

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

COMPUTER SCIENCE

doi: 10.17586/2226-1494-2022-22-1-33-46

УДК 004.04

Детектирование состояния зевоты у водителя транспортного средства при помощи модели сверточной нейронной сети

Игорь Борисович Лашков

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), Санкт-Петербург, 199178, Российская Федерация

igor-lashkov@ya.ru[✉], <https://orcid.org/0000-0001-6418-4660>

Аннотация

Предмет исследования. Среди факторов, потенциально являющихся причиной возникновения дорожно-транспортных происшествий, является усталость водителя, накапливающаяся во время поездки или присутствующая еще до ее начала. Распространенный признак усталости или утомления у водителя транспортного средства — зевота. Выявление признаков зевоты у человека способно охарактеризовать состояние усталости. Методы компьютерной обработки изображений активно применяются в задачах детектирования открытости рта и определения зевоты у человека. Однако такой подход обладает перечнем недостатков, к которым относятся различные окружающие условия работы и множество ситуационных вариантов зевоты для разных людей. Представлена схема детектора выявления признаков зевоты, заключающаяся в обработке изображений лица водителя с использованием методов анализа данных, компьютерной обработки изображений и сверточной модели нейронной сети. **Метод.** Сущность предложенного метода заключается в детектировании признаков состояния зевоты в поведении водителя в кабине транспортного средства на основе анализа последовательности изображений, полученных с видеокамеры. Показано, что возникновение состояния зевоты водителя сопровождается широким и продолжительным открытием рта. Длительная открытость рта сигнализирует о возникновении признаков зевоты у человека. Представлена концептуальная модель обнаружения открытости рта у человека, и разработана схема обработки и полуавтоматической разметки наборов данных YawDD и Kaggle Drowsiness Dataset. Разработанная модель сверточной нейронной сети показала точность равную 0,992 и полноту — 0,871 на тестовом 10 % наборе данных. **Основные результаты.** Предложенная схема детектора выявления признаков зевоты апробирована на тестовой выборке видео, сформированной набором данных YawDD: Yawning Detection Dataset. Данная схема детектирования успешно обнаружила 124 зевка среди всех видеофайлов из тестового набора данных. Доля правильно классифицированных объектов составила 98,2 %, точность — 96,1 %, полнота — 98,4 %, а F-мера — 97,3 % при определении признаков зевоты у водителя транспортного средства. **Практическая значимость.** Детектирование признаков зевоты в поведении водителя позволяет уточнить о нем информацию, и тем самым повысить эффективность существующих систем мониторинга водителя в кабине транспортного средства, ориентированных на предотвращение и снижение риска возникновения дорожно-транспортных происшествий. Предлагаемый подход может быть использован совместно с другими технологиями мониторинга поведения водителя при построении интеллектуальной системы поддержки водителя.

Ключевые слова

транспортное средство, водитель, зевота, камера, мониторинг, обработка информации

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-29-06099 (разработка методов поиска уязвимостей интерфейсов взаимодействия человека с искусственным интеллектом транспортной среды «умного города»). Концептуальная модель мониторинга поведения водителя была поддержана бюджетной темой СПб ФИЦ РАН № FFZF-2022-0005.

Ссылка для цитирования: Лашков И.Б. Детектирование состояния зевоты у водителя транспортного средства при помощи модели сверточной нейронной сети // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 1. С. 33–46. doi: 10.17586/2226-1494-2022-22-1-33-46

Abstract

Among the factors that usually cause road accidents in the world is driver fatigue, which accumulates during the trip or is present even before it begins. One of the most common signs of fatigue or tiredness of a vehicle driver is yawning. The detection of signs of yawning in human behavior is potentially able to further characterize its state of fatigue. Computer image processing methods are actively used to detect the openness of the mouth and yawning for a person. However, this approach has many disadvantages, which include different environmental conditions and a variety of situational yawning options for different people. The paper presents a scheme of a detector for determining signs of yawning, which is focused on processing images of the driver's face using data analysis methods, computer image processing, and a convolutional neural network model. The essence of the proposed method is to detect yawning in the driver's behavior in the cabin of a vehicle based on the analysis of a sequence of images obtained from a video camera. It is shown that the driver's yawning state is accompanied by a wide and prolonged openness of the mouth. Prolonged openness of the mouth signals the appearance of signs of yawning. A conceptual model for detecting the openness of the mouth for a vehicle driver is presented and a scheme for processing and labeling the YawDD and Kaggle Drowsiness Dataset datasets is developed. The developed convolutional neural network model showed an accuracy of 0.992 and recall of 0.871 on a test 10 % data set. The proposed scheme for detecting the yawning state has been validated on a test video subset extracted from the YawDD: Yawning Detection Dataset. This detection scheme successfully detected 124 yawns among all video files from the test dataset. The proportion of correctly classified objects is 98.2 % accuracy, precision is equal to 96.1 %, recall is 98.4 %, and F score is 97.3 % while detecting signs of yawning in driver behavior. Detecting signs of yawning in the driver's behavior allows one to clarify information about the driver and thereby to increase the effectiveness of existing driver monitoring systems in the vehicle cabin, aimed at preventing and reducing the risk of road accidents. The proposed approach can be combined with other technologies for monitoring driver behavior when building an intelligent driver support system.

Keywords

vehicle, driver, yawning, camera, monitoring, information processing

Acknowledgements

The research was carried out with the financial support of the RFBR within the scientific project No. 19-29-06099 (development of methods for finding vulnerabilities in human interaction interfaces with artificial intelligence of the "smart city" transport environment). The conceptual model of driver behavior monitoring was supported by the budget theme of the SPC RAS No. FFZF-2022-0005.

For citation: Lashkov I.B. Detection of yawning in driver behavior based a convolutional neural network. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 1, pp. 33–46 (in Russian). doi: 10.17586/2226-1494-2022-22-1-33-46

Введение

Дорожно-транспортные происшествия (ДТП) по-прежнему остаются одной из серьезных проблем в мире. Сложно найти страну, которую не затронули дорожные инциденты. Тысячи людей ежегодно погибают или получают серьезные ранения по причине того, что водители засыпают за рулем. Согласно аналитическим отчетам Всемирной организации здравоохранения, увеличение средней скорости движения на 1 % приводит к увеличению числа аварий со смертельным исходом на 4 % и числа аварий с серьезными травмами — на 3 %¹. Согласно публично открытым отчетам, в Европе за 2019 год в результате ДТП погибло почти 23 000 чел.². В соответствии с отчетом национального управления безопасности дорожного движения (БДД) США (NHTSA),

более 91 000 зарегистрированных полицией аварий в 2017 г. были связаны с засыпанием водителей за рулем транспортного средства (ТС)³. Результаты расследования, проведенного Национальным управлением БДД [1] в США, показали, что каждый третий респондент испытывал усталость во время вождения автомобиля. Среди тех, у кого были инциденты с вождением в состоянии сонливости, 10 % признались, что у них были такие инциденты в течение последнего месяца вождения. Установлено, что усталость водителя приводит к возникновению 22 % ДТП. В отсутствие какого-либо предупреждения усталость водителя повышает вероятность столкновения в шесть раз, чем при вождении в нормальном состоянии. Таким образом, исследования в области распознавания усталости водителя, несомненно, являются необходимыми для повышения БДД. За последние десятилетия предложено множество методов определения усталости водителей [2, 3], чтобы обратить внимание водителей на безопасное управление автомобилем и, в целом, повысить БДД. Поведенческие

¹ WHO, Road traffic injuries [Электронный ресурс]. URL: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries> (дата обращения: 09.08.2021).

² Road safety: Europe's roads are getting safer but progress remains too slow | Mobility and Transport [Электронный ресурс]. URL: https://ec.europa.eu/transport/media/news/2020-06-11-road-safety-statistics-2019_en (дата обращения: 09.08.2021).

³ N. Administration, «Drowsy Driving», NHTSA, 2019 [Электронный ресурс]. URL: <https://www.nhtsa.gov/risky-driving/drowsy-driving/> (дата обращения: 10.08.2021).

характеристики водителей при утомительном вождении включают моргание, кивание, закрытие глаз и зевоту. Среди водителей автотранспорта зевота – одна из ключевых форм проявления усталости при вождении [4].

Усталость обычно воспринимается как чувство сонливости. В контексте, связанном с работой человека, усталость — состояние умственного и/или физического истощения, которое снижает способность безопасно и эффективно выполнять работу [5]. Усталость может быть вызвана рядом факторов, включая недостаточный отдых, чрезмерную физическую или умственную активность, нарушения режима сна или чрезмерный стресс. Существует множество последствий, которые могут повлиять на отдельных людей и которые считаются непосредственно связанными с усталостью, такими как нарушение зрения, снижение зрительно-моторной координации, низкая мотивация, сниженное внимание, медленные рефлексы, недостаточная реакция или неспособность сосредоточиться. Усталость считается основной причиной ДТП, приводящих к человеческим жертвам. Данные за 2010 год от Министерства транспорта Великобритании¹ указывают на 1850 погибших и 22 660 серьезно раненых, при этом 20 % от общего числа несчастных случаев [6] произошли по причине усталости водителя за рулем автомобиля. Федеральное управление автомобильных дорог США утверждает, что около 100 000 аварий в год происходят по причине усталости водителя за рулем ТС [7].

