

НЕЙРОСЕТЕВАЯ МОДЕЛЬ РАСПОЗНАВАНИЯ ВЫРАЖЕНИЯ ЛИЦА ДЛЯ ОПРЕДЕЛЕНИЯ УСТАЛОСТИ ЧЕЛОВЕКА

Б. Курбанов, А.С. Катасёв, И.М. Шаяхметов, Б.Р. Зиннуров

Казанский национальный исследовательский технический университет
им. А.Н. Туполева-КАИ
Российская Федерация, 420111, г. Казань, ул. К. Маркса, д. 10

Аннотация. Статья посвящена использованию сверточной нейросетевой модели для определения усталости человека по выражению лица. Для этого используется архитектура сверточной нейронной сети ResNet50, которая обучается на изображениях и находит признаки, указывающие на усталость и бодрость человека. Подготовленный набор данных включает в себя 6000 изображений, из которых 3000 изображений соответствуют признакам усталого человека (на этих изображениях человек зеваает или закрывает глаза) и 3000 изображений – признакам усталого человека. При создании нейросетевой модели ResNet50 использован язык программирования Python, платформа для разработки Jupyter Notebook и среда разработки Anaconda3. Нейросетевая модель обучалась в течение 100 эпох, при этом каждый обучающий пример состоял из 32 элементов. Благодаря использованию алгоритма оптимизации Adam удалось обучить нейронную сеть для корректной классификации изображений на два класса: усталость и бодрость. При обучении нейросетевая модель достигла уровня точности классификации в 95%. В результате расчета значений метрик качества классификации Recall и Precision на тестовой выборке данных, используя матрицу ошибок, удалось получить высокие результаты. Для класса «усталый» значение метрики Recall составило 0.9426, а значение метрики Precision – 0.9587. Для класса «бодрый» значение метрики Recall составило 0.9576, а значение метрики Precision – 0.9417. Эти показатели свидетельствуют об адекватности построенной модели и возможности ее практического использования. В итоге можно сделать вывод, что проведенное исследование продемонстрировало успешное применение нейросетевой модели распознавания выражения лица для определения усталости человека с высокой степенью точности.

Ключевые слова: нейросетевая модель, сверточная нейронная сеть, определение усталости человека, выражение лица, ResNet50.

Введение

В настоящее время практически каждый человек в своей повседневной деятельности сталкивается с решением сложных задач, связанных с обработкой большого количества информации или с выполнением большого количества операций. Это быстро приводит к усталости человека, которая может негативно повлиять на качество выполняемой работы, а также на его работоспособность. Усталость также может возникать из-за недостатка сна, длительной физической активности, эмоционального стресса и других негативных факторов. В любом случае усталость оказывает негативное влияние на работоспособность человека и безопасность выполняемых им операций.

Снижение концентрации внимания, как следствие усталости, может привести, например, к ошибкам при управлении автомобилем, что, в свою очередь, может повлечь за собой серьезные последствия [1]. Поэтому важно определять состояние усталости человека, в том числе без его участия в этом процессе. Также определение усталости является актуальной задачей для обеспечения транспортной безопасности. Для ее решения целесообразно использовать интеллектуальные системы, основанные на методах классификации изображений, например, на сверточных нейронных сетях. Новизна предложенного в работе решения заключается в выборе и обогащении данных из открытого источника, а также в

обучении сверточной нейронной сети известной архитектуры на этих данных. В данной статье описана разработанная сверточная нейросетевая модель, позволяющая определять состояние усталости человека по выражению лица. Рассмотрим этапы построения и тестирования нейросетевой модели, а также достигнутые результаты моделирования.

Выбор и описание модели сверточной нейронной сети для определения усталости человека по выражению лица

Как известно [2], любая сверточная нейронная сеть состоит из нескольких слоев, каждый из которых выполняет определенную функцию. Сверточные нейронные сети имеют ряд преимуществ перед другими моделями машинного обучения. В частности, они позволяют исследовать изображения объектов во всей глубине, а не только поверхностно. В то же время, сверточная нейронная сеть обладает способностью запоминать и выделять наиболее характерные особенности объектов, что способствует повышению точности идентификации и классификации. Благодаря этому, сверточные нейронные сети могут более точно определять характеристики объектов и их параметры, учитывая контекст и взаимосвязи между ними на изображении [3-5]. Сверточная нейронная сеть обладает способностью различать черты лица, такие как глаза, нос и рот, и на основе этого делать выводы о классификации изображений. В задаче распознавания выражения лица для определения усталости человека сверточные нейронные сети могут быть использованы для автоматического извлечения признаков из изображения лица, что может повысить точность классификации. Помимо этого, сверточные нейронные сети могут обучаться на больших объемах данных, что может повысить качество модели [6].

