

УДК 004.89

РАСПОЗНАВАНИЕ ЭМОЦИЙ С ПОМОЩЬЮ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ОЦЕНКИ ПСИХОЛОГИЧЕСКОГО СОСТОЯНИЯ АВИАТОРОВ

© Дворянчиков Д.И.

*Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация*

e-mail: Didvoryanchikov@yandex.ru

Научный руководитель: Солдатова О.П., канд. техн. наук, доцент
*Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация*

Ключевые слова: распознавание эмоций, аффективные вычисления, свёрточные нейронные сети, алгоритм обучения нейронной сети.

Стресс в авиации и космонавтике – обычное явление, возникающее в силу высокой степени опасности профессиональной деятельности данной группы людей. Исследование психоэмоционального состояния людей экстремальных профессий является актуальной задачей. В последнее время для решения данной задачи активно используются методы искусственного интеллекта, так называемых аффективные вычисления. Основное решение данной проблемы достигается с помощью свёрточных нейронных сетей, обученных на аннотированных датасетах лицевых экспрессий. От выбора алгоритма обучения и архитектуры модели напрямую зависит точность классификации аффективных состояний.

В работе использовались модели с архитектурой ResNet. Благодаря своей уникальной структуре ResNet стала основой для множества современных моделей в компьютерном зрении, распознавании изображений, классификации и других задачах искусственного интеллекта [1].

Модель ResNet стала победителем соревнования ILSVRC-2015, ансамбль из вариантов которой показал частоту ошибок 3.6 %, что ниже, чем оцениваемая частота ошибок человека (5.1 %) [2].

Архитектура ResNet произвела революцию в мире глубокого обучения, решив одну из главных проблем – исчезающие градиенты в очень глубоких нейронных сетях. Основные компоненты ResNet:

- Остаточные блоки. Они помогают нейросети передавать информацию через несколько слоев, минуя некоторые из них. ResNet не соединяет слои последовательно, а использует связи skip connections.

- Свёрточные слои и функции активации. Каждый остаточный блок включает свёрточные слои с ядрами 3*3 и функции активации ReLU. Данные элементы добавляют нелинейность, позволяя модели улавливать сложные закономерности в данных.

- Нормализация и оптимизация. Для стабилизации обучения ResNet использует Batch Normalization. Данный метод нормализует выходы каждого слоя, ускоряет обучение и улучшает стабильность нейросети.

– Масштабируемость. ResNet легко адаптируется под разные задачи. Например, ResNet18 содержит 18 слоев, а ResNet50 – 50. Это позволяет выбирать модель в зависимости от сложности задачи и доступных вычислительных ресурсов.

Для обучения свёрточной сети в данной работе были выбраны два алгоритма обучения: AdamW и SGD с моментом.

AdamW – это модификация алгоритма Adam, которая отделяет регуляризацию весов от процесса обновления градиента [3].

SGD с моментом – модификация алгоритма SGD, которая добавляет в правило обновления параметр momentum.

Для обучения моделей ResNet (ResNet18 и ResNet50) было использовано два датасета: FER-2013 и RAF-DB.

Датасет FER-2013 представляют собой изображения лиц в оттенках серого размером 48x48 пикселей с аннотацией семи базовых эмоций. Обучающий набор состоит из 28 709 примеров, а общедоступный тестовый набор из 3589 примеров.

Датасет RAF-DB содержит 15 000 цветных изображений лиц в разрешении 100x100 пикселей с аннотацией семи базовых эмоций. Изображения в этой базе данных сильно различаются по возрасту, полу и этнической принадлежности субъектов, позам головы, условиям освещения, окклюзиям, операциям постобработки и так далее.

Каждый датасет, модель и архитектура приводят к своему собственному результату. Цель данной работы – сравнить эффективность архитектур и алгоритмов обучения для точного распознавания эмоций. Чтобы оценить эффективность обучения моделей, мы будем учитывать два ключевых показателя:

– Ассурасу – это доля правильно распознанных эмоций на изображениях от общего количества.

– Loss (CrossEntropyLoss) – это показатель, который демонстрирует, насколько сильно отличается реальный ответ сети от эталонного ответа.

Результаты экспериментов с разделением на количество эпох для каждого датасета и модели предоставлены в таблицах 1–2.

Таблица 1 – Результаты экспериментов для датасета FER-2013

Количество эпох обучения	10		20		30	
Алгоритмы обучения	accuracy	loss	accuracy	loss	accuracy	loss
AdamW (ResNet18)	0.5779	1.1457	0.5978	1.5107	0.6030	1.9049
SGD с моментом (ResNet18)	0.5681	1.1523	0.5925	1.3796	0.6028	1.7479
AdamW (ResNet50)	0.5918	1.1097	0.6091	1.5039	0.6082	1.9080
SGD с моментом (ResNet50)	0.5748	1.2174	0.6006	1.6964	0.6236	1.9722

Таблица 2 – Результаты экспериментов для датасета RAF-DB

Количество эпох обучения	10		20		30	
Алгоритмы обучения	accuracy	loss	accuracy	loss	accuracy	loss
AdamW (ResNet18)	0.7907	0.6956	0.8171	0.7120	0.8152	0.9740
SGD с моментом (ResNet18)	0.7458	0.7394	0.7748	0.7735	0.7780	0.8610
AdamW (ResNet50)	0.7813	0.7045	0.8044	0.8074	0.8064	0.8850
SGD с моментом (ResNet50)	0.7663	0.7161	0.7790	0.9561	0.7872	0.9648

По результатам экспериментов, следует, что лучшим вариантом будет использование датасета RAF-DB с моделью ResNet18 и алгоритмом обучения AdamW на 20 эпохах, дающую точность в 81,7 % и меньшим loss.

Библиографический список

1. CPA.LIVE, Как работает ResNet: разбор архитектуры и обучение на реальных примерах. – URL: <https://cpa.live/nejroseti/chto-takoe-resnet>.
2. Машинное обучение, Модель ResNet. – URL: <https://deepmachinelearning.ru/docs/Neural-networks/Convolutional-architectures/ResNet>.
3. Habr. Методы оптимизации в машинном и глубоком обучении. От простого к сложному. – URL: <https://habr.com/ru/articles/813221>.