

ПЕРСОНАЛИЗАЦИЯ ПОЛЬЗОВАТЕЛЬСКИХ ДАННЫХ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ СИСТЕМЫ ПОМОЩИ ВОДИТЕЛЮ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Фам Туан Ань

Санкт-Петербургский национальный
исследовательский университет ИТМО
anhpt@itmo.ru

PERSONALIZATION OF USER DATA TO INCREASE THE ACCURACY OF THE DRIVER ASSISTANCE SYSTEM BASED ON MACHINE LEARNING

Pham Tuan Anh

Summary. This paper discusses a machine learning approach to personalizing user data to improve the accuracy of a driver assistance system that aims to adapt to driver preferences, driving styles, skills, and driving patterns. Driving sleep warning has become a mandatory feature in modern automotive systems. EAR methods for determining the state of eye openness are complex and inaccurate for people with small and narrow eyes, as, for example, in most Asians. Therefore, the author proposes a system for personalizing user data based on machine learning to improve the accuracy of determining the state of closed/open eyes of Asian faces. The system uses the front camera of the smartphone to collect user data and operate in real time.

Keywords: driver assistance system, personalization of user data, drowsiness, system accuracy, machine learning.

Аннотация. В данной работе рассматривается подход к персонализации пользовательских данных для повышения точности системы помощи водителю на основе машинного обучения, который нацелен на адаптацию к предпочтениям водителей, стилям вождения, навыкам и моделям вождения. Предупреждение о сне при вождении стало обязательной функцией в современных автомобильных системах. Методы же EAR для определения состояния открытости глаз сложны и не точны для людей с маленькими и узкими глазами, как, например, у большинства азиатов. Поэтому автор предлагает систему персонализации пользовательских данных на основе машинного обучения для повышения точности определения состояния закрытости/открытости глаз азиатских лиц. Система использует фронтальную камеру смартфона для сбора пользовательских данных и работы в режиме реального времени.

Ключевые слова: система помощи водителю, персонализация пользовательских данных, сонливость, точность системы, машинное обучение.

Введение

Современные системы помощи водителю развиваются в направлении все более сложных функций помощи, применяемых с более широким охватом и значительно увеличивающимся числом возможных пользователей в связи с более широким проникновением на рынок. Чтобы справиться с таким большим разнообразием условий использования и моделей использования, были разработаны методы персонализации, обеспечивающие оптимальный пользовательский опыт и поставляемую системную функцию.

Усталость всегда опасна для водителей и безопасности дорожного движения. Недостаток сна, приводит к сонливости во время вождения, легко может стать причиной несчастного случая со смертельным исходом. Но вождение не является основной причиной устало-

сти, потому что водители часто устали от долгой работы, недосыпания вообще и недосыпания из-за апноэ во сне в частности или сменной работы и т.д.

Вождение в состоянии усталости может привести к более медленной реакции организма, отсутствию концентрации, ошибкам в вычислении скорости и расстояния, снижению бдительности и неадекватности суждений, а также сонливости — за секунды, и это может привести к катастрофическим последствиям.

Были приняты правила, касающиеся обязательного оснащения европейской системы предупреждения о безопасном вождении стандартами безопасности дорожного движения. Эти системы требуют специального оборудования, обладают высокой точностью и применимы к новым автомобилям. В случае активных транспортных средств водитель может легко получить доступ

к системе предупреждения о сонливости с помощью смартфона, который может обеспечить раннее предупреждение об усталости.

В рамках данного исследования рассматриваются следующие вопросы:

- ◆ Какие существуют подходы машинного обучения для классификации состояния лица водителя?
- ◆ Как персонализировать характеристики лица пользователя?
- ◆ Роль персонифицированного распознавания лиц в современных системах?
- ◆ Какие методы используются для определения сонливости человека?
- ◆ Какова эталонная модель системы поддержки принятия решений водителем на основе персональных данных и методов машинного обучения.

Анализ некоторых существующих работ

С начала 2000-х годов автомобильная промышленность посвятила много времени и ресурсов на сотрудничество с исследователями из области компьютерного зрения и машинного обучения для создания соответствующих систем обнаружения сна.

Авторы статьи [1] разработали систему мониторинга сонливости водителя в режиме реального времени. Система использует камеру смартфона для записи лица водителя на видео. Система фиксирует лицевые ориентиры на обнаруженном лице, а затем вычисляет пропорции глаз, коэффициент раскрытия рта и соотношение длины носа. В зависимости от рассчитанных значений система определяет состояние сонливости на основе выбранных значений адаптивного порога. Для анализа данных авторами использовались алгоритмы SVM и Histogram of Oriented Gradients (HOG).