После проведения сравнительного анализа моделей сверточных нейронных сетей для решения задачи определения усталости человека по выражению лица была выбрана модель ResNet50. Особенностью данной модели является наличие соединений быстрого доступа «shortcut connections», которые позволяют пропускать данные через слои без изменений, что помогает бороться с проблемой затухания градиентов. Благодаря этой особенности модель способна обрабатывать большой объем данных и достигать высокой точности при классификации изображений. Сеть ResNet50 состоит из 50 слоев, которые используются для классификации изображений. Информация о структуре и слоях этой нейронной сети представлена на рис.1.

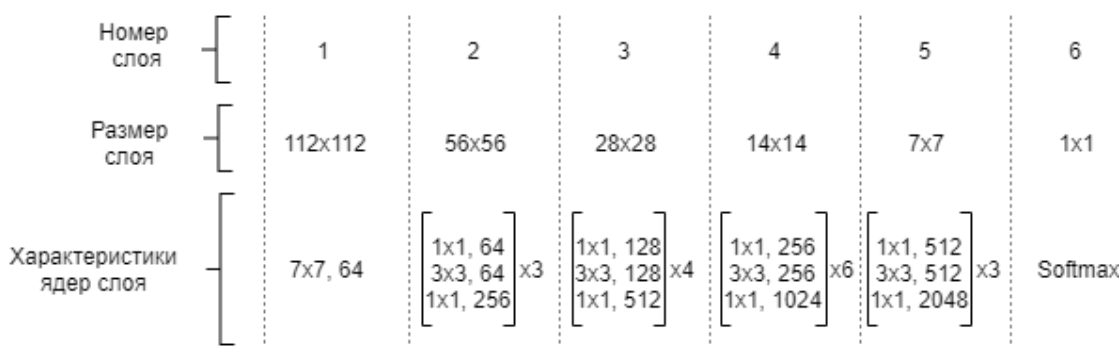


Рис.1. Слои архитектуры ResNet50

Рассмотрим характеристики слоев сети:

- 1) первый слой свертки имеет размер 112×112, а размер ядер слоя 7×7, 64;
- 2) второй слой свертки имеет размер 56×56, а размер ядер слоя 1×1, 64, 3×3, 64 и 1×1, 256 (слои повторяются 3 раза, образуя 9 слоев);
- 3) третий слой свертки имеет размер 28×28, а размер ядер слоя 1×1, 128, 3×3, 128 и 1×1, 512 (слои повторяются 4 раза, образуя 12 слоев);

4) четвертый слой свертки имеет размер 14×14 , а размер ядер слоя 1×1 , 256 , 3×3 , 256 и 1×1 , 1024 (слои повторяются 6 раз, образуя 18 слоев);

5) пятый слой свертки имеет размер 7×7 , а размер ядер слоя 1×1 , 512 , 3×3 , 512 и 1×1 , 2048 (слои повторяются 3 раза, образуя 9 слоев);

б) выходной слой имеет размер 1×1 с функцией активации softmax.

Таким образом, сверточная нейронная сеть ResNet50 состоит из 50 слоев, что и определяет ее название.

Далее рассмотрим описание этапа поиска и подготовки исходных данных для анализа и построения сверточной нейросетевой модели.

Подготовка исходных данных для построения нейросетевой модели

Для того, чтобы построить модель нейронной сети, которая будет распознавать выражение лица для определения усталости человека, необходимы качественные исходные данные. В работе в качестве источника данных выбрана онлайн платформа Kaggle [7]. На сайте Kaggle представлена обширная база данных, включающая изображения лиц людей в различных функциональных состояниях, включая усталость. Многие наборы данных предоставляют маркированные изображения, которые могут быть использованы для обучения сверточной нейронной сети.

Для решения задачи определения усталости человека были отобраны следующие наборы данных: «Drowsiness_dataset», «fatigue_detection», «Drowsiness Prediction Dataset». Отобранные наборы данных содержат изображения лиц с признаками усталости: закрытые глаза и зевота. Кроме того, использован собственный набор данных, содержащий фотографии людей различных возрастных категорий, этнических групп и полов. Важно подчеркнуть, что собственный набор данных дополняет отобранные наборы данных и вносит разнообразие, которое может положительным образом повлиять на эффективность модели сверточной нейронной сети.