Предмет исследования

Интеллектуальные технологии в транспорте, выполняющие мониторинг и оказывающие помощь водителю в управлении ТС [8], в последние годы стали активно исследуемой тематикой в области повышения БДД. В связи с этим автопроизводители стремятся оборудовать свои автомобили различными видами передовых вспомогательных систем для водителя, которые способствуют безопасной эксплуатации и помогают водителю в различных небезопасных ситуациях. Контроль безопасности включает: съезд и удержание в полосе движения автомобиля; предупреждение о возникающих препятствиях, ТС и пешеходах; об экстренном торможении в случае, если водитель не предпринимает действий, чтобы избежать столкновения; визуализацию дорожных знаков на приборной панели автомобиля и др. Эффективный мониторинг поведения водителей позволяет повысить БДД не только для самих водителей, но и дает преимущества страховым и логистическим компаниям с точки зрения сбережений и потребностей общества.

Сонливость является немаловажным физиологическим состоянием человека, по причине которого происходит около 24 % всех ДТП в мире [9]. Вождение автомобиля в состоянии сонливости повышает риск возникновения дорожной аварии в 4–6 раз в сравнении с состоянием бдительности. Одним из предупрежда-

ющих сигналов, по возникновению которого можно судить о состоянии усталости у водителя, является зевота. Зевота — активно исследуемая тематика в научной сфере, предоставляющая неинвазивный механизм определения физиологического состояния человека.

Исследователи предложили различные методы для обнаружения состояния зевоты. Существующие методы обнаружения зевоты сосредоточены на измерении открывания рта при анализе одного статического изображения в единицу времени. При исследованиях применяются методы компьютерной обработки изображений для поиска и локализации области рта человека, необходимых для выявления признаков зевоты, основываясь на различиях в характеристиках между открытым и закрытым ртом. Нейронные модели глубокого обучения начинают получать распространение в области систем интеллектуальных транспортных средств [10] и, в том числе, находят применение при обнаружении состояния открытости и закрытости рта. Подходы, основанные на работе со статическим изображением, сталкиваются со значительными трудностями и не всегда могут обеспечить приемлемое качество и скорость распознавания признаков небезопасного поведения по одному изображению. Для решения перечисленных трудностей в настоящей работе предложен подход к выявлению зевоты, основанный на распознавании набора статических изображений лица. Временной промежуток от возникновения признаков усталости у водителя за рулем с признаками зевоты до серьезных травм, включая летальный исход, может быть достаточно коротким. Таким образом, эффективность алгоритма обнаружения зевоты у водителя при управлении ТС во времени имеет решающее значение. Эффективность алгоритма — свойство алгоритма, которое связано с численной оценкой его качества, которая, как правило, определяется при помощи метрик полноты, точности, правильно классифицированных объектов, F-меры.

Цель работы — разработка подхода для детектирования признаков зевоты в поведении водителя в кабине ТС на основе обработки и анализа изображений с видеокамеры. Предложенный подход основан на мониторинге поведения водителя в кабине ТС и позволяет заблаговременно выявить признаки небезопасного поведения человека при вождении.

Научная значимость работы заключается в разработке подхода, ориентированного на выявление признаков состояния зевоты у водителя ТС с использованием алгоритмов анализа и обработки данных, позволивших сформировать единый датасет данных. Датасет включает в себя: наборы данных, представленные открытым и доступным для загрузки источником видео YawDD, а также аннотированным набором данных на платформе Kaggle, с разделением на классы «зевающий» и «не зевающий» водитель; модели нейронной сети, обученной с использованием данного сформированного датасета, цель которого — детектирование открытости рта на изображениях; схемы выявления признаков зевоты у водителя ТС на временном промежутке. Одновременное использование данных о визуальном (внешнем) состоянии водителя из разных источников данных способно повысить точность и надежность работы системы под-

¹ D. f. T. UK, Reported Road Casualties in Great Britain: 2010 Annual Report, Department for Transport UK, United Kingdom, Statistical Release – 29 September 2011.

держки водителя. Данное исследование представляет продолжение существующих исследований авторов настоящей работы в области детектирования усталости водителя за рулем ТС [11]. Полученные результаты могут быть использованы при построении алгоритма выявления признаков состояния зевоты для системы мониторинга поведения водителя в кабине ТС [12], например Drive Safely [13].

Исследования в области систем мониторинга состояния зевоты водителя

Существующие методы обнаружения зевоты можно условно разделить на две категории: методы компьютерной обработки изображений и методы, использующие в своей работе алгоритмы машинного обучения и математические модели нейронных сетей.

Методы компьютерной обработки изображений. Такие методы основаны на извлечении геометрических и цветовых признаков. Исследователи [14] предложили метод, который обнаруживает признаки зевоты через изменения в соотношении сторон области границ рта. Область границ рта измеряется на основе самой темной области между губами. Такая область сегментируется с использованием адаптивного порога. Для выявления ситуации, когда при зевоте рот водителя закрыт рукой, использован оператор локальных бинарных шаблонов (Local Binary Patterns) для классификации и извлечения признаков области рта, с целью дальнейшей оценки такого состояния при помощи классификатора нейронной сети. Чтобы убедиться, что закрытие рта рукой вызвано именно зевотой человека, а не каким-то другим его действием или состоянием, производится анализ появления морщин на лице. Была проведена серия экспериментов, чтобы найти предельное значение соотношений сторон. Когда соотношение сторон рта меньше фиксированного порогового значения, состояние рта классифицируется как зевота. Состояние зевоты детектируется, когда происходят значимые искажения в рассмотренных областях лица. Предложенный алгоритм обнаруживает зевоту при открытии рта длительностью от 3 до 5 с. Если в рассматриваемый момент открытость рта меньше порогового значения, рот считается закрытым. Тогда на следующем шаге происходит выявление состояния закрытия рта рукой человеком и наличие искажения в верхней области лица. Если оба условия выполняются, алгоритмом фиксируется состояние зевоты у человека, а в ином случае — ее отсутствие в рассматриваемый момент.

Достаточно распространенным подходом к определению состояния зевоты у человека — наблюдение за величиной открытости рта на протяжении некоторого промежутка времени, приблизительно равного несколькими секундами. При решении задачи детектирования признаков зевоты большая часть существующих исследований концентрируется на детектировании области лица человека, и затем остальных его частей, таких как рот, нос и глаза, с последующим обнаружением состояния зевоты. Один из популярных подходов к распознаванию лиц — применение метода Виола–Джонса (Viola–Jones) [15, 16], представленного в библиотеке

компьютерного зрения OpenCV [17], и обученного детектора Хаара. Существенным недостатком использования такого детектора лиц является то, что идентификация лиц среди множества кандидатов, вырезанных с кадров при различном масштабе с помощью скользящего окна и других методов, сопровождается значительной вычислительной нагрузкой на целевые устройства. Также стоит отметить, что данный каскадный детектор лиц способен находить область лиц только при относительно небольшом определенном угле обзора камеры. Эффективность работы такого детектора лиц непосредственно зависит от внешних окружающих условий, к которым относятся окружающая освещенность, наличие бликов, теней и др. При недостаточной освещенности или наличии каких-либо объектов, частично или полностью перекрывающих видимость лица человека в кадре, точность и другие показатели работы данного алгоритма могут быть значительно снижены.

Работы [18, 19] основаны на детектировании открытости рта по лицевым характеристикам и множеству 68 опорных точек на лице. Отметим работу [20], в которой предложен подход, заключающийся в локализации и сегментации области интереса по кадрам входного видеофайла. Алгоритм Виола–Джонса использован для определения местоположения области лица и рта человека на изображении. Чтобы провести границу между примерами положительного и отрицательного класса, при которых лицо человека на изображении присутствует и там, где его нет, анализ и классификация объектов выполнены при помощи алгоритма машинного обучения адаптивного бустинга Adaboost [21]. С целью обнаружения состояния зевоты область губ была локализована и выбрана из изображения. Для определения краев губ использован метод активных контуров [22]. Принцип активных контуров был впервые предложен в 1998 г. [23]. Активный контур — параметрическая кривая, которая стремится изменить свое положение таким образом, чтобы ее энергия оказалась минимальной. Далее к полученному сигналу применяется функция пересечения нуля, чтобы определить состояние зевоты у водителя.

Методы, основанные на применении математических моделей нейронных сетей. В других исследованиях активно прорабатываются подходы, основанные на использовании не только методов компьютерной обработки изображений, но и глубоких нейронных сетей. В [24] описан алгоритм для определения усталости и сонливости водителя по признакам зевоты и состоянию закрытости глаз. Обнаружение открытости рта и зевоты водителя становится возможным благодаря нахождению кончика носа путем получения информации о глубине лица на изображении. Зная расположение кончика носа, нижняя часть лица может быть отделена (отрезана) от всего изображения. На полученном изображении с помощью адаптивной маски, которая действует как детектор границ, область рта обнаружена с точностью 91 %. Два метода, один из которых основан на максимальной глубине, а другой — на модели активного контура, были применены для задачи поиска области рта, и их результаты были объединены для выявления открытости рта, продемонстрировав

точность в 95 %. Несмотря на то, что среднее количество последовательных кадров в видео, на которых обнаружена зевота, составляет 85, в качестве порогового значения видеок кадров для определения состояния зевоты было выбрано 45, обозначающих число таких кадров, на которых рот человека открыт, описывающих проявление признаков зевоты. Далее, как только найдена верхняя область лица, происходит обнаружение области глаз на изображении при помощи алгоритма Виолы–Джонса. Затем, используя предварительно обученную сверточную нейронную сеть для выделения признаков VGG (Visual Geometry Group), открытость/закрытость глаз была классифицирована, показав точность в 90 %. В такой нейронной модели были обучены только последние два полно связанных слоя и слой softmax для распознавания открытости/закрытости глаз человека. Входное изображение для нейронной сети имеет размер 24×24 пикселей. Исследователи выбрали значение для показателя PERCLOS (PERcentage of eye CLOSure), равным доле времени, в течение которого глаза закрыты, на уровне в 0,5 (т. е. глаза закрыты 15 кадров из доступных 30). Также, с целью снижения количества ошибок второго рода (false negative), было предложено игнорировать 20 % кадров (что соответствует трем кадрам), на которых было обнаружено, что глаза водителя закрыты, среди всех 15 кадров, на которых отмечено, что глаза закрыты. Путем объединения результатов алгоритмов обнаружения зевоты и детектирования закрытости глаз принимается решение о предупреждении водителя. Использование информации о глубине, которая не чувствительна к изменениям освещенности, является основным из преимуществ предложенного решения. Разработанный алгоритм не сработает в тех случаях, когда водитель прикрывает рот рукой или он носит очки.

