

НЕЙРОСЕТЕВАЯ МОДЕЛЬ ОПРЕДЕЛЕНИЯ СОНЛИВОСТИ ЧЕЛОВЕКА ПО ВЫРАЖЕНИЮ ЛИЦА

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 исследованию применения сверточной нейросетевой модели YOLOv5 для выявления сонливости человека на основе выражения его лица. Модель обучалась на изображениях лиц людей, распознавая признаки их сонливости и бодрости. Набор данных, состоящий из 2300 изображений, был разделен на два класса: сонливые люди (они либо зевают, либо закрывают глаза) и бодрые. Для разработки модели использовался язык Python, обучение проводилось на платформе Google Colab с применением данных из Kaggle. Процесс обучения продолжался в течение 30 эпох, при этом каждая партия данных состояла из 32 изображений. В результате использования различных улучшений YOLOv5 удалось добиться высокого уровня классификации изображений. Средняя точность модели составила 96,8%. Оценка качества классификации с использованием метрик Recall и Precision на тестовых данных продемонстрировала высокие результаты. Для класса «сонливые» значение Recall составило 83,4%, а Precision – 89,1%. Для класса «бодрые» значение Recall – 93,6%, а Precision – 87,6%. Это свидетельствует о высокой надежности и полезности модели. Исследование подтвердило эффективность использования нейросетевой модели для определения сонливости человека по выражению лица, что открывает широкие перспективы для внедрения данной технологии в системы транспортной безопасности. Это может способствовать снижению риска дорожно-транспортных происшествий и улучшению общей безопасности на дорогах.

Ключевые слова: нейросетевая модель, определение сонливости человека, выражение лица, YOLOv5.

Введение

Сонливость стала серьезной проблемой в современном обществе. Концентрация внимания важна в любом виде деятельности человека, например, при управлении автомобилем [1]. Различные исследования [2-4] показывают, что увеличение случаев сонливости все чаще приводит к серьезным последствиям. Одна из самых распространенных причин дорожно-транспортных происшествий – это сонливость за рулем [5]. Сонливость – это физиологическое состояние человека, характеризующееся повышенной потребностью во сне. Оно может быть вызвано различными факторами, такими как недостаток сна, монотонная деятельность, физическое или эмоциональное утомление. Сонливость оказывает значительное воздействие на человеческий организм и его функционирование. Признаками сонливости являются опущенные веки, закрытые глаза, усталый взгляд, отсутствие живой мимики и пониженный тонус мышц лица. Сонливый водитель создает серьезную угрозу для безопасности всех участников дорожного движения. Для предотвращения опасных ситуаций важно научиться выявлять состояние сонливости у человека и сигнализировать об этом [6]. Это актуализирует необходимость разработки и использования современных интеллектуальных технологий в данной предметной области, основанных на нечеткой логике [7, 8], нейронных сетях [9, 10], сверточных нейронных сетях [11, 12] и др. В данной статье рассматривается решение указанной задачи с помощью построения сверточной нейросетевой модели.

Описание модели сверточной нейронной сети для определения сонливости человека по выражению лица

Задачу определения сонливости человека по выражению лица можно отнести к задаче классификации [13]. Интеллектуальная система определения сонливости человека должна определить класс, к которому относится объект: сонливый или бодрый.

Технологии искусственных нейронных сетей стали неотъемлемой частью многих исследований, особенно в области машинного зрения [14-16]. Наиболее известной технологией в задачах обнаружения и распознавания объектов на изображениях является YOLO (You Only Look Once) – семейство предобученных моделей с использованием глубокого обучения [17]. Особенность данных моделей заключается в том, что на вход подается сразу все изображение, проходя через нейронную сеть только один раз, в отличие от других моделей, где этот процесс происходит многократно. Благодаря такой особенности YOLO позволяет использовать ее в задачах, где требуется быстрое определение класса в режиме реального времени. YOLO использует сверточные сети, обрабатывающие пиксели изображения и конвертирующие их в информацию о классах и координатах объектов.

Архитектура нейронной сети YOLOv5 представлена на рисунке 1.