Проведение предварительной обработки данных является одним из наиболее важных этапов в процессе обучения модели нейронной сети. Данный процесс включает в себя ряд последовательных действий, которые позволяют подготовить данные для последующего анализа и получения более точных результатов [8].

Предварительная обработка данных включает в себя устранение шума и выбросов, а также приведение изображений к единому размеру и формату (например, JPEG).

Далее необходимо разделить данные на две выборки: обучающую и тестовую. В процессе обучения модели будет использоваться обучающая выборка, а тестовая – для оценки ее точности и эффективности. В процентном соотношении 80% исходного набора данных будет входить в обучающую выборку, а 20% – в тестовую. Также важно учитывать случайность отбора, чтобы избежать переобучения модели.

В результате предварительной обработки данных создан набор изображений, состоящий из двух классов: изображения уставших и бодрых людей. Набор содержит 3000 изображений, которые указывают на усталость человека (он зевает или закрывает глаза). Кроме того, в него входят 3000 изображений бодрого человека. В результате, общий объем данных составил 6000 подготовленных к анализу изображений.

Построение нейросетевой модели

Сверточная нейронная сеть ResNet50 построена с помощью языка программирования Python, который используется на платформе Anaconda3 и в среде разработки Jupyter Notebook. Построение нейросетевой модели осуществлялось в течение 100 эпох, при этом размер обучающих примеров достигал 32 элемента за одну итерацию. В процессе построения модели использовались веса ImageNet и был использован алгоритм Adam, благодаря

которому удалось обучить сверточную нейронную сеть так, чтобы она правильно разделяла изображения на два класса – «усталый человек» и «человек, находящийся в состоянии бодрствования». В результате построения модель достигла высокого уровня точности, который позволяет классифицировать выражение лица для определения усталости человека. Стоит отметить, что в процессе построения модели были учтены многие аспекты, такие как оптимизация скорости обучения и подбор оптимальных параметров.

Приведенный на рисунке 2 график демонстрирует процесс обучения сверточной нейросетевой модели.

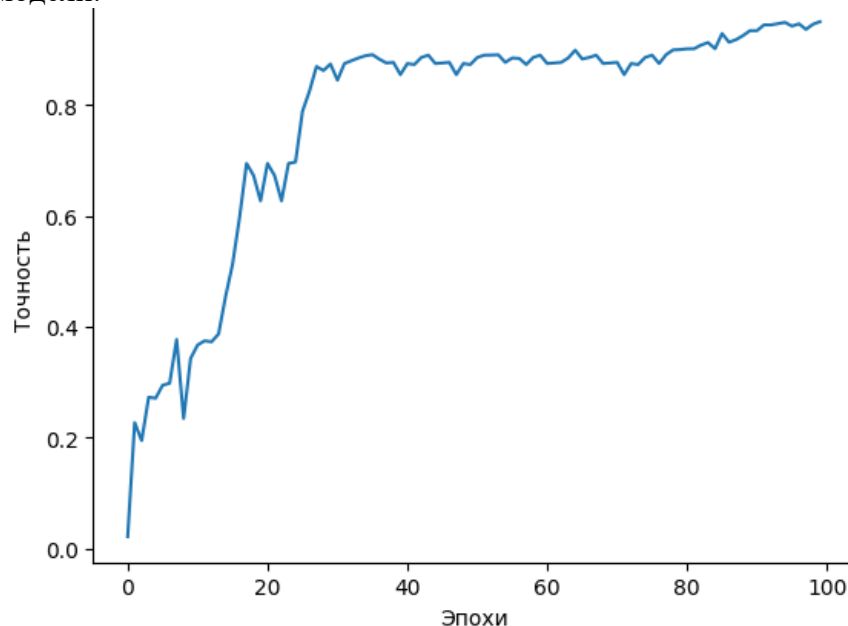


Рис. 2. График процесса обучения сверточной нейросетевой модели

Как показано на рисунке, по оси абсцисс обозначены эпохи обучения сверточной нейросети, а по оси ординат – точность классификации. В процессе обучения модели была достигнута точность классификации на уровне 95%, что является высоким показателем для сверточных нейронных сетей.

Тестирование оценка эффективности нейросетевой модели

На рис.3 представлены примеры использования построенной сверточной нейросетевой модели для определения усталости человека по выражению лица.