Авторы статьи [2] предложили систему driver fatigue detection system (DFS), которая контролирует водителя и подает предупреждающие сигналы, если водитель засыпает за рулем. Система встроена на смартфоне. Камера используется для определения следующих состояний: определение зевоты, кивание головой, движение век. Для анализа данных авторы использовали каскады классификации Haar-like, алгоритмы сопоставления с образцом.

Авторы статьи [3] предложили систему, определяющую сонливость водителя в режиме реального времени. Эта система разработана как приложение для смартфона с ОС Android, использующее камеру для определения лица водителя. Точность определения сонливости этой системой оценивается в 93,37%. Система адекватно

работает при естественном освещении и независимо от того, используются ли для водителя аксессуары, такие как очки, наушники или кепка.

В статье [4] исследуется способность водителя обнаруживать сонливость с помощью распознавания выражений лица для анализа на основе одного кадра и динамического моделирования на основе Hidden Markov Model (HMM). Они попросили двух водителей управлять безразличными сценариями как в сонных, так и в несонливом состояниях и собрали обучающие и тестовые изображения для анализа и оценки их подхода. Затем, чтобы обнаружить сонливость водителя, сначала применив алгоритмы Viola-Jones и Camshift, чтобы определить лицо водителя. После обнаружения лица они извлекли черты лица с помощью Gabor Wavelets. Затем они использовали Adaboost для выбора извлеченных функций. Эти разделы абсолютно одинаковы для обеих систем классификации. Затем они выполнили однокадровый анализ и динамическое моделирование на основе HMM. В заключение, динамический подход дает большую производительность, чем анализ на основе одного кадра. Поэтому в этой статье говорится, что выражение лица лучше распознается при анализе последовательности кадров.

В статье [5] исследователи сосредоточены на нескольких физических характеристиках, чтобы повысить точность определения уровней сонливости. Кроме того, многие другие исследования были извлечены только по одному объекту. В этой статье в основном используется Local Binary Pattern (LBP) при извлечении областей глаз и рта. REDE — это название предложенного алгоритма этой статьи. REDE превзошел предыдущие четыре исследования по выявлению усталости. Поскольку не было общедоступной базы данных для определения уровня утомляемости, они создали свою собственную базу данных, полученную от добровольцев. При предварительной обработке данных извлечения областей двух глаз и рта используется Dlib. Затем значения LBP извлекаются как для глаз, так и для рта и после этого собственный вектор функций LBP.

В статье [6] рассматривается возможность обнаружения и прогнозирования состояния водителя, такого как сонливость, путем разработки искусственной нейронной сети для этих задач. Кроме того информация, вводимая в алгоритм, собирается из различных источников данных, начиная от физиологических, поведенческих и заканчивая психологическими данными о водителе, а также информацией о характеристиках транспортного средства. Различные наборы данных из разных источников были оценены для определения наилучшей степени успех в обнаружении и прогнозировании нарушений. Данные собираются от участников симулятора вождения. Обобщение и межличностная изменчивость

рассматриваются как сложная задача для оценки водителей, данные которых не были обучены заранее.

В статье [7] предлагается единая структура сбора и анализа данных, называемая DarNet. Он использует ансамбли нейронных сетей для классификации и данные на основе датчиков смартфона для обнаружения и классификации аномального поведения за рулем. Кроме того, они признают проблемы конфиденциальности, понижая дискретизацию собранных видеоданных, что по-прежнему обеспечивает разумную точность классификации 80%.

В статье [8] представлена система, которая обнаруживает сонливости водителей и показано, что существуют потенциальные социальные проблемы, связанные с применением методов Convolutional Neural Networks (CNN) в системах интеллектуального анализа данных, которые отслеживают характер вождения и рекомендуют перерывы для отдыха, когда это необходимо, путем выделения проблем в обнаружении темнокожих лиц водителей. Это особенно важная проблема в африканских условиях, где водителей больше темнокожих. К сожалению, общедоступные наборы данных часто отражаются в различных культурных контекстах и, следовательно, не охватывают все этнические группы, что может привести к ложным определениям или моделям с расовой предвзятостью. В этой работе оценивается производительность, полученная путем обучения сложной модели нейронной сети на часто используемых наборах данных обнаружения сонливости водителя и тестировании на наборах данных, специально выбранных для более широкого представления. Следовательно, необходимо иметь решения для улучшения результатов и уменьшения влияния таких факторов, как раса, пол и возраст, на точность алгоритма в системе.