Большинство современных методов обнаружения зевоты основаны на одном статическом изображении, и могут вызывать ложные срабатывания, поскольку зевота представляет собой действие, а не статическое состояние. Так, для решения этой проблемы в работе [25] представлен метод детектирования состояния усталости, осуществляющий мониторинг уровня сонливости, и способный отображать выходные данные сверточной нейронной сети (CNN) на уровне кадров в вероятность (уверенность) нахождения человека в таком состоянии. Задача в данном исследовании сформулирована как бинарная классификация, заключающаяся в определении класса для каждого кадра видео: «зевающий» или «не зевающий». В качестве классификатора выступает глубокая нейронная сеть, построенная на основе архитектуры сети ResNet-50. Обучение нейронной модели выполнено следующим способом. Параметры данной нейронной модели инициализированы с использованием параметров, взятых от модели, обученной на наборе данных ImageNet. Обучение модели произведено с использованием алгоритма стохастического градиентного спуска для оптимизации функции потерь перекрестной энтропии. На этапе тестирования нейронной модели генерируется 10 случайных примеров тестовых изображений, и окончательный результат прогнозирования записан путем голосования большинством голосов из

10 прогнозов для этих примеров. При проведении начальных экспериментов использован открытый набор данных YawDD [26], содержащий 322 видеофрагмента, записанных внутри кабины ТС. Результаты работы [26] подтверждают достоверность гипотезы о том, что обнаружение зевоты следует проводить в течение длительного непрерывного захвата видео, а не на отдельно взятых изображениях. Среди недостатков данного метода стоит выделить нестабильное качество его работы в условиях низкого уровня освещенности.

В [27] предложен подход к обнаружению зевоты на основе видеофайлов. Такой подход состоит из трех основных компонентов. Первый компонент представляет собой модуль предварительной обработки входных данных. Этот этап включает распознавание лиц и сегментацию, нормализацию размера изображения и шумоподавление. Алгоритм Виолы–Джонса используется для обнаружения и сегментации области лица, чтобы уменьшить влияние информации о внешнем виде фона на изображении. Размер последовательных кадров должен быть одинаковым после сегментации. Затем однородные последовательные кадры исключаются с помощью алгоритма быстрой медианной фильтрации для уменьшения шума изображения. Обработанные кадры подаются на вход второго компонента — модуля выбора ключевых кадров. Ключевые кадры выбираются посредством вычисления гистограммы изображения и обнаружения выбросов для устранения избыточных кадров. Далее ключевые кадры поступают на вход в нейронную сеть распознавания лиц 3D-LTS (3D deep learning network with a low time sampling characteristic). В качестве третьего компонента выступает модуль классификации движений лица. В этом модуле 3D-сеть глубокого обучения характеристикой LTS построена с использованием ключевых видеок кадров для выявления различных незначительных движений лица. LTS — это характеристика изменения скорости дискретизации временных измерений для улучшения распознавания едва заметных лицевых действий. 3D-LTS можно разделить на две части. Кратковременные пространственно-временные характеристики лица извлекаются с помощью 3D-свертки. Двухнаправленная долговременная кратковременная память (LSTM, Bidirectional long short-term memory) используется для создания долговременных последовательностей с помощью извлеченных кратковременных пространственно-временных функций для улучшения эффекта распознавания. Распознавание лицевых действий способно обнаруживать признаки зевоты у человека и уменьшать количество ложных срабатываний.

Существующие исследования показали, что усталость водителя — одна из основных причин возникновения ДТП. Исследования, основанные на применении методов компьютерного зрения для выявления признаков физической усталости водителя, таких как зевота, продемонстрировали хороший потенциал при разработке систем технологий выявления таких состояний. Методы, основанные на компьютерном зрении, популярны, хорошо изучены в научном сообществе и зачастую не требовательны к вычислительным ресурсам и применимы к изображению, полученному с

видеокамеры, установленной перед водителем. Такие методы не лишены недостатков, к которым относятся условия окружающей освещенности и индивидуальные различия в цвете и форме лица человека.

В настоящей работе предложено разработать новый подход к обнаружению зевоты, основанный на распознавании признаков проявления состояния зевоты с использованием специализированных подходов к анализу и обработке данных, и сверточной нейронной сети, обученной на разных наборах данных, позволяющей выделять опорные признаки и паттерны на изображениях. Особое внимание сосредоточено на подготовке набора данных, используемого для обучения модели, и архитектуре самой модели.

Подготовка данных для обучения глубокой нейронной сети определения открытости рта человека

Входными источниками данных для задачи детектирования величины открытости рта человека послужили следующие открытые информационные ресурсы.

- 1) Набор данных YawDD: Yawning Detection Dataset:
 - 322 видео в формате RGB, записанные при частоте в 30 кадров/с; по 3–4 видео соотносится с каждым водителем (4,8 ГБ); видеокамера установлена под зеркалом заднего вида;
 - 29 видеофайла в формате RGB, записанные при частоте в 30 кадров/с; каждое видео ассоциировано с каждым водителем; камера установлена напротив водителя, на бортовой панели автомобиля;
 - три вида вождения: нормальное вождение (нет разговоров); разговор или пение; зевота;
 - мужчины и женщины за рулем автомобиля, с очками/солнцезащитными очками и без, различные этнические группы;
 - рамки лица разного размера и смещения;
 - различные условия освещения и разные погодные условия (облачно, солнечно, и др.);

— переменный угол поворота головы 0–30°;

— неразмеченный набор данных.

- 2) Набор данных Kaggle: Drowsiness Dataset¹:

— 1450 изображений;

— две категории: «зевающий»/«не зевающий»;

— мужчины и женщины за рулем автомобиля, с очками/солнцезащитными очками и без, различные этнические группы;

— переменный угол поворота головы 0–30°;

— размеченный набор данных.

Поскольку набор данных YawDD состоит только из видеофайлов, неаннотированных покaдрово, т. е. является неразмеченным, требуется классифицировать каждое изображение из видео на предмет присутствия одного из двух состояний у водителя: рот открыт и рот закрыт. Схема подготовки YawDD датасета и формирования собственного набора данных представлена на рис. 1.

На вход данному алгоритму, изображенному на схеме (рис. 1), поступает список RGB видеофайлов в разрешении 640×480 пикселей с частотой 30 кадров/с, на которых изображен водитель за рулем ТС. Данные видеофайлы записаны в двух положениях видеокамеры: со стороны заднего зеркала и в анфас (фронтально). С целью уменьшения шума на изображении и сужении области интереса изображения, конечным результатом в сформированном наборе данных должно стать изображение самого лица человека без каких-либо посторонних объектов. Чтобы выполнить поиск и обнаружение положения лица на изображении требуется воспользоваться методами компьютерной обработки изображений или возможностями глубоких нейронных сетей. Среди алгоритмов поиска объекта лица человека на изображении выбор происходит между Dlib [28], Faceboxes [29], SSD (Single Shot MultiBox Detector) [30],

¹ Drowsiness_dataset | Kaggle [Электронный ресурс]. URL: <https://www.kaggle.com/dheerajperumandla/drowsiness-dataset> (дата обращения: 06.08.2021).