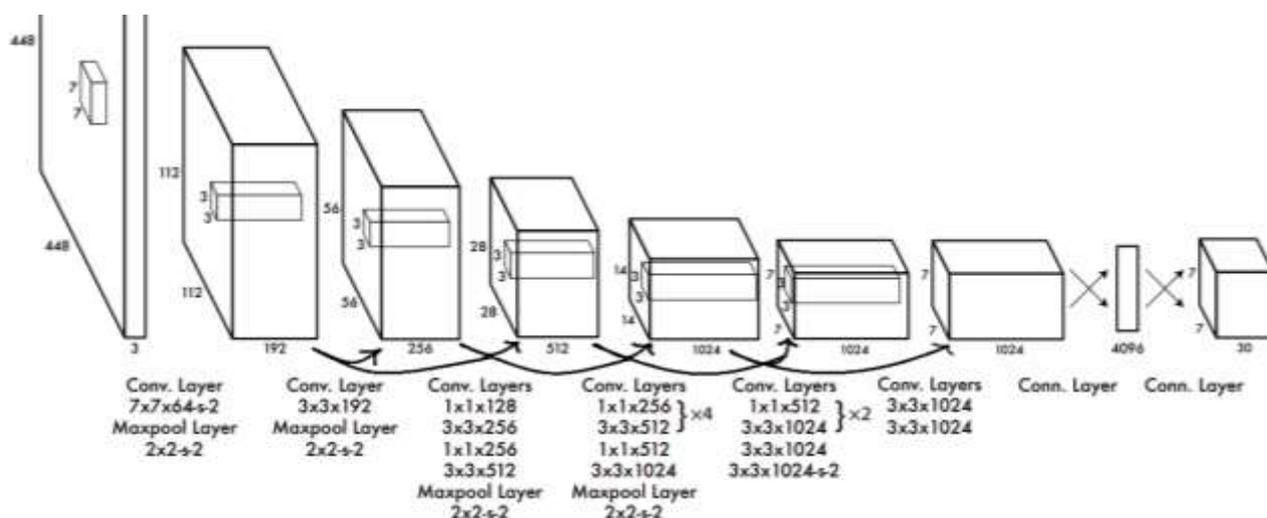


Рис. 1. Слои архитектуры YOLO

Модель включает следующие слои [18]:

- сверточные слои: основной строительный блок модели (сверточные слои выполняют операцию свертки, применяются для выделения признаков из изображений);
- слой сжатия MaxPooling: это один из видов слоя сжатия, который используется для сокращения размерности данных и увеличения области восприятия нейронов;
- соединения пропуска (skip-connections): это специальные соединения, которые помогают сохранить информацию о пространственном разрешении и предотвратить потерю информации при сжатии данных;
- слои объединения: эти слои используются для объединения информации из разных частей сети и улучшения обнаружения объектов разного размера.

Подготовка данных для построения и тестирования модели

Для качественного обучения модели YOLOv5 необходимо подготовить набор данных, состоящий из изображений и соответствующих координат объектов, которые необходимо обнаружить [19]. В работе для обучения модели выбраны наборы данных из от-

крытого источника Kaggle. Важно отметить, что этнический признак может влиять на определение «сонливости» по выражению лица. Поэтому при выборе набора данных учитывалось наличие изображений людей разной этнической принадлежности. Исходные изображения из выбранного набора были разделены на две папки – train (обучающая выборка) и test (тестовая). Каждая папка содержала еще две папки: images (изображения) и labels – папка с текстовыми файлами, содержащими метки объектов на этих изображениях в формате YOLO. Этот формат является стандартным для обучения моделей YOLO и позволяет указать координаты объектов и их классы в текстовых файлах, что делает процесс обучения более гибким и позволяет использовать разные методы аугментации данных. Для создания меток объектов использовался инструмент с открытым исходным кодом LabelImg [20]. После подготовки изображений и меток был создан файл dataset.yaml для указания пути до обучающей и тестовой выборок данных, а также количество классов и их метки. Общий объем данных составил 2300 изображений, которые были разделены на два класса: сонливые и бодрые люди.

На рисунках 2 и 3 представлены примеры изображений из набора данных.