Методы распознавания лиц и объектов на лице

В системах компьютерного зрения обнаружение объектов основано на методах выделения признаков из данных пикселей с использованием различных методов. Сначала необходимо обнаружить лицо водителя, а затем определить объектов на лице, чтобы создать черты, интерпретируя их в алгебраических и алгоритмических процессах. Некоторые из широко используемых методов извлечения признаков:

- ♦ *Интенсивность света:* на изображениях лица в оттенках серого область глаз обычно темнее, чем нос и щеки, из-за разницы в высоте. Эти области можно изобразить белыми и черными прямоугольниками, которые называются Haar-like features. Самый известный алгоритм, использующий Haar-like features, — это алгоритм Viola-Jones [9].

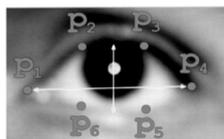
Он в основном определяет наиболее важные функции Haar с помощью Cascading AdaBoost (Adaptive Boosting) и работает с этими функциями при обнаружении лиц.

- ♦ *Собственные лица (Eigenfaces):* если изображение лица искажено, оно не будет полностью случайным, в изменениях будут некоторые закономерности. Эти изменения могут отражать лицевые члены или расстояние между ними. Таким образом, в компьютерном зрении они называются собственными лицами, и самым известным примером этого является функция 2D Габора, которая используется для улучшения топологической структуры человеческого лица, то есть носа в виде вершин, глаза в виде впадины. Когда изображение лица свернуто с помощью подходящего фильтра Габора, который в основном представляет собой синусоидальную плоскую волну, характеристики лица могут стать более подчеркнутыми, и это может быть использовано при обнаружении лиц [10].
- ♦ *Дескрипторы функций:* в компьютерном зрении существуют некоторые алгоритмы, такие как SIFT (Scalar Invariant Feature Transform) [11], SURF (Speeded-Up Robust Features) [12], для описания локальных особенностей изображений, инвариантных к масштабированию, ориентации, искажению и освещению. Поскольку это сложные и дорогостоящие алгоритмы и процедуры в вычислительном отношении, их сложно использовать в сценариях реального времени. Существует также еще один характерный дескриптор HOG (Histogram of Oriented Gradient) [13], использующий распределение направлений градиентов в качестве функций. Это один из самых быстрых и надежных подходов в области компьютерного зрения, поэтому известная библиотека Dlib использует его реализацию для обнаружения лиц с помощью линейного классификатора.

Методы классификации сонливости

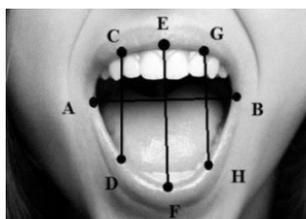
После выделения функций из необработанных данных и построения наборов обучающих данных, есть варианты классификаторов для прогнозирования уровня сонливости. Некоторые из них перечислены ниже:

- ♦ *Базовая установка пороговых значений:* Используются пороги для обнаружения сна, такие как анализ закрытия глаз или анализ зевоты с открыванием рта.
- ♦ *Анализ закрытия глаз.* Для определения состояния глаз можно использовать показатели EAR (Eye Aspect Ratio) — представляет собой отношение длины глаз к ширине глаз. Длина глаз рассчитывается путем усреднения двух четких



$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Рис. 1. Координаты глаз и формула Eye Aspect Ratio (EAR)



$$MAR = \frac{|EF|}{|AB|}$$

Рис. 2. Координаты глаз и формула Mouth Aspect Ratio (MAR)



Рис. 3. Индекс EAR ошибочно легко определяет состояние глаз азиатов.

вертикальных линий поперек глаз, как показано на рисунке 1.