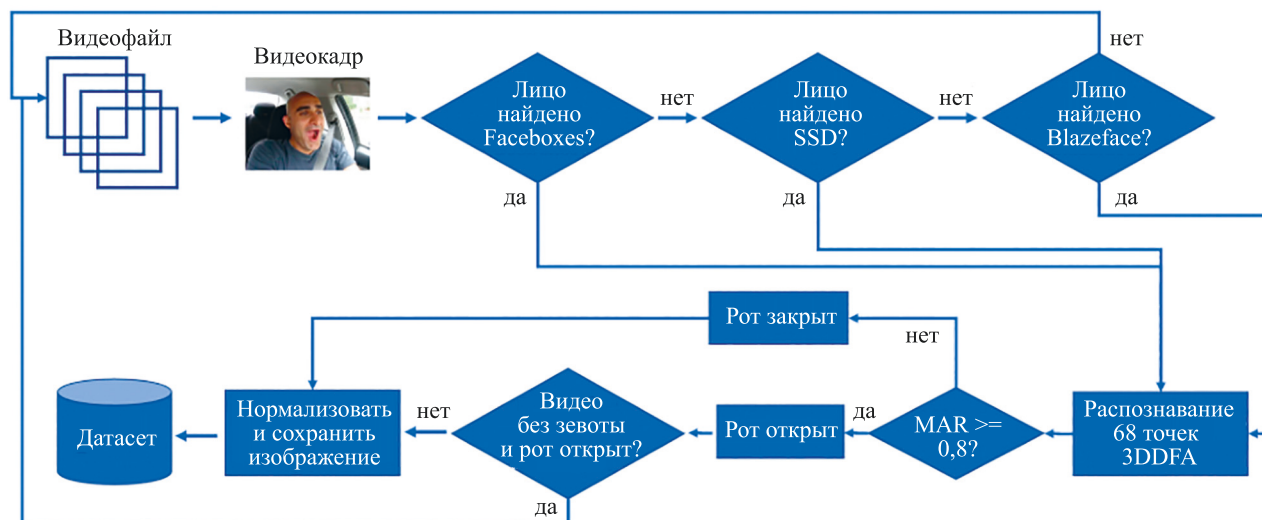


Рис. 1. Схема подготовки набора данных из YawDD

Fig. 1. Scheme of processing of the YawDD dataset

Blazeface [31]. Выполнено сравнение эффективности и точности работы вышеперечисленных алгоритмов поиска лиц на подмножестве тестовых образцов из набора данных YawDD. Алгоритм поиска Dlib показал на порядок хуже результаты по сравнению с остальными решениями. Среди недостатков фреймворка Dlib можно выделить неспособность найти лица при переменном и недостаточном освещении, при повороте головы влево или вправо, а также некорректно найденные объекты на изображении, не относящиеся к лицу. Остальные решения для поиска лица на изображении показали достаточно стабильные результаты по точности работы. Для задачи поиска лица на изображении использованы перечисленные алгоритмы, поскольку разные решения способны независимо друг от друга находить рамку лица одного и того же человека с незначительным смещением и различиями по высоте и ширине. Данные о рамке лица человека, полученные в результате работы разных алгоритмов, способны разнообразить выборку изображений и улучшить качество всего набора данных, используемого в дальнейшем для обучения нейронной сети и для предсказания открытости рта на целевом изображении. Отметим, что алгоритмы поиска лиц на изображениях исполняются поочередно для каждого последующего изображения, тем самым позволяя получить различные рамки лица. Как только лицо найдено одним из трех алгоритмов, оно передается на вход алгоритму 3D Face alignment [32], отвечающему за поиск 68 лицевых опорных точек. Такой набор точек позволяет определить положение и размеры рта на изображении и вычислить долю открытости рта MAR (mouth aspect ratio):

$$MAR = \frac{\|p_2 - p_8\| + \|p_3 - p_7\| + \|p_4 - p_6\|}{2\|p_1 - p_5\|},$$

где p_1-p_8 — лицевые точки, соответствующие расположению, показанному на рис. 2.

По результатам экспериментов на наборе данных YawDD получено пороговое значение показателя $MAR = 0,8$ для определения события открытости рта. На рис. 3 приведен пример возникновения события зевоты, характеризующейся продолжительным открытием рта на временном промежутке, состоящим из последовательности видеокадров с 55 по 110, на которой показатель MAR выше установленного порога.

Отметим, что открытость рта человека при возникновении состояния зевоты характеризуется большей величиной и иной продолжительностью непрерывной

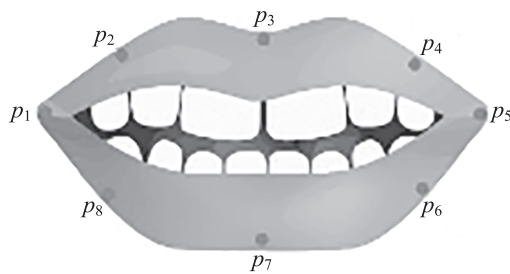


Рис. 2. Положение лицевых точек для области рта (MAR)

Fig. 2. Location of facial points related to the mouth (MAR)

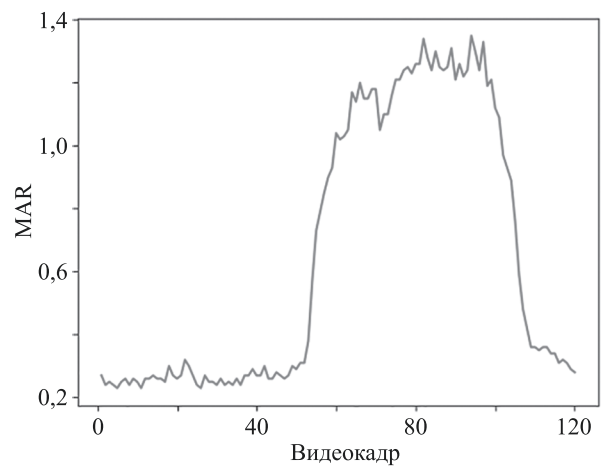


Рис. 3. Пример проявления состояния зевоты на временном промежутке

Fig. 3. Example of yawning occurrence within time period

открытости рта (рис. 3), чем при других видах активностей человека, таких как «разговор» (рис. 4) или «пение». На рис. 4 приведен пример возникновения состояния «разговор» на временном промежутке, характеризующимся непродолжительным открытием рта на данном промежутке, состоящим из последовательности видеокадров от 100 до 200, на котором показатель MAR изменяется незначительно и варьируется в интервале от 0,35 до 0,65.

Видео, полученные из набора данных YawDD, описываются очень кратко и могут только подсказать, зевает ли человек на записанном видеофрагменте или нет. Если по имеющемуся описанию видео становится понятно, что человек не зевает за рулем автомобиля, и при этом алгоритм фиксирует открытость рта у водителя, которая может определяться при разговоре с пассажиром или по телефону, такое изображение пропускается из обработки, и алгоритм переходит к следующему кадру из видео. Если же такое условие не выполняется, изображение лица человека проходит предфинальную нормализацию — переводится в формат полутонового

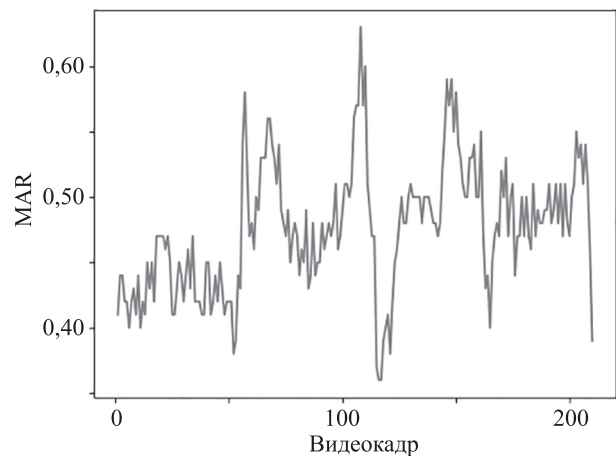


Рис. 4. Пример разговора человека на временном промежутке

Fig. 4. Example of a person talking within time period

изображения (grayscale), трансформируется в размер ширины и высоты 100×100 пикселей и сохраняется в новый полученный набор данных, запоминая данные о показателе MAR.

В результате проведенной обработки набора данных YawDD Dataset количество изображений лиц водителя с закрытым ртом составило приблизительно 49 000 изображений, а с открытым ртом — 11 000 изображений. Также в результирующий набор данных (на схеме компонент «Датасет») вошли изображения лиц водителей, представленные существующим открытым набором данных «Kaggle: Drowsiness Dataset»: 1448 изображений лица водителя, из которых 723 — с открытым ртом, и 725 изображений — с закрытым.

Концептуальная модель определения открытости рта водителя

Разработанная концептуальная модель определения открытости рта представлена на рис. 5. Подготовленные наборы примеров изображений лиц должны быть перемешаны, чтобы исключить неявный порядок следования изображений и равномерно распределить их для следующего этапа обработки. Как только наборы данных для обучения и тестирования нейронной сети готовы, на следующем этапе происходит разделение данных на три разных подмножества выборки: обучающая, проверочная и тестовая (training-validation-test). На обучающей выборке происходит первичная настройка параметров модели, ее весов и смещений посредством применения алгоритма градиентного спуска. Проверочная выборка — это та выборка, которая не принимала участие в процессе обучения модели и на которой оценивалось качество созданной модели и оптимизация ее гиперпараметров. Наборы данных, составляющих тестовую выборку, использовались для независимой оценки получившейся нейронной модели,

получения конечных результатов качества модели и выбора наилучшей модели из некоторого множества. В качестве обучающей, проверочной и тестовой выборки для нейронной модели использован единый набор данных, полученный из подготовленного на предыдущем этапе набора данных «YawDD Dataset», размеченного в полуавтоматизированном режиме, а также открытого размеченного набора данных «Kaggle Drowsiness Dataset». С учетом особенностей подготовки исходных наборов данных, а также на основе общепринятых практик в машинном обучении, наборы обучающей, проверочной и тестовой выборки были определены в соотношениях 80, 10 и 10 % соответственно.

Разработанная модель имеет структуру сверточной глубокой нейронной сети (DNN, Deep Neural Network). При разработке данной модели за основу взята архитектура модели MNIST, специализирующейся на распознавании рукописных цифр [33]. Модификация архитектуры MNIST, по большей мере, основана на корректировке весов слоев Flatten, Dense, Dropout, представляющих собой эффективный и достаточно простой способ уменьшения эффекта переобучения модели и улучшения обобщения рассматриваемых при обучении образцов изображений. Так, например, наилучший показатель для слоя Dropout составил 60 %.

Количество параметров, участвующих при обучении нейронной сети, примерно составляет 440 000. Обучение финальной версии модели выполнено на 80 эпохах (epochs) при количестве обучающих примеров за одну итерацию (batch size) — 64, а в качестве алгоритма оптимизации использована реализация Adam с шагом 0,001. Построение нейронной сети произведено при помощи фреймворка Keras совместно с TensorFlow. Функция $f(p)$ определяет выходной результат нейронной сети, описываемый вещественным числом в диапазоне от 0 до 1, где 0 — модель уверена, что рот человека закрыт; а 1 — рот открыт.