Рис.2. Примеры изображений из набора данных с бодрым человеком



Рис.3. Примеры изображений из набора данных с сонливым человеком

Визуальный сравнительный анализ этих рисунков показывает разницу между бодрым и сонливым человеком. В последнем случае глаза человека либо не полностью открыты, либо закрыты. Кроме того, общее выражение лица сонливого человека отличается от выражения лица человека, находящегося в состоянии бодрствования. Указанные особенности (визуальные признаки сонливости) будут учтены при обучении сверточной нейросетевой модели. В обучающем наборе данных для каждого изображения были созданы метки в формате YOLO, которые указывают координаты глаз и других черт лица. Это позволит сверточной нейросетевой модели научиться идентифицировать эти черты и связывать их с состоянием сонливости или бодрствования человека.

Обучение, тестирование и оценка адекватности нейросетевой модели

Обучение модели нейронной сети является трудозатратным процессом, в связи с чем для обучения использовалась платформа Google Colab [21]. В данной работе при построении нейронной сети YOLOv5 использовался язык программирования Python [22]. Процесс обучения модели длился в течение 30 эпох, при этом в каждой итерации использовалось по 32 обучающих примера.

На рисунке 4 представлен график отношения точности и достоверности обучения сверточной нейросетевой модели.

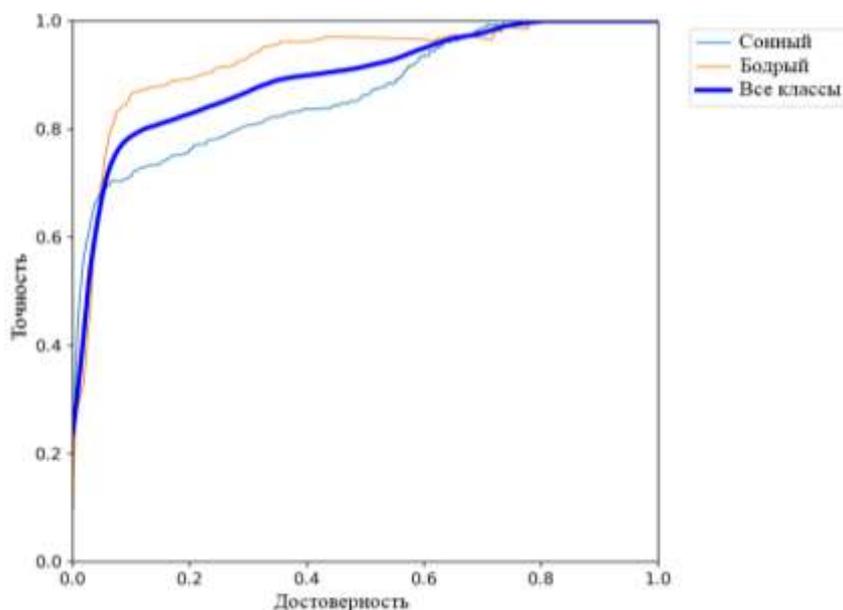


Рис. 4. График отношения точности-достоверности обучения YOLOv5

График включает в себе кривые, иллюстрирующие точность определения состояний человека: сонный (drowsy) и бодрый (awake). Жирная кривая (с меткой «all classes») отображает среднюю точность модели по двум классам.

На рисунках 5 и 6 представлен результат применения обученной нейросетевой модели для определения сонливости человека по выражению лица.



Рис. 5. Результат применения модели (класс бодрый)