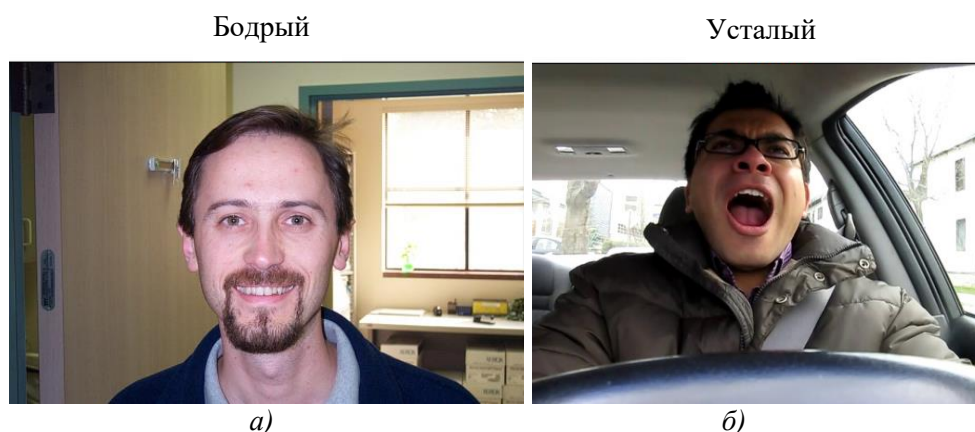


Рис. 3. Примеры использования построенной нейросетевой модели: а) распознавание состояния бодрости, б) распознавание состояния усталости

Согласно результатам модели, на рис.3а человек классифицируется как бодрый, а на рис.3б - как усталый. В обоих случаях классификация была проведена верно.

Для окончательной оценки эффективности построенной модели использован тестовый набор данных. В табл.1 представлены результаты тестирования модели по метрике «точность классификации».

Таблица 1. Результаты тестирования нейросетевой модели

Класс	Точность классификации на тестовой выборке данных
Бодрый	94,17%
Усталый	95,83%
Среднее значение	95%

Из таблицы видно, что средняя точность модели составила 95%. Следовательно, построенная модель является адекватной и может быть эффективно использована для решения поставленной задачи.

Для дополнительной оценки адекватности нейросетевой модели построена матрица ошибок классификации, представленная на рис. 4.

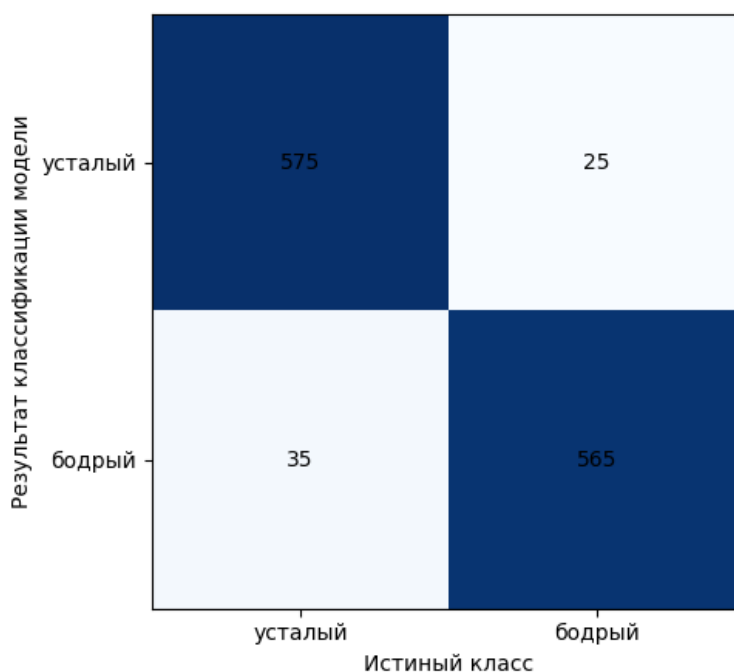


Рис.4. Матрица ошибок классификации

По полученным данным вычислены значения метрик качества классификации:

- Recall = $575 / (575 + 35) = 0.9426$ (для класса усталый);
- Recall = $565 / (565 + 25) = 0.9576$ (для класса бодрый);
- Precision = $575 / (575 + 25) = 0.9587$ (для класса усталый);
- Precision = $565 / (565 + 35) = 0.9417$ (для класса бодрый).

Полученные значения метрик качества классификации также указывают на адекватность построенной модели, которая правильно классифицирует большинство объектов.