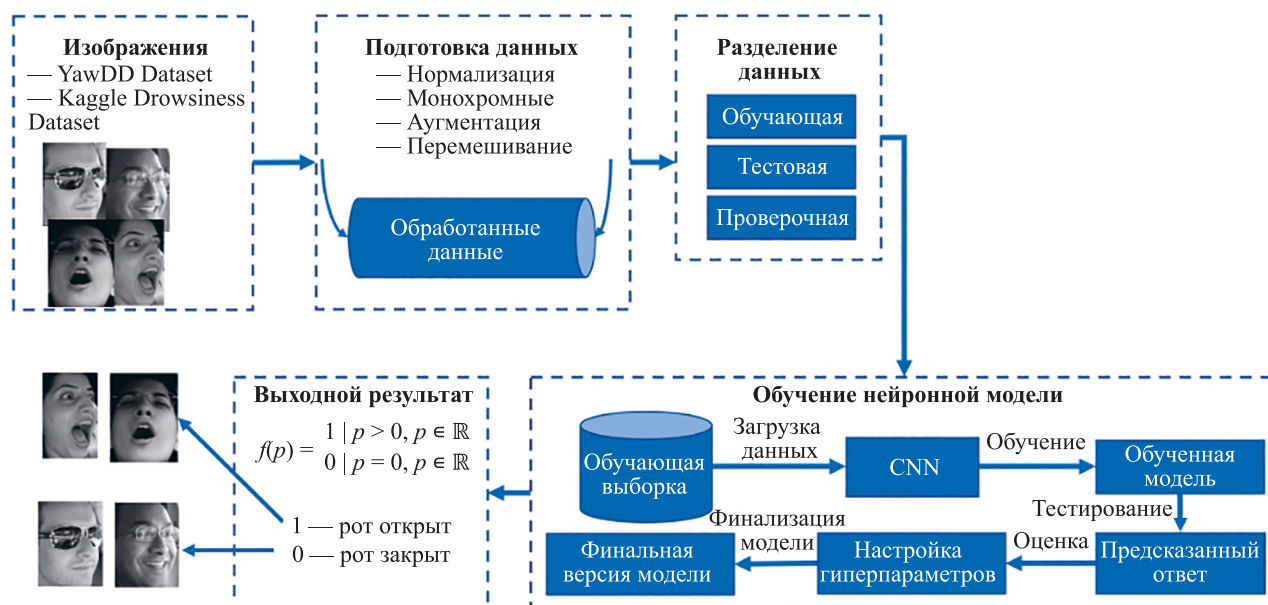


Рис. 5. Концептуальная модель обнаружения открытости/закрытости рта

Fig. 5. Reference model for the detection of mouth openness/closeness

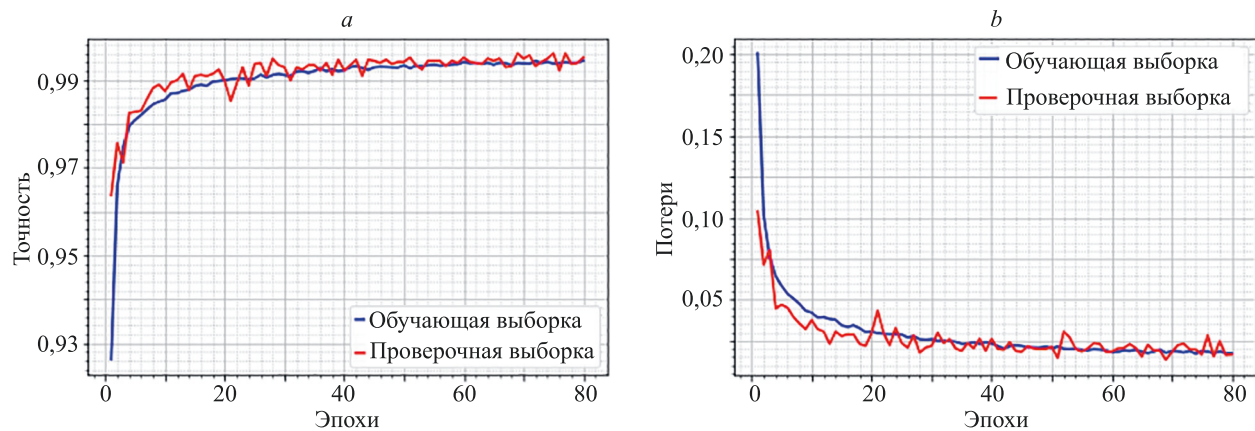


Рис. 6. Графики точности нейронной сети (a) и потерь (b) на этапе обучения

Fig. 6. Plots for accuracy of neural network (a) and loss (b) on the training stage

Эффективность обучения нейронной сети приведена на рис. 6. Можно заметить, что эффект переобучения модели сведен к минимуму, а точность (доля правильных ответов, accuracy) обучающей и проверочной выборок составили более 0,99 (рис. 6, a), а потери (loss) (рис. 6, b) около 0,02 единиц.

Оценка работы модели нейронной сети по детектированию открытости рта человека

Оценка полученной модели сверточной нейронной сети произведена на тестовой 10 % выборке, образованной набором данных YawDD и размеченным датасетом Kaggle Dataset. В табл. 1 при сравнении результатов работы модели нейронной сети на разных наборах данных использованы метрики доли правильно классифицированных объектов (accuracy), точности (precision), полноты (recall) и F-меры. Отметим, что на сегодняшний день в открытом доступе отсутствуют размеченные наборы данных, одновременно включающие как состояние зевоты, так и другие действия человека, представляющие собой «разговор», «улыбку» и другие проявления эмоций. Таким образом, оценка получившейся модели глубокого обучения на существующих наборах данных является не совсем корректной. Несмотря на это, для получения более полной оценки разработанной нейронной модели, модель была апробирована на открытом наборе данных¹, содержащем изображения лиц людей в состоянии «разговор» и «улыбка». Результаты проведенного эксперимента показали, что модель нейронной сети безошибочно детектирует отсутствие событий открытости рта, которые бы характеризовали проявление состояния зевоты для всех изображений из данного набора данных. Другой набор данных, использованный при оценке нейронной модели — Emotion Detection Dataset [34], опубликован в открытом доступе и содержит более 32 000 изображений, классифицированных по семи типам эмоций, проявляемых человеком: злость, раздражение, страх,

радость, нейтральное состояние, грусть и удивление. При проверке всех изображений на отсутствие продолжительного открытия рта, характерного для состояния зевоты, из данного набора точность бинарного классификатора составила 0,806. Полученный результат объясняется в том числе тем, что все изображения в наборе представлены в недостаточном разрешении 48×48 пикселей, а модель нейронной сети обучена на изображениях лиц людей в разрешении 100×100 пикселей.

Из табл. 2 видно, что полученная модель нейронной сети превосходит по метрикам оценки качества модель, опубликованную в работе [11]. Традиционный подход, основанный на алгоритмах компьютерного зрения, к распознаванию доли открытости рта по 68 лицевым координатным точкам [35] показывает значительно хуже результат по всем используемым статистическим метрикам — точность, полнота, F-мера. Проведено сравнение с подходами, базирующимися на методах машинного обучения — на методе опорных векторов (Support vector machines, SVM) [36, 37]. Результаты оценки метода [36] показывают, что метод опорных векторов значительно превосходит по качеству подход [35], но не достигает результатов, полученных новой разработанной нейронной моделью.

Размер нейронной сети на файловом диске составил 1,7 МБ в формате ONNX. Исходный код нейронной модели, а также данные и остальные результаты сравнения моделей опубликованы в открытом репозитории Github².

Оценка скорости работы разработанной нейронной модели проведена на персональном ноутбуке, и на смартфонах различной мощности. Сравнение характеристик быстродействия работы полученной модели в различных режимах и настройках, включающих конфигурацию целевого устройства, используемый акселератор, формат представления нейронной модели, среднее время работы модели на устройстве и версию фреймворка машинного обучения для исполнения модели, приведено в табл. 3. Заметим, что время работы

¹ Talking Face Video [Электронный ресурс]. URL: https://personalpages.manchester.ac.uk/staff/timothy.f.cootes/data/talking_face/talking_face.html (дата обращения: 11.10.2021).

² HippoYD | Github [Электронный ресурс]. URL: <https://github.com/iglaweb/HippoYD> (дата обращения: 09.10.2021).

Таблица 1. Сравнение результатов работы модели нейронной сети на разных наборах данных
Table 1. Comparison of the results obtained from the neural network model on different datasets

Набор данных	Отрицательный класс («рот закрыт»)	Положительный класс («рот открыт»)	Метрика			
			Доля правильно классифицированных объектов	Точность	Полнота	F-мера
Тестовая выборка на основе YawDD и Kaggle Drowsiness Dataset	4900	1100	0,932	0,992	0,871	0,927
Talking Face Video	5 000	—	1,000	—	—	—
Emotion Detection Dataset [34]	32 000	—	0,806	—	—	—

Таблица 2. Сравнение работы предложенной модели нейронной сети и других подходов к детектированию открытости рта у человека
Table 2. Comparison of the neural network model and other approaches aimed at detecting mouth openness of a person

Модель	Метрика			
	Доля правильно классифицированных объектов	Точность	Полнота	F-мера
Предложенная в настоящей работе	0,932	0,992	0,871	0,927
Предыдущее исследование авторов настоящей работы [11]	0,740	0,910	0,540	0,680
Традиционный подход [35]	0,280	0,317	0,384	0,347
Метод опорных векторов [36]	0,803	0,921	0,662	0,770

Таблица 3. Сравнение быстродействия модели нейронной сети
Table 3. Comparison of model performance of a neural network model

Конфигурация	Акселератор	Формат модели	Среднее время, мс	TensorFlow/TFLite версия
Macbook Pro, CPU 2 GHz Quad-Core Intel Core i5	CPU	TFLite (Floating)	5	2.3
	CPU	TFLite (Quantized)	8	2.3
	CPU	Keras H5 (Floating)	30	2.3
	CPU	ONNX	2	2.3
	CPU	Frozen pb	4	2.3
	Wasm (Safari 14.0, Firefox 84)	TensorFlowJS	30	2.3
Смартфон Xiaomi MI8	GPU/CPU 3 потока	TFLite (Floating)	4	2.4
	CPU 3 потока	TFLite (Quantized)	10	2.4
Смартфон Xiaomi Redmi 9	GPU/CPU 3 потока	TFLite (Floating)	11	2.4
	CPU 3 потока	TFLite (Quantized)	9	2.4
Смартфон Samsung Galaxy S20+	GPU/CPU 3 потока	TFLite (Floating)	2	2.4
	CPU 3 потока	TFLite (Quantized)	2	2.4

нейронной модели на смартфоне редко составляет более 10 мс, что, в свою очередь, не превышает лимит количества обрабатываемых кадров в 30 кадров/с.