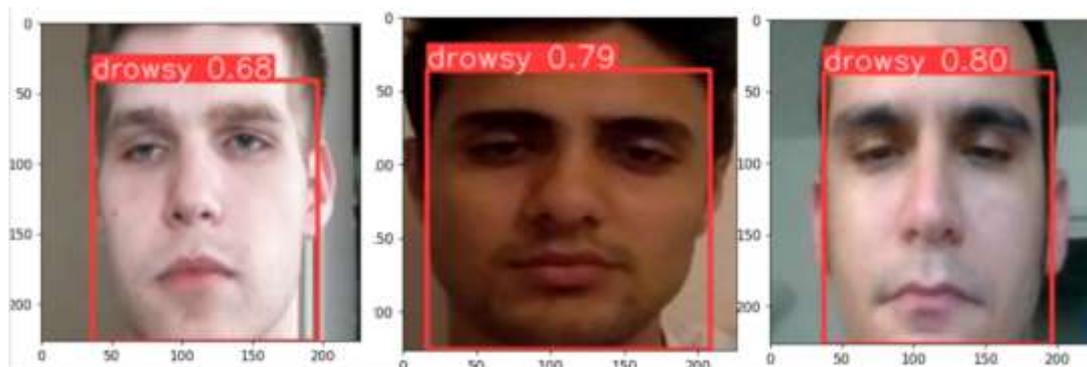


Рис. 6. Результат применения модели (класс сонливый)

На рисунке 5 сверточная нейросетевая модель определила состояния людей как «awake» (бодрый), а на рисунке 6 – как «drowsy» (сонный). В обоих случаях результат классификации является верным.

После обучения произведена оценка обобщающей способности модели на тестовых данных. Точность определения класса «бодрый» составила 97,9%, а класса «сонливый» – 95,7%. Средняя точность модели составила 96,8%, что является высоким результатом для использования в системе определения сонливости человека по выражению лица.

Для оценки эффективности модели построена матрица ошибок, а также рассчитаны метрики полнота, точность, меткость и F1-мера [23, 24]. На рисунке 7 представлена матрица ошибок обученной модели.

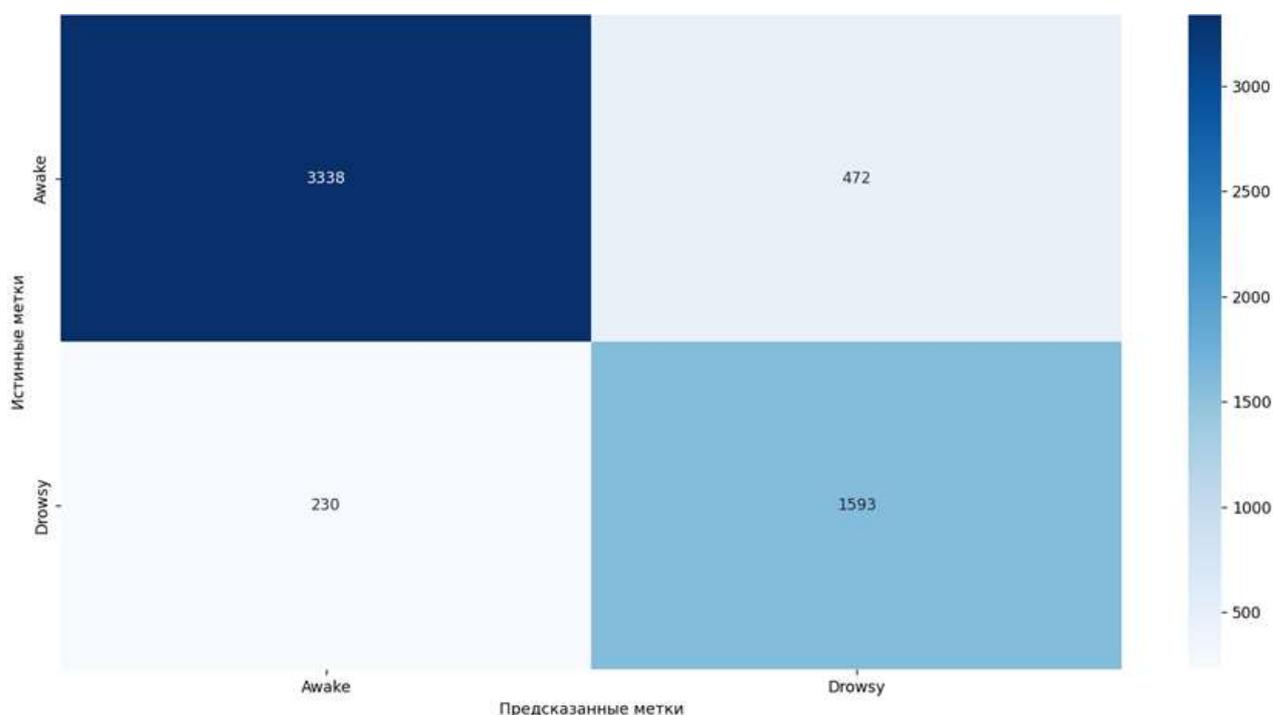


Рис.7. Матрица классификации для построенной модели

Матрица ошибок позволяет вычислять метрики оценки качества работы модели, такие как полнота (recall) и точность (precision). Эти метрики позволяют оценить, насколько хорошо модель справляется с задачей оценки сонливости.