Заключение

Результаты проведенных исследований показали, что построенная модель является адекватной и может быть применена в реальных условиях, например, для определения усталости водителей. Дальнейшая разработка интеллектуальной системы, основанной на нейросетевой модели, потребует улучшения алгоритмов обработки изображений, добавления возможности анализа видеоизображений, что положительно скажется на точности определения усталости. Это позволит расширить спектр применимости системы и сделать ее эффективной для применения в различных предметных областях.

Список литературы

1. Маркин Э.В. Повышение уровня адаптации и усталости организма человека к различным условиям внешней среды / Э.В. Маркин, И.С. Мысишин // Инновации в образовании: Материалы Международной научно-практической конференции. – Орел, 2009. – С. 284-288.
2. Аведьян Э.Д. Сравнительный анализ структур полносвязных и сверточных нейронных сетей и их алгоритмов обучения / Э.Д. Аведьян, А.И. Галушкин, С.А. Селиванов // Информатизация и связь. – 2017. – № 1. – С. 18-30.
3. Черненький И.М. Сегментация почечных структур по изображениям контрастной компьютерной томографии с помощью сверточной нейронной сети / И.М. Черненький, М.М.Черненький, Д.Н. Фиев, Е.С. Сирота // Сеченовский вестник. – 2023. – Т. 14, № 1. – С. 39-49.
4. Хусаинов Р.М. Нейросетевая модель и программный комплекс распознавания объектов дорожной инфраструктуры / Р.М. Хусаинов, Н.Г. Талипов, А.С. Катасёв // Информационные технологии. – 2023. – Т. 29, № 9. – С. 484-491.
5. Хусаинов Р.М. Нейросетевая технология анализа транспортных потоков в автоматизированных системах управления дорожным движением / Р.М. Хусаинов, Н.Г. Талипов, А.С. Катасёв, Д.В. Шалаева // Программная инженерия. – 2023. – Т. 14, № 10. – С. 513-519.
6. Немков Р.М. Динамическое изменение воспринимающих свойств сверточных нейронных сетей как обучение с шумом и его влияние на обобщающую способность / Р.М. Немков, О.С. Мезенцева // Нейрокомпьютеры: разработка, применение. – 2015. – № 2. – С. 12-19.
7. Карпова А.Е. Kaggle - платформа для анализа данных / А.Е. Карпова // Вестник магистратуры. – 2018. – № 12-4 (87). – С. 48-49.
8. Катасёв А.С. Нейросетевая биометрическая система распознавания изображений человеческого лица / А.С. Катасёв, Д.В. Катасёва, А.П. Кирпичников // Вестник Технологического университета. – 2016. – Т. 19, № 18. – С. 135-138.

NEURAL NETWORK MODEL OF FACIAL EXPRESSION RECOGNITION TO DETERMINE HUMAN FATIGUE

B. Kurbanov, A.S. Katasev, I.M. Shayahmetov, B.R. Zinnurov

Kazan National Research Technical University named after A. N. Tupolev-KAI
10, st. Karl Marx, Kazan, 420111, Russian Federation

Annotation. The article is devoted to the use of a convolutional neural network model to determine a person's fatigue by facial expression. For this purpose, the architecture of the convolutional neural network ResNet50 is used, which is trained on images and finds signs indicating a person's fatigue and vigor. The prepared dataset includes 6000 images, of which 3000 images correspond to a tired person (in these images the person yawns or closes his eyes) and 3000 images correspond to a cheerful person.

When creating the ResNet50 neural network model, the Python programming language, the Jupyter Notebook development platform and the Anaconda3 development environment were used. The neural network model was trained for 100 epochs, with each training example consisting of 32 elements. Thanks to the use of the Adam optimization algorithm, it was possible to train a neural network to correctly classify images into two classes: fatigue and vigor. During training, the neural network model achieved a classification accuracy level of 95%. As a result of calculating the values of classification quality metrics Recall and Precision on a test data sample using an error matrix, it was possible to obtain high results. For the “tired” class, the value of the Recall metric was 0.9426, and the Precision metric was 0.9587. For the “vigorous” class, the value of the Recall metric was 0.9576, and the Precision metric was 0.9417. These indicators indicate the adequacy of the constructed model and the possibility of its practical use. As a result, we can conclude that the study demonstrated the successful use of a neural network model for facial expression recognition to determine human fatigue with a high degree of accuracy.

Keywords: neural network model, convolutional neural network, human fatigue detection, facial expression, ResNet50.

Статья представлена в редакцию 18 декабря 2023 г.