- ◆ Анализ зевоты: зевота — один из первых показателей, который приходит на ум при обнаружении сонливости. В вычислительном отношении похожий на EAR, MAR измеряет отношение длины рта к ширине рта (рис. 2). Когда зевает, в результате чего его MAR в этом состоянии становится выше, чем обычно. $|AB|$ представляет евклидово расстояние между двумя позициями ориентиров.
- ◆ Обычные инструменты машинного обучения: хотя можно использовать более простые подходы, такие как Logistic Regression, Decision Tree, k-NN (K-Nearest Neighbors), NB (Naïve-Bayes) и получить удовлетворительные результаты [14], большинство ученых, вероятно, предпочитают более сложные инструменты, например SVM (Support Vector Machine), Random Forest, HMM (Hidden Markov Model), AdaBoost (Adaptive Boosting). Среди них самые популярные - SVM и HMM.
- ◆ Инструменты глубокого обучения: это математические модели, которые состоят из связанных искусственных нейронов, вдохновленных био-

логической нейронной сетью с коннекционистским подходом. Они используются для моделирования сложных отношений и шаблонов между входными и выходными данными с использованием обучаемых весов связей. Искусственная нейронная сеть (Artificial neural network - ANN) с возрастающей сложностью и несколькими промежуточными уровнями называются DNNs (Deep Neural Networks). В области обнаружения сонливости во многих сценариях ученые предпочитают подтип DNN, которые называются CNN (Convolutional Neural Network) в качестве классификатора, почти так же часто, как SVM и HMM. В дополнение к обычным слоям DNN, CNN также состоят из некоторых сверточных слоев, которые представляют собой свернутые входные изображения с фильтрами, чтобы присвоить важность некоторым аспектам/объектам изображения, чтобы снизить вычислительные затраты [15].

Предлагаемая методология

В этом разделе будут рассматриваться методы машинного обучения для улучшения результата обнаружения состояния глаз: закрываются /открываются.

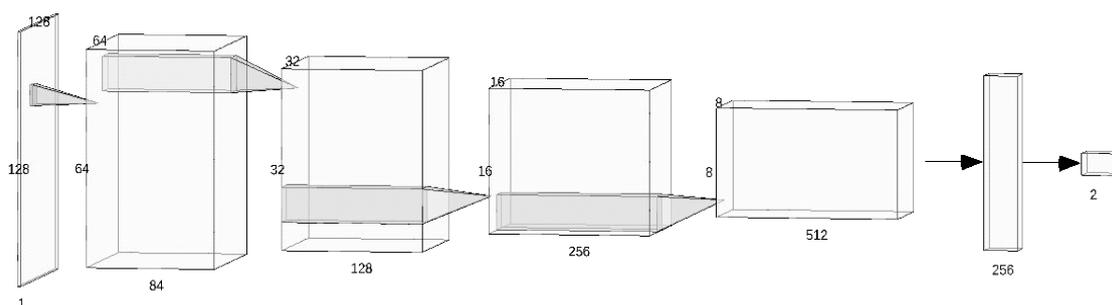


Рис. 4. Предлагаемая глубокая модель CNN

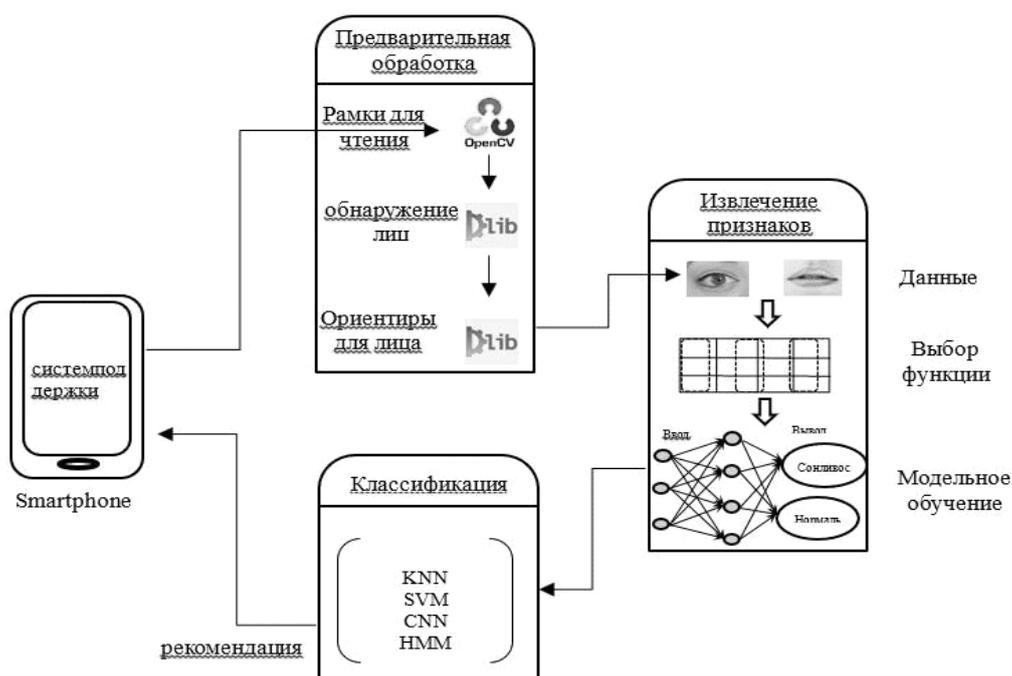


Рис. 5. Персонализация системы поддержки принятия решений

Состояние сонливости является состоянием глаз, рот закрыть или рот открыт большее. В соответствии с индексом EAR — Если значение EAR меньше определенного порога, обычно 0,3, состояние глаза определяется как закрытое.