Схема детектора выявления признаков состояния зевоты по видеоряду с камеры

Определение состояния зевоты целесообразно рассматривать не на одном кадре, а с учетом последовательности изображений, полученных с видеокамеры.

В качестве следующего шага после построения нейронной модели детектирования открытости рта, предложена схема детектора выявления признаков состояния зевоты у водителя ТС, показанная на рис. 7. Данная схема представляет собой демонстрационный пример детектирования состояния зевоты, при условии, что характеристики зевоты (вероятность открытости рта $p_1, p_2, \dots, p_k$) могут быть распределены для каждого водителя и в каждый момент времени ($t_{e1}, t_{e2}, \dots, t_{ei}$) по-разному. Состояние зевоты человека принято рас-

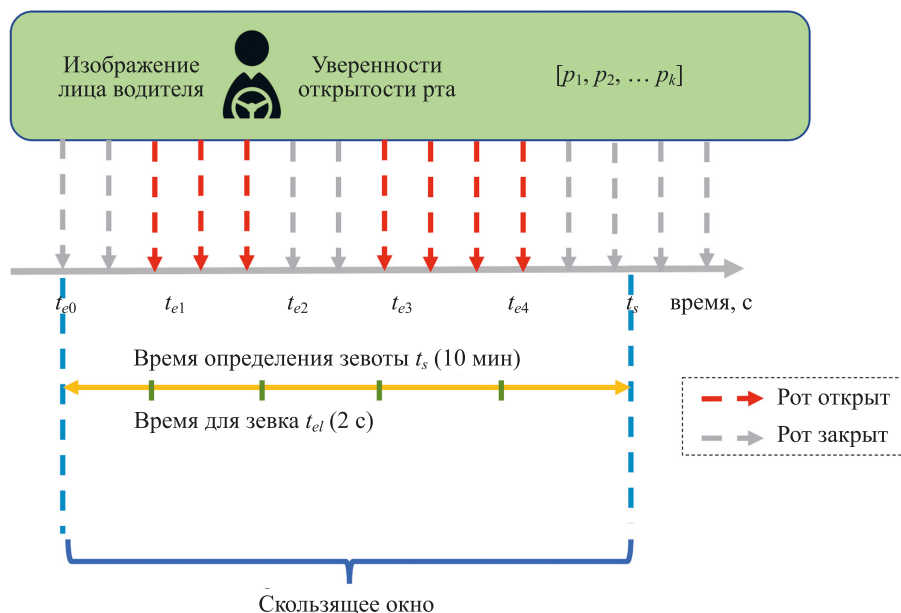


Рис. 7. Схема детектора выявления признаков состояния зевоты
Fig. 7. Scheme for the detection of yawning

смаивать на промежутке t_s равном 10 мин. В течение этого промежутка времени с использованием алгоритма скользящего окна [38] определяется число зевков, совершенных водителем. Если водитель совершает три зевка в течение данного интервала, фиксируется состояние зевоты, и водитель может быть оповещен о таком опасном состоянии. На основе существующих научных исследований время определения одного зевка t_{el} выбрано в промежутке до 2,5 с. При условии, что рот водителя был открыт более 80 % времени в рамках данного промежутка времени, принимается решение, что водитель совершил один зевок. Открытость рта водителя определяется на основе ранее представленной нейронной модели глубокого обучения, выходным значением для которой является уверенность открытости рта.

В качестве эксперимента по анализу эффективности работы предложенной схемы, она была апробирована на тестовой выборке видео, сформированной из набора данных YawDD. Общее количество видеофайлов, на которых отмечен хотя бы один зевок, составляет 90 видеофайлов. Размер тестовой выборки равен 35 видеофайлам, что составляет 10 % от количества всех видеофайлов в исходном наборе. Общее количество зевков, которые совершили водители, составило 128 раз. Данная схема детектирования успешно обнаружила 124 зевка среди всех видеофайлов из тестового набора

данных. Доля правильно классифицированных объектов составила 98,2 %, точность — 96,1 %, полнота — 98,4 %, а F-мера — 97,3 % при определении признаков состояния зевоты у водителя ТС.

С целью сглаживания результатов работы нейронной сети, другим вариантом выявления признаков состояния зевоты на промежутке времени является анализ последовательности видеокадров с использованием алгоритма скользящего среднего. Суть подхода заключается в усреднении вероятностей принадлежности изображений к одному из классов, «рот открыт» и «рот закрыт» работы модели нейронной сети, на промежутке времени. Размер скользящего окна, включающего набор усредняемых значений, определен в 2 с и включает все видеокадры, вошедшие в данный период времени. Такой размер окна для детектирования состояния зевоты выбран на основе проанализированных ранее исследований. В результате работы данного алгоритма изображение относится к классу «рот открыт» или «рот закрыт», усредненное значение вероятностей которого является максимальным в рассматриваемый момент времени. Данный алгоритм также был протестирован на наборе данных YawDD: Yawning Detection Dataset. Доля правильно классифицированных объектов достигла 94,7 %, точность — 86 %, полнота — 97,4 %, а F-мера — 91,4 % при определении признаков состояния зевоты.

Таблица 4. Сравнение подходов к детектированию состояния зевоты, %
Table 4. Comparison of approaches of recognizing yawning states

Подход	Метрика			
	Правильно классифицированные объекты	Точность	Полнота	F-мера
Схема детектирования зевоты	98,2	96,1	98,4	97,3
Алгоритм скользящего среднего	94,7	86,0	97,4	91,4

Исходя из оценки результатов работы обоих алгоритмов детектирования состояния зевоты (табл. 4) видно, что первоначальная схема детектирования зевоты, описанная на рис. 7, превосходит по всем перечисленным статистическим показателям алгоритм скользящего среднего.

Заключение

В работе предложена схема эффективного выявления признаков зевоты у водителя в кабине транспортного средства с использованием методов анализа и обработки данных, компьютерной обработки изображений и сверточной модели нейронной сети. В рамках данного подхода представлена концептуальная модель обнаружения открытости/закрытости рта у человека за рулем транспортного средства, и разработана схема обработки и разметки наборов данных YawDD и Kaggle Drowsiness Dataset. Разработанная модель сверточной нейронной сети показала точность (precision) равную 0,992 и полноту (recall) — 0,871 на тестовом 10 % наборе данных, что является приемлемым результатом для дальнейшей интеграции и апробации в интеллектуальной системе поддержки водителю в реальных ус-

ловиях вождения. Представленные результаты работы по совмещению данных из разных наборов способны найти применение при построении технологии детектирования признаков зевоты для систем поддержки водителя и улучшить эффективность их работы. В качестве дальнейших перспектив по улучшению полученных результатов можно рассмотреть возможность разработки гибридного подхода, основанного на одновременном использовании анализа открытости/закрытости рта, состояния глаз человека, и оценки его позы в пространстве при выявлении зевоты и состояния сонливости в целом.

Физиологическое состояние водителя позволяет проанализировать и оценить текущую обстановку в кабине транспортного средства. Присутствие признаков небезопасного поведения человека за рулем может потенциально снизить безопасность для водителя и других участников дорожного движения. В целях повышения безопасности и экономических факторов информация об управлении транспортным средством способна найти применение в работе страховых и логистических компаний, специализирующихся на грузоперевозках, а также в области пассажирских перевозок.

Литература

1. Rau P.S. Drowsy driver detection and warning system for commercial vehicle drivers: Field operational test design, data analyses, and progress // *Proc. Nat. Highway Traffic Saf. Admin.* 2005. P. 05–0192.
2. Dua M., Singla R., Singla R., Raj S., Jangra A. Deep CNN models-based ensemble approach to driver drowsiness detection // *Neural Computing and Applications*. 2021. V. 33. N 8. P. 3155–3168. <https://doi.org/10.1007/s00521-020-05209-7>
3. Bakheet S., Al-Hamadi A. A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve bayesian classification // *Brain Sciences*. 2021. V. 11. N 2. P. 240. <https://doi.org/10.3390/brainsci11020240>
4. Li L., Chen Y., Li Z. Yawning detection for monitoring driver fatigue based on two cameras // *Proc. 12th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. 2009. P. 12–17. <https://doi.org/10.1109/ITSC.2009.5309841>
5. Daquin G., Micallef J., Blin O. Yawning // *Sleep Medicine Reviews*. 2001. V. 5. N 4. P. 299–312. <https://doi.org/10.1053/smr.2001.0175>
6. Jackson P., Hilditch C., Holmes A., Reed N., Merat N., Smith L. *Fatigue and Road Safety: A Critical Analysis of Recent Evidence*. Department for Transport, London, 2011. 88 p.
7. Resendes R., Martin K.H. *Saving Lives Through Advanced Safety Technology*. Federal Highway Administration, US, Washington, DC, 2003.
8. Wang L., Sun P., Xie M., Ma S., Li B., Shi Y., Su Q. Advanced driver-assistance system (ADAS) for intelligent transportation based on the recognition of traffic cones // *Advances in Civil Engineering*. 2020. P. 8883639. <https://doi.org/10.1155/2020/8883639>
9. Zhang W., Murphey Y.L., Wang T., Xu Q. Driver yawning detection based on deep convolutional neural learning and robust nose tracking // *Proc. of the International Joint Conference on Neural Networks (IJCNN)*. 2015. P. 1–8. <https://doi.org/10.1109/IJCNN.2015.7280566>
10. Chevalier Y., Fenzl F., Kolomeets M., Rieke R., Chechulin A., Krauss C. Cyberattack detection in vehicles using characteristic functions, artificial neural networks, and visual analysis // *Информатика и автоматизация*. 2021. T. 20. № 4. С. 845–868. <https://doi.org/10.15622/ia.20.4.4>
11. Hasan F., Kashevnik A. State-of-the-Art Analysis of modern drowsiness detection algorithms based on computer vision // *Proc. of the 29th Conference of Open Innovations Association (FRUCT)*. 2021. P. 141–149. <https://doi.org/10.23919/FRUCT52173.2021.9435480>
12. Kashevnik A., Lashkov I., Parfenov V., Mustafin N., Baraniuc O. Context-based driver support system development: Methodology and

