

РАСПОЗНАВАНИЕ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ - «СЧАСТЛИВЫЙ ЧЕЛОВЕК»

Чжуню Го, В.С. Садов

Белорусский государственный университет, Минск, Беларусь

E-mail: guozhongyou1@gmail.com

В данной работе исследованы методы использования нейронных сетей для распознавания эмоции «счастливый человек» по изображению лица. Мы использовали набор данных FER-2013, который содержит 28709 примеров для обучения нейронной сети и 3589 примеров для ее тестирования. Обучающая и тестовая выборки включают семь базовых эмоциональных категорий (злость, отвращение, страх, радость, грусть, удивление, нейтральность). Использована сверточная нейронная сеть (CNN) для извлечения особенностей лица, Точности распознавания эмоций составила 69%.

Ключевые слова: распознавание эмоций; базовые эмоции; глубокое обучение; нейронная сеть; модель.

ВВЕДЕНИЕ

Распознавание эмоций по лицу – это область, использующая технологии нейронных сетей для анализа и распознавания выражений лица. С развитием искусственного интеллекта нейронные сети достигли значительных успехов в области компьютерного зрения, и распознавание эмоций по лицу стало распространенным направлением исследований.

Цель распознавания эмоций по лицу заключается в точном определении эмоционального состояния человека путем анализа физиологических характеристик лица, и определения по ним качественного и количественного состава отображаемых базовых эмоций.

Полученные данные позволяют спрогнозировать поведенческую линию человека, оценить его способность к тем или иным действиям. Эта технология имеет широкое применение во многих областях, включая взаимодействие человека с компьютером, эмоциональные вычисления, маркетинговые исследования и т. д.

ИСПОЛЬЗУЕМАЯ БАЗА ДАННЫХ

В данном исследовании использовалась база данных FER-2013 для анализа изображений, посредством сверточной нейронной сети. Размер изображений в этом наборе данных составляет 48x48 пикселей, и они представлены только черно-белыми изображениями с одним цветовым каналом. Эта база данных содержит 28709 примеров в обучающем наборе и 3589 [1] примеров в тестовом наборе. Оба набора данных включают изображения лиц разных людей, выражающих 7 базовых эмоций (гнев, отвращение, страх, радость, грусть, удивление, нейтральность) [1].

В табл. 1 приведен количественный и качественный состав базы данных FER-2013, а на рис. 1 примеры фотографий с выражениями основных базовых эмоций.

Таблица 1

Наполнение базы (выборки данных) FER-2013

Базовые эмоции	Train	Test
Радость	7215	1774
Нейтральность	4965	1223
Грусть	4830	1247
Страх	4097	1024
Злость	3995	958
Удивление	3171	831
Отвращение	436	111

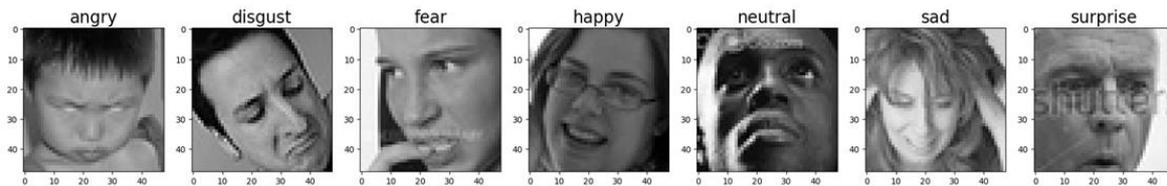


Рис. 1. Примеры образцов из базы данных FER-2013

МЕТОД ИССЛЕДОВАНИЙ

В данном исследовании использовалась нейронная сеть архитектуры ResNet50V2 для анализа данных. Основная идея ResNet50V2 [2] заключается в использовании блоков с остаточными соединениями (Residual Connection) для решения проблемы затухания и всплеска градиента в процессе обучения глубоких нейронных сетей. В традиционных сверточных нейронных сетях с увеличением количества слоев градиент постепенно уменьшается, что делает обучение сложным. Однако благодаря остаточным соединениям сеть позволяет градиенту быстрее распространяться, что эффективно решает проблемы его затухания и всплеска.

Архитектура сети ResNet50V2 включает в себя сверточные слои, слой пакетной нормализации (Batch Normalization), функции активации и блоки с остаточными соединениями. Блок с остаточными соединениями является ключевым компонентом ResNet и состоит из нескольких сверточных слоев и остаточного соединения. Внутри каждого блока обычно находятся два сверточных слоя размером 3x3 и одно остаточное соедине-

ние, а также слои пакетной нормализации и функции активации для улучшения выразительности сети. На рис. 2 и в табл. 2 приведены графики функции потерь и точности при обучении сети на наборе данных из базы FER-2013