Однако с лицами с маленькими и узкими глазами, такими как у Азиатов, при применении формулы EAR часто бывает сложно рассчитать состояние глаза в этом случае и обычно определяется неправильно — глаза всегда закрыты (рис. 3).

Для более точного определения состояния глаз у азиатов автор рекомендует использовать сеть CNN для изучения состояния закрытости и открытости глаз следующим подходом:

1. Алгоритм обнаружения лица Виолы-Джонса используется для обнаружения лица на изображении

и передает его в качестве входных данных в алгоритм обнаружения глаз Виолы-Джонса.

2. После обнаружения лица алгоритм обнаружения глаз Виолы-Джонса используется для извлечения области глаза из изображений лиц и передает ее в качестве входных данных в CNN.
3. CNN с четырьмя сверточными слоями используется для извлечения глубоких особенностей, и эти особенности передаются на полностью связанный слой.
4. Слой Softmax в CNN классифицирует изображения на сонные и несонные.

Дизайн слоев предложенной глубокой модели CNN:

В предлагаемой нами работе модель Deep CNN разработана с 4 сверточными слоями и одним полностью связанным слоем. Извлеченные ключевые изображения размером 128 X 128 передаются на вход. Выход класси-

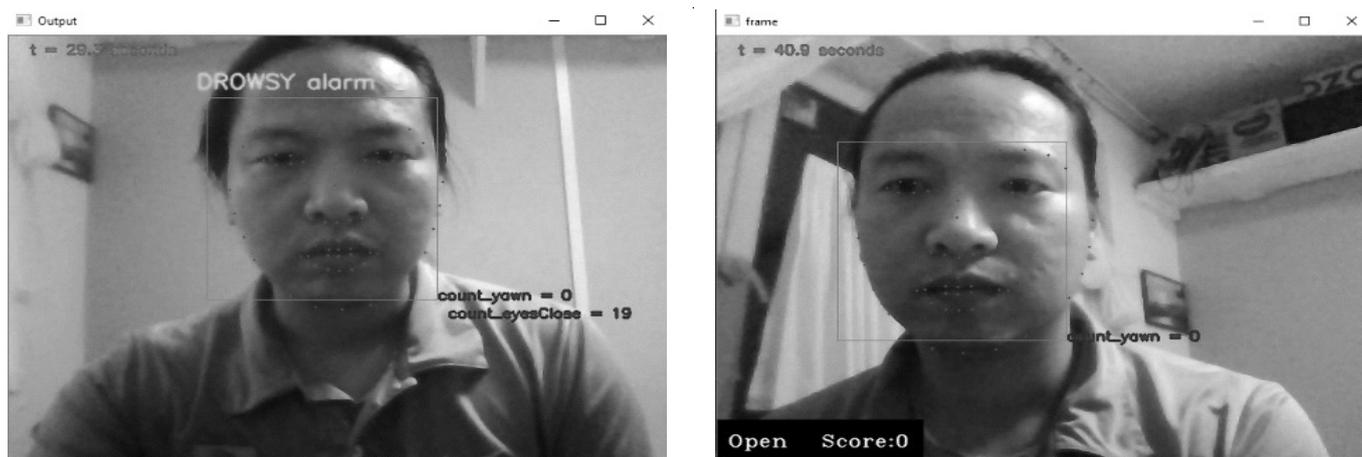


Рис. 6. Результаты до и после персонализации пользовательских данных

фикатора — это два состояния, поэтому выходной слой имеет только два выхода. В предлагаемой нами структуре CNN256 выходов полностью связанного слоя — это глубокие особенности, извлеченные из входных изображений глаза. Последние 2 выхода могут быть линейными комбинациями глубинных элементов.

На основе информации из обзора литературы, касающейся распространенных методов машинного обучения и соответствующих данных, собранных со смартфонов, была предложена эталонная модель, которая показывает, как современные подходы машинного обучения могут использоваться для улучшения качества определения сонливости у азиатов в системы поддержки принятия решений водителем (рис. 4).