References

1. Rau P.S. Drowsy driver detection and warning system for commercial vehicle drivers: Field operational test design, data analyses, and progress. *Proc. Nat. Highway Traffic Saf. Admin.*, 2005, pp. 05–0192.
2. Dua M., Singla R., Singla R., Raj S., Jangra A. Deep CNN models-based ensemble approach to driver drowsiness detection. *Neural Computing and Applications*, 2021, vol. 33, no. 8, pp. 3155–3168. <https://doi.org/10.1007/s00521-020-05209-7>
3. Bakheet S., Al-Hamadi A. A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve bayesian classification. *Brain Sciences*, 2021, vol. 11, no. 2, pp. 240. <https://doi.org/10.3390/brainsci11020240>
4. Li L., Chen Y., Li Z. Yawning detection for monitoring driver fatigue based on two cameras. *Proc. 12th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2009, pp. 12–17. <https://doi.org/10.1109/ITSC.2009.5309841>
5. Daquin G., Micallef J., Blin O. Yawning. *Sleep Medicine Reviews*, 2001, vol. 5, no. 4, pp. 299–312. <https://doi.org/10.1053/smr.2001.0175>
6. Jackson P., Hilditch C., Holmes A., Reed N., Merat N., Smith L. *Fatigue and Road Safety: A Critical Analysis of Recent Evidence*. Department for Transport, London, 2011, 88 p.
7. Resendes R., Martin K.H. *Saving Lives Through Advanced Safety Technology*. Federal Highway Administration, US, Washington, DC, 2003.
8. Wang L., Sun P., Xie M., Ma S., Li B., Shi Y., Su Q. Advanced driver-assistance system (ADAS) for intelligent transportation based on the recognition of traffic cones. *Advances in Civil Engineering*, 2020, pp. 8883639. <https://doi.org/10.1155/2020/8883639>
9. Zhang W., Murphey Y.L., Wang T., Xu Q. Driver yawning detection based on deep convolutional neural learning and robust nose tracking. *Proc. of the International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1–8. <https://doi.org/10.1109/IJCNN.2015.7280566>
10. Chevalier Y., Fenzl F., Kolomeets M., Rieke R., Chechulin A., Krauss C. Cyberattack detection in vehicles using characteristic functions, artificial neural networks, and visual analysis. *Informatics and Automation*, 2021, vol. 20, no. 4, pp. 845–868. <https://doi.org/10.15622/ia.20.4.4>
11. Hasan F., Kashevnik A. State-of-the-Art Analysis of modern drowsiness detection algorithms based on computer vision. *Proc. of the 29th Conference of Open Innovations Association (FRUCT)*, 2021, pp. 141–149. <https://doi.org/10.23919/FRUCT52173.2021.9435480>

- case study // *Proc. of the 21st Conference of Open Innovations Association (FRUCT)*. 2017. P. 162–171. <https://doi.org/10.23919/FRUCT.2017.8250179>
13. Kashevnik A., Lashkov I., Gurtov A. Methodology and mobile application for driver behavior analysis and accident prevention // *IEEE Transactions on Intelligent Transportation Systems*. 2020. V. 21. N 6. P. 2427–2436. <https://doi.org/10.1109/TITS.2019.2918328>
 14. Ibrahim M.M., Soraghan J.J., Petropoulakis L., Di Caterina G. Yawn analysis with mouth occlusion detection // *Biomedical Signal Processing and Control*. 2015. V. 18. P. 360–369. <https://doi.org/10.1016/j.bspc.2015.02.006>
 15. Ji Y., Wang S., Lu Y., Wei J., Zhao Y. Eye and mouth state detection algorithm based on contour feature extraction // *Journal of Electronic Imaging*. 2018. V. 27. N 5. P. 051205. <https://doi.org/10.1117/1.JEI.27.5.051205>
 16. Omidyeganeh M., Shirmohammadi S., Abtahi S., Khurshid A., Farhan M., Scharcanski J., Hariri B., Laroche D., Martel L. Yawning detection using embedded smart cameras // *IEEE Transactions on Instrumentation and Measurement*. 2016. V. 65. N 3. P. 570–582. <https://doi.org/10.1109/TIM.2015.2507378>
 17. Khan S., Akram A., Usman N. Real time automatic attendance system for face recognition using face API and OpenCV // *Wireless Personal Communications*. 2020. V. 113. N 1. P. 469–480. <https://doi.org/10.1007/s11277-020-07224-2>
 18. Jose J., Vimali J.S., Ajitha P., Gowri S., Sivasangari A., Jinila B. Drowsiness detection system for drivers using image processing technique // *Proc. of the 5th International Conference on Trends in Electronics and Informatics (ICOEI)*. 2021. P. 1527–1530. <https://doi.org/10.1109/ICOEI51242.2021.9452864>
 19. Al-Madani A.M., Gaikwad A.T., Mahale V., Ahmed Z.A., Shareef A.A.A. Real-time driver drowsiness detection based on eye movement and yawning using facial landmark // *Proc. of the International Conference on Computer Communication and Informatics (ICCCI)*. 2021. P. 1–4. <https://doi.org/10.1109/ICCCI50826.2021.9457005>
 20. Akrouf B., Mahdi W. Yawning detection by the analysis of variational descriptor for monitoring driver drowsiness // *Proc. of the 2nd International Image Processing, Applications and Systems Conference (IPAS)*. 2016. P. 1–5. <https://doi.org/10.1109/IPAS.2016.7880127>
 21. Zharmagambetov A., Gabidolla M., Carreira-Perpinán M.A. Improved multiclass AdaBoost for image classification: The role of tree optimization // *Proc. of the IEEE International Conference on Image Processing (ICIP)*. 2021. P. 19–22. <https://doi.org/10.1109/ICIP42928.2021.9506569>
 22. Ковшов Е.Е., Завистовская Т.А. Разработка автоматизированной системы контроля и управления доступом на основе анализа динамики носогубных мышц лица человека // *Современные проблемы науки и образования*. 2013. № 2. С. 185.
 23. Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models // *International Journal of Computer Vision*. 1988. V. 1. N 4. P. 321–331. <https://doi.org/10.1007/BF00133570>
 24. Yazdi Z.M., Soryani M. Driver drowsiness detection by identification of yawning and eye closure // *Automotive Science and Engineering*. 2019. V. 9. N 3. P. 3033–3044. <https://doi.org/10.22068/ijae.9.3.3033>
 25. Kassem H.A., Chowdhury M.U., Abawajy J., Al-Sudani A.R. Yawn based driver fatigue level prediction // *EPiC Series in Computing*. 2020. V. 69. P. 372–382. <https://doi.org/10.29007/67kk>
 26. Abtahi S., Omidyeganeh M., Shirmohammadi S., Hariri B. YawDD: A yawning detection dataset // *Proc. of the 5th ACM Multimedia Systems Conference*. 2014. P. 24–28. <https://doi.org/10.1145/2557642.2563678>
 27. Yang H., Liu L., Min W., Yang X., Xiong X. Driver yawning detection based on subtle facial action recognition // *IEEE Transactions on Multimedia*. 2021. V. 23. P. 572–583. <https://doi.org/10.1109/TMM.2020.2985536>
 28. Suwarno S., Kevin K. Analysis of face recognition algorithm: Dlib and openCV // *Journal of Informatics and Telecommunication Engineering*. 2020. V. 4. N 1. P. 173–184. <https://doi.org/10.31289/jite.v4i1.3865>
 29. Zhang S., Zhu X., Lei Z., Shi H., Wang X., Li S.Z. FaceBoxes: A CPU real-time face detector with high accuracy // *Proc. of the IEEE International Joint Conference on Biometrics (IJCB)*. 2017. P. 1–9. <https://doi.org/10.1109/BTAS.2017.8272675>
 30. Zhang K., Zhang Z., Li Z., Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks // *IEEE Signal*
 12. Kashevnik A., Lashkov I., Parfenov V., Mustafin N., Baraniuc O. Context-based driver support system development: Methodology and case study. *Proc. of the 21st Conference of Open Innovations Association (FRUCT)*, 2017, pp. 162–171. <https://doi.org/10.23919/FRUCT.2017.8250179>
 13. Kashevnik A., Lashkov I., Gurtov A. Methodology and mobile application for driver behavior analysis and accident prevention. *IEEE Transactions on Intelligent Transportation Systems*, 2020, vol. 21, no. 6, pp. 2427–2436. <https://doi.org/10.1109/TITS.2019.2918328>
 14. Ibrahim M.M., Soraghan J.J., Petropoulakis L., Di Caterina G. Yawn analysis with mouth occlusion detection. *Biomedical Signal Processing and Control*, 2015, vol. 18, pp. 360–369. <https://doi.org/10.1016/j.bspc.2015.02.006>
 15. Ji Y., Wang S., Lu Y., Wei J., Zhao Y. Eye and mouth state detection algorithm based on contour feature extraction. *Journal of Electronic Imaging*, 2018, vol. 27, no. 5, pp. 051205. <https://doi.org/10.1117/1.JEI.27.5.051205>
 16. Omidyeganeh M., Shirmohammadi S., Abtahi S., Khurshid A., Farhan M., Scharcanski J., Hariri B., Laroche D., Martel L. Yawning detection using embedded smart cameras. *IEEE Transactions on Instrumentation and Measurement*, 2016, vol. 65, no. 3, pp. 570–582. <https://doi.org/10.1109/TIM.2015.2507378>
 17. Khan S., Akram A., Usman N. Real time automatic attendance system for face recognition using face API and OpenCV. *Wireless Personal Communications*, 2020, vol. 113, no. 1, pp. 469–480. <https://doi.org/10.1007/s11277-020-07224-2>
 18. Jose J., Vimali J.S., Ajitha P., Gowri S., Sivasangari A., Jinila B. Drowsiness detection system for drivers using image processing technique. *Proc. of the 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, 2021, pp. 1527–1530. <https://doi.org/10.1109/ICOEI51242.2021.9452864>
 19. Al-Madani A.M., Gaikwad A.T., Mahale V., Ahmed Z.A., Shareef A.A.A. Real-time driver drowsiness detection based on eye movement and yawning using facial landmark. *Proc. of the International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1–4. <https://doi.org/10.1109/ICCCI50826.2021.9457005>
 20. Akrouf B., Mahdi W. Yawning detection by the analysis of variational descriptor for monitoring driver drowsiness. *Proc. of the 2nd International Image Processing, Applications and Systems Conference (IPAS)*, 2016, pp. 1–5. <https://doi.org/10.1109/IPAS.2016.7880127>
 21. Zharmagambetov A., Gabidolla M., Carreira-Perpinán M.A. Improved multiclass AdaBoost for image classification: The role of tree optimization. *Proc. of the IEEE International Conference on Image Processing (ICIP)*, 2021, pp. 19–22. <https://doi.org/10.1109/ICIP42928.2021.9506569>
 22. Kovshov E.E., Zavistovskaya T.A. Development of automated system of an access control based on analysis of the nasolabial muscles dynamics of the human face. *Modern problems of science and education*, 2013, no. 2, pp. 185–185. (in Russian)
 23. Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision*, 1988, vol. 1, no. 4, pp. 321–331. <https://doi.org/10.1007/BF00133570>
 24. Yazdi Z.M., Soryani M. Driver drowsiness detection by identification of yawning and eye closure. *Automotive Science and Engineering*, 2019, vol. 9, no. 3, pp. 3033–3044. <https://doi.org/10.22068/ijae.9.3.3033>
 25. Kassem H.A., Chowdhury M.U., Abawajy J., Al-Sudani A.R. Yawn based driver fatigue level prediction. *EPiC Series in Computing*, 2020, vol. 69, pp. 372–382. <https://doi.org/10.29007/67kk>
 26. Abtahi S., Omidyeganeh M., Shirmohammadi S., Hariri B. YawDD: A yawning detection dataset. *Proc. of the 5th ACM Multimedia Systems Conference*, 2014, pp. 24–28. <https://doi.org/10.1145/2557642.2563678>
 27. Yang H., Liu L., Min W., Yang X., Xiong X. Driver yawning detection based on subtle facial action recognition. *IEEE Transactions on Multimedia*, 2021, vol. 23, pp. 572–583. <https://doi.org/10.1109/TMM.2020.2985536>
 28. Suwarno S., Kevin K. Analysis of face recognition algorithm: Dlib and openCV. *Journal of Informatics and Telecommunication Engineering*, 2020, vol. 4, no. 1, pp. 173–184. <https://doi.org/10.31289/jite.v4i1.3865>
 29. Zhang S., Zhu X., Lei Z., Shi H., Wang X., Li S.Z. FaceBoxes: A CPU real-time face detector with high accuracy. *Proc. of the IEEE International Joint Conference on Biometrics (IJCB)*, 2017, pp. 1–9. <https://doi.org/10.1109/BTAS.2017.8272675>