Для класса «бодрый» (awake):

$$\text{Recall} = 3338 / (3338 + 230) = 93,6\%$$

$$\text{Precision} = 3338 / (3338 + 472) = 87,6\%$$

Для класса «сонливый» (drowsy):

$$\text{Recall} = 1593 / (1593 + 472) = 83,4\%$$

$$\text{Precision} = 1593 / (1593 + 230) = 89,1\%$$

Метрика Accurasy (меткость) показывает насколько правильно модель идентифицирует или исключает определенный класс. Вычисляется как отношение числа правильных классификаций к объему всех данных. Для разработанной модели значение меткости Accurasy = $(3338 + 1593) / (3338+472 +1593 + 230) = 87,5\%$.

Метрика F1-мера (F1-score) – это статистический показатель, используемый для оценки точности бинарных классификационных моделей, который учитывает, как точность, так и полноту тестируемой модели. Для разработанной модели можно вычислить F1-меру, используя значения Recall и Precision:

Для класса «бодрый» (awake):

$$\text{F1-мера} = 2 * (93,6\% * 87,6\%) / (93,6\% + 87,6\%) = 90,5\%$$

Для класса «сонливый» (drowsy):

$$\text{F1-мера} = 2 * (83,4\% * 89,1\%) / (83,4\% + 89,1\%) = 86,2\%$$

Результаты оценки эффективности построенной модели представлены в таблице 1.

Таблица 1. Значения метрик качества классификации для построенной модели

Класс	Точность (Precision)	Полнота (Recall)	F1-мера (F1-score)	Меткость (Accurasy)
Сонливый (drowsy)	87,4%	86,2%	81,9%	87,5%
Бодрый (awake)	87,6%	93,6%	90,5%	

Полученные результаты указывают на адекватность построенной модели с точки зрения решения задачи определения сонливости человека по выражению лица.

Заключение

В данной статье описана построенная нейросетевая модель определения сонливости человека по выражению лица. Полученные результаты показали, что модель достаточно точно классифицирует большинство случаев состояния сонливости и бодрости человека. Таким образом, модель является эффективным инструментом для определения сонливости человека по выражению лица в различных предметных областях [25-28]. В дальнейшем для совершенствования модели можно будет применить более сложные методы аугментации данных, например, MixUp [29] или CutMix [30], которые помогут увеличить разнообразие обучающего набора данных и улучшить точность классификации. Кроме того, необходимо провести отдельный анализ влияния этнических особенностей на точность классификации модели. Это позволит создать более точную, функциональную и универсальную модель для определения сонливости человека по выражению лица.

Список литературы

1. Попов А.В. Исследование проявлений утомления среди водителей грузовых автомобилей в Российской Федерации / А.В. Попов, Р.О. Соколов // Грузовик. – 2021. – № 1. – С. 37-41.
2. Yung M. Examining the fatigue-quality relationship in manufacturing / M. Yung, A. Kolus, R. Wells, W. Neumann // Applied Ergonomics. – 2020. – Vol. 82. – Article 102919.

