке – 0.854. Графики точности и потерь изображены на рис. 1.

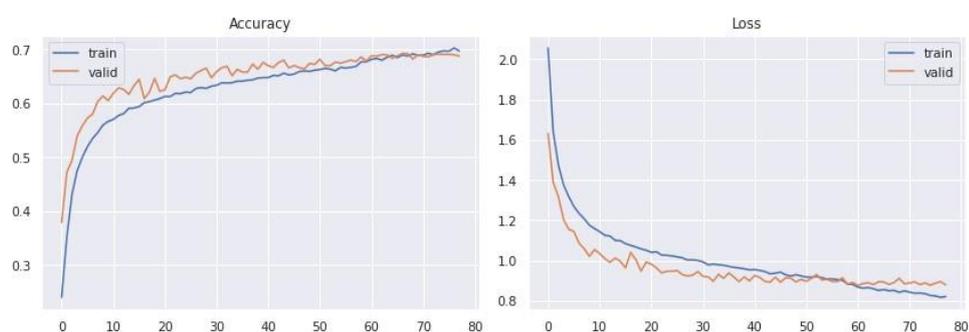


Рисунок 1. Графики точности и потерь на основе результатов обучения нейронной сети по ключевым точкам лица

Матрица ошибок изображена на рис. 2. Она позволяет визуализировать, насколько хорошо модель справляется с задачей классификации, путем сравнения предсказанных классов с фактическими классами. По матрице ошибок заметно, что лучше всего распознаются эмоции радость, удивления и нейтральности. Также заметно, что эмоция грусти может классификатором ошибочно приниматься за нейтральную. Вероятно по причине того, что эти

эмоции похожи внешне, и по ключевым точкам могут быть плохо различимы (рисунок 2).

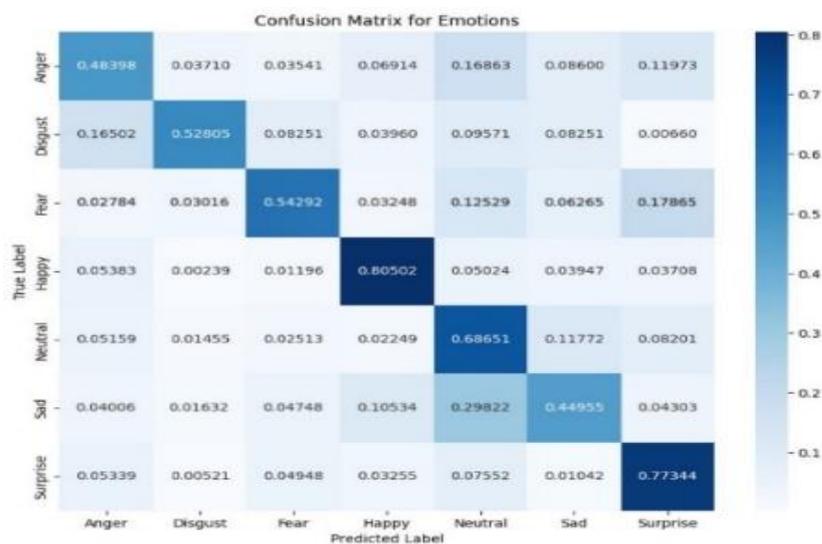


Рисунок 2. Матрица ошибок, которая отображает точность прогнозирования модели (классификатора) в отношении 7 классов эмоций на основе ключевых точек

Полученные результаты

Проанализировав кадры с выражением лжи и правды с помощью обученной модели, была создана статистика: минимальное, максимальное и среднее значения процентного соотношения для каждой эмоции.

Результаты анализа представлены в таблице 1. Исходя из таблицы 1, микровыражения на лице лгущего человека чаще всего являются следствием выражения эмоций радости, нейтральности и грусти, а для человека, говорящего правду, так же более характерны эмоции радости и нейтральности. Для человека, который говорит правду, можно заметить, что эмоции злости, отвращения и страха не особо свойственны.

Таблица 1. Статистика предсказанных эмоций моделью на основе ключевых точек лица

Эмоции/значения	Минимальное, %	Максимальное, %	Среднее, %
Ложь			
Злость	1.80210	54.21466	4.13143
Отвращение	0.16751	10.78209	0.81886
Страх	0.24844	24.67541	1.04491
Радость	1.71094	89.84230	28.32151
Нейтральность	0.07402	35.97861	22.58203
Грусть	4.16396	57.62893	18.80372
Удивление	0.00033	40.19944	1.93037
Правда			
Злость	0.01584	9.09733	2.39784
Отвращение	0.00044	4.41773	0.30386
Страх	0.01319	7.70059	1.84514
Радость	4.34429	99.7719	50.94182
Нейтральность	0.03365	69.45223	27.11486
Грусть	0.08124	36.32472	9.89575
Удивление	0.06451	32.41095	7.50070

На рисунке 3 изображены наиболее характерные результаты оценки эмоционального фона по фотографии лица, полученные по разработанной методике.



Рисунок 3. Результаты предсказания эмоционального фона нейросетевой моделью, обученной по ключевым точкам лица

Заключение

В ходе данной работы был рассмотрен алгоритм классификации эмоций на основе ключевых точек на изображении лица. Точность классификации - 70%. Точность обуславливается правильным выбором модели для обучения (свёрточной нейронной сети); разделением данных на тренировочную, тестовую и валидационную выборки; правильным подбором слоев нейросети;

а также правильным подбором ее параметров. Однако модель, созданная для обучения на ключевых точках, может быть еще эффективнее при корректном выборе набора данных: изображения должны быть лучшего качества.

Был проведен анализ микровыражений, отобранных из кадров видео, где различные люди лгут и говорят правдивую информацию. Было выявлено, что в видео, где люди лгут, содержится больше микровыражений, чем в видео, где говорят правду. А также было обнаружено, что интенсивность проявления таких эмоций, как радость, грусть, злость и страх, выше у «лживых», чем у «правдивых» микровыражений, что может быть связано с попытками скрыть истинную эмоцию либо страхом разоблачения. Однако стоит учитывать, что проявление эмоций зависит от человека и его мимических особенностей, что усложняет задачу классификации.

Литература

1. Экман П., Узнай лжеца по выражению лица. / Пол Экман. - СПб.: Питер, 2019. - 272 с.
2. Экман П., Психология лжи. Обмани меня, если сможешь. / Пол Экман. - СПб.: Питер, 2021. - 384 с.
3. FER2013+balanced [Электронный ресурс]. - Режим доступа: <https://www.kaggle.com/datasets/prilia/fer2013pluscleanedaugmballanced1>. Дата доступа: 12.10.2023.