- Processing Letters. 2016. V. 23. N 10. P. 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>
31. Bazarevsky V., Kartynnik Y., Vakunov A., Raveendran K., Grundmann M. Blazeface: Sub-millisecond neural face detection on mobile gpus // arXiv.org. 2019. arXiv:1907.05047.
 32. Bulat A., Tzimiropoulos G. How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230,000 3d facial landmarks) // Proc. of the 16th IEEE International Conference on Computer Vision (ICCV). 2017. P. 1021–1030. <https://doi.org/10.1109/ICCV.2017.116>
 33. Cohen G., Afshar S., Tapson J., Van Schaik A. EMNIST: Extending MNIST to handwritten letters // Proc. of the International Joint Conference on Neural Networks (IJCNN). 2017. P. 2921–2926. <https://doi.org/10.1109/IJCNN.2017.7966217>
 34. Goodfellow I.J., Erhan D., Carrier P.L., Courville A., Mirza M., Hamner B., Cukierski W., Tang Y., Thaler D., Lee D., Zhou Y., Ramaiah C., Feng F., Li R., Wang X., Athanasakis D., Shave-Taylor J., Milakov M., Park J., Ionescu R., Popescu M., Grozea C., Bergstra J., Xie J., Romaszko L., Xu B., Chuang Z., Bengio Y. Challenges in representation learning: A report on three machine learning contests // Lecture Notes in Computer Science. 2013. V. 8228. P. 117–124. https://doi.org/10.1007/978-3-642-42051-1_16
 35. Ramos A.L., Erandio J.C., Enteria E.M., Del Carmen N., Enriquez L.J., Mangilaya D.H. Driver drowsiness detection based on eye movement and yawning using facial landmark analysis // International Journal of Simulation: Systems, Science & Technology. 2019. V. 20. N S2. P. 37. <https://doi.org/10.5013/IJSSST.a.20.S2.37>
 36. Savaş B.K., Becerikli Y. Real time driver fatigue detection based on SVM algorithm // Proc. of the 6th International Conference on Control Engineering and Information Technology (CEIT). 2018. P. 1–4. <https://doi.org/10.1109/CEIT.2018.8751886>
 37. Saradadevi M., Bajaj P. Driver fatigue detection using mouth and yawning analysis // International Journal of Computer Science and Network Security. 2008. V. 8. N 6. P. 183–188.
 38. Castella F.R. Sliding window detection probabilities // IEEE Transactions on Aerospace and Electronic Systems. 1976. V. AES-12. N 6. P. 815–819. <https://doi.org/10.1109/TAES.1976.308363>
 39. Zhang K., Zhang Z., Li Z., Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 2016, vol. 23, no. 10, pp. 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>
 40. Bazarevsky V., Kartynnik Y., Vakunov A., Raveendran K., Grundmann M. Blazeface: Sub-millisecond neural face detection on mobile gpus. *arXiv.org*, 2019, arXiv:1907.05047.
 41. Bulat A., Tzimiropoulos G. How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230,000 3d facial landmarks). *Proc. of the 16th IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 1021–1030. <https://doi.org/10.1109/ICCV.2017.116>
 42. Cohen G., Afshar S., Tapson J., Van Schaik A. EMNIST: Extending MNIST to handwritten letters. *Proc. of the International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 2921–2926. <https://doi.org/10.1109/IJCNN.2017.7966217>
 43. Goodfellow I.J., Erhan D., Carrier P.L., Courville A., Mirza M., Hamner B., Cukierski W., Tang Y., Thaler D., Lee D., Zhou Y., Ramaiah C., Feng F., Li R., Wang X., Athanasakis D., Shave-Taylor J., Milakov M., Park J., Ionescu R., Popescu M., Grozea C., Bergstra J., Xie J., Romaszko L., Xu B., Chuang Z., Bengio Y. Challenges in representation learning: A report on three machine learning contests. *Lecture Notes in Computer Science*, 2013, vol. 8228, pp. 117–124. https://doi.org/10.1007/978-3-642-42051-1_16
 44. Ramos A.L., Erandio J.C., Enteria E.M., Del Carmen N., Enriquez L.J., Mangilaya D.H. Driver drowsiness detection based on eye movement and yawning using facial landmark analysis. *International Journal of Simulation: Systems, Science & Technology*, 2019, vol. 20, no. S2, pp. 37. <https://doi.org/10.5013/IJSSST.a.20.S2.37>
 45. Savaş B.K., Becerikli Y. Real time driver fatigue detection based on SVM algorithm. *Proc. of the 6th International Conference on Control Engineering and Information Technology (CEIT)*, 2018, pp. 1–4. <https://doi.org/10.1109/CEIT.2018.8751886>
 46. Saradadevi M., Bajaj P. Driver fatigue detection using mouth and yawning analysis. *International Journal of Computer Science and Network Security*, 2008, vol. 8, no. 6, pp. 183–188.
 47. Castella F.R. Sliding window detection probabilities. *IEEE Transactions on Aerospace and Electronic Systems*, 1976, vol. AES-12, no. 6, pp. 815–819. <https://doi.org/10.1109/TAES.1976.308363>

Автор

Лашков Игорь Борисович — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), Санкт-Петербург, 199178, Российская Федерация, orcid.org/0000-0001-6418-4660, igor-lashkov@ya.ru

Author

Igor B. Lashkov — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, orcid.org/0000-0001-6418-4660, igor-lashkov@ya.ru

Статья поступила в редакцию 09.08.2021
Одобрена после рецензирования 30.11.2021
Принята к печати 21.01.2022

Received 09.08.2021
Approved after reviewing 30.11.2021
Accepted 21.01.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»