Чтобы иметь возможность персонализировать пользовательские данные, при запуске системы пользователям предлагалось смотреть прямо в камеру и выполнять следующие требования: прямое зрение, закрытые глаза, зевота. Система будет изучать характеристики глаз каждого пользователя и со временем повысит точность каждого пользователя (рис. 5). Кроме того, отслеживая историю активности, включая алгоритмы нечеткой логики, система будет определять характеристики и более точно прогнозировать преждевременную утомляемость на основе таких показателей, как скорость закрытия глаз, активные привычки, активное время и т.д. чтобы выявить потенциальные особенности и предсказать будущее.

Заключение

В этой статье наше внимание сосредоточено на подходе к персонализации пользовательских данных в переносных системах помощи водителю. Основной целью персонализации является повышение точности и удобства использования и, следовательно, приемлемости системы для водителей. Это особенно важно в приложениях, связанных с безопасностью, где оповещения и их время должны быть адаптированы к навыкам и потребностям водителя, чтобы предотвратить неиспользование системы. Кроме того, персонализация может способствовать повышению безопасности систем помощи, например, предлагая возможность увеличить время предупреждения на основе наблюдения за отдельным водителем. Наконец, это средство для улучшения систем комфорта, которые могут быть адаптированы к предпочтительным стилям вождения.

Персонализация пользовательских данных становится все более важным требованием в большинстве современных систем. Изучая особенности самого пользователя, персонализация повышает точность определения состояний закрытости и открытости глаз азиатских лиц. В будущем будут использоваться методы машинного обучения и персонализации данных для разработки алгоритмов определения эмоциональных состояний или употребления психоактивных веществ во время вождения — Это тоже проблемы для водителей и очень актуальные задачи для исследования.

ЛИТЕРАТУРА

1. Kumar A., Patra R. Driver drowsiness monitoring system using visual behaviour and machine learning //2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). — IEEE, 2018. — С. 339–344.
2. Abulkhair M. et al. Mobile platform detect and alerts system for driver fatigue //Procedia Computer Science. — 2015. — Т. 62. — С. 555–564.

3. Galarza E. E. et al. Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone //International Conference on Information Theoretic Security. — Springer, Cham, 2018. — С. 563–572.
4. Tadesse E., Sheng W., Liu M. Driver drowsiness detection through HMM based dynamic modeling //2014 IEEE International conference on robotics and automation (ICRA). — IEEE, 2014. — С. 4003–4008.
5. Li K. et al. Accurate fatigue detection based on multiple facial morphological features //Journal of Sensors. — 2019. — Т. 2019.
6. de Naurois C. J. et al. Detection and prediction of driver drowsiness using artificial neural network models //Accident Analysis & Prevention. — 2019. — Т. 126. — С. 95–104.
7. Streiffer C. et al. Darnet: a deep learning solution for distracted driving detection //Proceedings of the 18th acm/ifip/usenix middleware conference: Industrial track. — 2017. — С. 22–28.
8. Ngxande M., Tapamo J. R., Burke M. Detecting inter-sectional accuracy differences in driver drowsiness detection algorithms //2020 International SAUPEC/RobMech/ PRASA Conference. — IEEE, 2020. — С. 1–6.
9. Viola P. et al. Robust real-time object detection //International journal of computer vision. — 2001. — Т. 4. — №. 34–47. — С. 4.
10. Tadesse E., Sheng W., Liu M. Driver drowsiness detection through HMM based dynamic modeling //2014 IEEE International conference on robotics and automation (ICRA). — IEEE, 2014. — С. 4003–4008.
11. Naz S., Ziauddin S., Shahid A. R. Driver fatigue detection using mean intensity, SVM, and SIFT //IJMAI. — 2019. — Т. 5. — №. 4. — С. 86–93.
12. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features //European conference on computer vision. — Springer, Berlin, Heidelberg, 2006. — С. 404–417.
13. Dalal N., Triggs B. Histograms of oriented gradients for human detection //2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). — IEEE, 2005. — Т. 1. — С. 886–893.
14. Nakamura T., Maejima A., Morishima S. Detection of driver's drowsy facial expression //2013 2nd IAPR Asian Conference on Pattern Recognition. — IEEE, 2013. — С. 749–753.
15. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Communications of the ACM. — 2017. — Т. 60. — №. 6. — С. 84–90.

© Фам Туан Ань (anhpt@itmo.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»