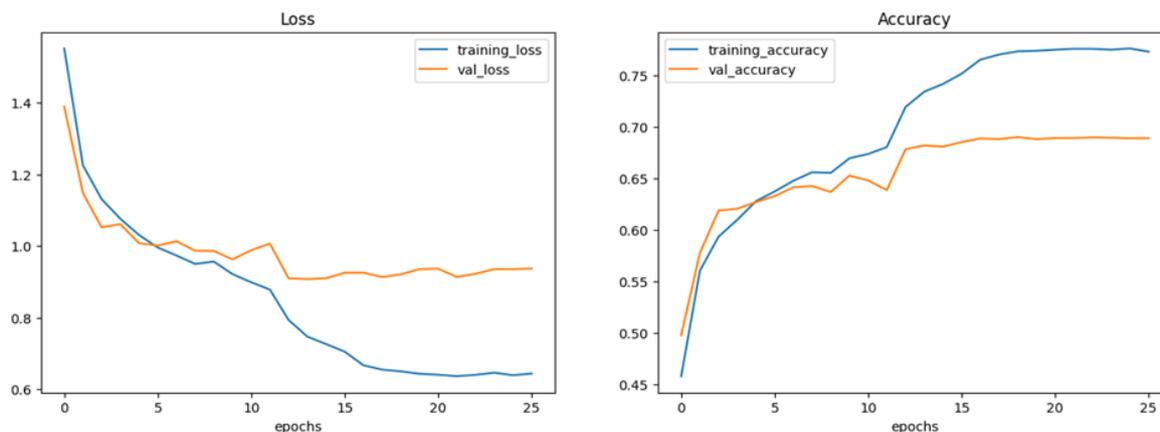


Рис. 2. Графики функции потерь и точности при обучении сети

Test Loss: 0,91942; Test Accuracy: 69,03%

Таблица 2

Результаты оценки базовых эмоций из базы данных FER-2013

Эмоции	Precision	Recall	f1-score	Support
Злость	0,62	0,64	0,63	958
Отвращение	0,78	0,56	0,65	111
Страх	0,56	0,45	0,5	1024
Радость	0,88	0,88	0,88	1774
Нейтральность	0,62	0,71	0,67	1233
Грусть	0,56	0,57	0,57	1247
Удивление	0,8	0,79	0,79	831
accuracy			0,69	7178
macro avg	0,69	0,66	0,67	7178
weighted avg	0,69	0,69	0,69	7178

Для идентификации эмоционального состояния «счастливый человек» по распознанным базовым эмоциям будем использовать следующую модель [3]:

Счастье = $(K_{радость} + K_{удивлен.} + K_{нейтральн.}) / (K_{грусть} + K_{страх} + K_{отвращен.} + K_{гнев.}) \geq 3$,
 где K – процентное содержание соответствующей базовой эмоции.

Результаты анализа фотографий, приведенных на рис. 3, на наличие среди них изображения лиц с выражением счастья приведен в табл. 3.

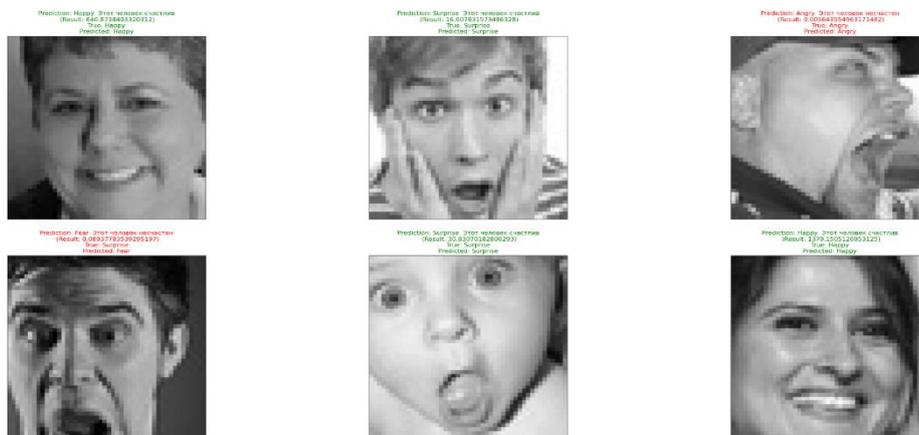


Рис. 3. Тестовая выборка с изображениями лиц счастливого человека (с пояснениями зеленого цвета)

Таблица 3

Результаты анализа тестовой выборка на наличие лиц счастливых людей

№ рис.	Счастливы?	Результат анализа по модели	Оценка эмоции в базе данных	Оценка эмоции разработанной системой
1	да	640,87	Радость	Радость
2	да	16,00	Удивление	Удивление
3	нет	0,005	Злость	Злость
4	нет	0,089	Удивление	Страх
5	да	30,83	Удивление	Удивление
6	да	1379,15	Радость	Радость

ВЫВОДЫ

В нашем исследовании мы использовали ResNet50V2 для распознавания эмоций на лицах, и с помощью оптимизации дизайна сети, обучающей стратегии и использования набора данных FER-2013, был достигнут удовлетворительный результат распознавания – 69,03%.

Эмоциональное состояние «счастливый человек» можно определить по совокупности выявленных базовых эмоций в соответствии с моделью [3].

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Kaggle. FER-2013 [Электронный ресурс]. URL: <https://www.kaggle.com/datasets/msambare/fer2013>.
2. He, K., Zhang, X., Ren, S., & Sun, J. Identity Mappings in Deep Residual Networks. In European Conference on Computer Vision, 2016.
3. Фредриксен, Б. Сила позитивных эмоций: Новейшие исследования раскрывают соотношение 3:1, которое изменит вашу жизнь/ Нью-Йорк MJF: Crown, 2009.