3. Xiang J. A Review of Researches on Sleepiness at Workplace / J. Xiang, Y. Maolin//Journal of Human Resource and Sustainability Studies. – 2016. – Vol. 4, No. 3. – P.155-161.
4. Саркисян Н.В. Расширенное исследование подхода к профилактике сонливости водителей за рулем посредством ольфакторной стимуляции / Н.В. Саркисян, А.Ю. Сорокин // International Journal of Medicine and Psychology. – 2021. – Т. 4, № 8. – С.87-91.
5. Левичев В. Основные причины ДТП / В. Левичев // Мир дорог. – 2021. – № 136. – С. 42-43.
6. Булыгин А.О. Анализ современных исследований в области детектирования утомления водителя в кабине транспортного средства / А.О. Булыгин, А.М. Кашевник // Системы анализа и обработки данных. – 2021. – № 3 (83). – С. 19-36.
7. Катасёва Д.В. Формирование баз знаний интеллектуальных систем на примере нейронечеткого анализа медицинских данных / Д.В. Катасёва, А.О. Барина // Вестник Технологического университета. – 2022. – Т. 25, № 2. – С. 67-70.
8. Катасёв А.С. Аппроксимация объектов с дискретным выходом на основе нечетко-продукционных баз знаний / А.С. Катасёв // Вестник Казанского государственного технического университета им. А.Н. Туполева. – 2013. – № 4. – С. 212-217.
9. Воробьева Ю.Н. Нейросетевая модель выявления DDos-атак / Ю.Н. Воробьева, Д.В. Катасёва, А.С. Катасёв, А.П. Кирпичников // Вестник Технологического университета. – 2018. – Т. 21, № 2. – С. 94-98.
10. Катасёв А.С. Нейронечеткая модель и программный комплекс автоматизации формирования нечетких правил для оценки состояния объектов / А.С. Катасёв // Автоматизация процессов управления. – 2019. – № 1 (55). – С. 21-29.
11. Катасёв А.С. Сверточная нейросетевая модель определения усталости человека по выражению лица / А.С. Катасёв, Б. Курбанов // Вестник Технологического университета. – 2023. – Т. 26, № 3. – С. 67-71.
12. Калинин Н.Н. Оценка производительности моделей машинного обучения, основанных на библиотеках Tensorflow и Dlib при определении состояния сонливости на Raspberry PI / Н.Н. Калинин, Ю.Г. Скурыдин, Е.М. Скурыдина // Высокопроизводительные вычислительные системы и технологии. – 2023. – Т. 7, № 1. – С. 106-114.
13. Алексеев А.А. Классификация текстовых документов на основе технологии Text Mining / А.А. Алексеев, А.С. Катасёв, А.Е. Кириллов, А.П. Кирпичников // Вестник Технологического университета. – 2016. – Т. 19, № 18. – С. 116-119.
14. Кондратьев С.Е. Применение методов компьютерного зрения и машинного обучения для распознавания технических требований на 2d-чертежах / С.Е. Кондратьев, О.В. Ульянин, Е.М. Абакумов // Информационные технологии в проектировании и производстве. – 2021. – № 2 (182). – С. 35-43.
15. Романчиков А.М. Определение траектории движения локомотива для систем машинного зрения / А.М. Романчиков, В.А. Гросс, П.Е. Мащенко, К.С. Болдырев // Железнодорожный транспорт. – 2020. – № 11. – С. 61-63.
16. Хасаншин И.Я. Применение технологий машинного зрения для распознавания одиночных прямых ударов в боксе / И.Я. Хасаншин, Д.С. Уткин, Д.Н. Дербин // Наука и спорт: современные тенденции. – 2022. – Т. 10, № 2. – С. 43-48.
17. Jiang P. A Review of Yolo Algorithm Developments / P.Jiang, D. Ergu, F. Liu, Y. Cai, B. Ma // Procedia Computer Science. – 2022. – Vol. 199. – P.1066-1073.
18. Terven J. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS / J. Terven, D. Esparza // Deep Learning in Image Analysis and Pattern Recognition. – 2023. – Vol. 5. – P.1680-1716.

19. Менциев А.У. Основные процессы алгоритма YOLOV5 в системах видеонаблюдения / А.У. Менциев, Т.Г. Айгулов, Ш.А. Абдурашидов // Приборы и системы. Управление, контроль, диагностика. – 2023. – № 12. – С.15-18.
20. Mathew M.P. Leaf-based disease detection in bell pepper plant using YOLO V5 / M.P. Mathew, T.Ya. Mahesh // Signal, Image and Video Processing. – 2022. – Vol. 16, No. 3. – P. 841-847.
21. Mishra H.R. Secure identification at your fingertips: building a face recognition system with Google Colab / H.R. Mishra, Sh.Vadgama, H. Pandya // International Journal for Research in Applied Science and Engineering Technology. – 2023. – Vol. 11, No. 3. – P. 1040-1044.
22. Лобарёв Д.С. Поиск оптимального решения дифференциальной линейно-квадратичной задачи управления средствами Python в облачной среде Google Colab / Д.С. Лобарёв, Д.В. Толбухин // Научно-технический вестник Поволжья. – 2021. – № 12. – С. 208-213.
23. Ахметвалеев А.М. Модель коллектива нейронных сетей и программный комплекс для определения функционального состояния человека / А.М. Ахметвалеев, А.С.Катасёв, М.А. Подольская // Прикаспийский журнал: управление и высокие технологии. – 2018. – № 1 (41). – С. 69-85.
24. Буланов В.А. Метрики оценки качества модели классификации на примере задачи кредитного скоринга / В.А. Буланов, О.Е. Фомичева // Приборы и системы. Управление, контроль, диагностика. – 2020. – № 3. – С. 33-37.
25. Катасёв А.С., Катасёва Д.В., Сибгатуллин А.А. Нейросетевая модель оценки функционального состояния водителей в системах транспортной безопасности / А.С. Катасёв, Д.В. Катасёва, А.А. Сибгатуллин // Электроника, фотоника и киберфизические системы. – 2023. – Т. 3, № 1. – С. 69-80.
26. Баринаова А.О. Формирование и использование базы знаний для оценки функционального состояния водителей автотранспортных средств / А.О. Баринаова, Д.В. Катасёва, А.С. Катасёв // Вестник Технологического университета. – 2020. – Т. 23, № 10. – С. 75-78.
27. Лашков И.Б. Детектирование состояния зевоты у водителя транспортного средства при помощи модели сверточной нейронной сети / И.Б. Лашков // Научно-технический вестник информационных технологий, механики и оптики. – 2022. – Т. 22, № 1. – С. 33-46.
28. Ахметвалеев А.М., Катасёв А.С. Нейросетевая модель определения функционального состояния опьянения человека в решении отдельных задач обеспечения транспортной безопасности / А.М. Ахметвалеев, А.С. Катасёв // Компьютерные исследования и моделирование. – 2018. – Т. 10, № 3. – С. 285-293.
29. Рюмина Е.В. Метод генерации обучающих данных для компьютерной системы обнаружения защитных масок на лицах людей / Е.В. Рюмина, Д.А. Рюмин, М.В.Маркитантов, А.А. Карпов // Компьютерная оптика. – 2022. – Т. 46, № 4. – С. 603-611.
30. Хохлов М.Р. Применение технологии компьютерного зрения в системах автоматизированного складского учёта / М.Р. Хохлов // ИТ-Стандарт. – 2020. – № 4 (25). – С. 48-54.

NEURAL NETWORK MODEL OF DETERMINATION DROWSYNESS OF A PERSON BASED ON FACIAL EXPRESSIONS

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 study of the application of the YOLOv5 convolutional neural network model for detecting human drowsiness based on their facial expressions. The model was trained on images of people's faces, recognizing signs of their drowsiness and vigor. The dataset, consisting of 2300 images, was divided into two classes: sleepy people (they either yawn or close their eyes) and vigorous. The Python language was used to develop the model, the training was carried out on the Google Colab platform using data from Kaggle. The training process lasted for 30 epochs, with each data batch consisting of 32 images. As a result of using various improvements, YOLOv5 managed to achieve a high level of image classification. The average accuracy of the model was 96.8%. Evaluation of the classification quality using the Recall and Precision metrics on the test data showed high results. For the sleepy class, the Recall value was 83.4%, and the Precision value was 89.1%. For the "vigorous" class, the Recall value is 93.6%, and the Precision value is 87.6%. This indicates the high reliability and usefulness of the model. The study confirmed the effectiveness of using a neural network model to determine a person's drowsiness by facial expression, which opens up broad prospects for implementing this technology in transport safety systems. This can help reduce the risk of road accidents and improve overall road safety.

Keywords: neural network model, human sleepiness detection, facial expression, YOLOv5.

Статья представлена в редакцию 30 сентября 2024 г.