

А. М. Ахметвалеев, И. В. Ахметвалеева, Р. Ф. Закиев

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВОЙ МОДЕЛИ YOLOV5 ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ РАДУЖКИ И ЗРАЧКА ГЛАЗА

Ключевые слова: функциональное состояние человека, состояние усталости водителя, пуриллометрия, нейронные сети, сегментация радужки и зрачка, сглаживание временного ряда, компьютерное зрение, нейросетевая модель YoLoV5.

В данной статье представлено применение нового подхода к сегментации изображений радужной оболочки и зрачка глаза человека с использованием предобученной нейросетевой модели YoLoV5. Не смотря на применение современных нейросетевых технологий и методов компьютерного зрения для обработки изображений и сегментаций глаза, не все из них гарантируют получение высокоточного результата искомым окружностей. Приведены примеры некорректной сегментации радужной оболочки глаза и зрачка, полученные с применением фремоворка Mediarpipe и алгоритма FloodFill. Для устранения артефактов, шумов и выбросов в данных ранее использовались сглаживающие фильтры. Вместе с тем, применение сглаживающих фильтров способствует искажению обрабатываемых ими данных, в особенности при работе в системах реального времени, обладающих свойством длительности наблюдения. Поскольку диагностика функционального состояния человека методом пуриллометрии чувствительна к качеству исходных данных, применение сглаживающих фильтров в ряде диагностических случаев не является допустимым. Для повышения точности сегментации изображений радужки и глаза, отказа от искажения данных сглаживающими фильтрами, авторским коллективом разработан модуль компьютерного зрения на базе механизма сегментации окружностей глаза с помощью предобученной нейросетевой модели YoLoV5. Приведены примеры сегментации окружностей глаза нейронной сетью YoLoV5. Проведена оценка точности сегментации изображений радужки и зрачка глаза для чего сравнивалось количество выбросов и пропусков значений временных рядов, полученных с использованием рассматриваемых в статье подходов. Предложенный в статье подход применения нейросетевой модели YoLoV5 реализован в решении задачи мониторинга состояния усталости водителя и применен в прототипе специализированного устройства, разработанного авторским коллективом.

А. М. Akhmetvaleev, I. V. Akhmetvaleeva, R. F. Zakiev

APPLICATION OF THE YOLOV5 NEURAL NETWORK MODEL TO INCREASE THE ACCURACY OF IRIS AND PUPIL IMAGE SEGMENTATION

Keywords: human functional state, driver fatigue state, pupillometry, computer vision, neural networks, iris and pupil segmentation, time series smoothing, computer vision, neural network model YoLoV5.

This article presents the application of a new approach to segmentation of images of the iris and pupil of the human eye using the pretrained neural network model YoLoV5. Despite the use of modern neural network technologies and computer vision methods for image processing and eye segmentation, not all of them guarantee obtaining a highly accurate result of the desired circles. Examples of incorrect segmentation of the iris and pupil obtained using the Mediarpipe framework and the FloodFill algorithm are given. Smoothing filters were previously used to eliminate artifacts, noise and outliers in the data. At the same time, the use of smoothing filters contributes to the distortion of the data they process, especially when working in real-time systems with the property of observation duration. Since the diagnosis of the functional state of a person by the pupillometry method is sensitive to the quality of the original data, the use of smoothing filters in a number of diagnostic cases is not acceptable. To improve the accuracy of iris and eye image segmentation and avoid data distortion by smoothing filters, the authors developed a computer vision module based on the eye circumference segmentation mechanism using the pre-trained YoLoV5 neural network model. Examples of eye circumference segmentation by the YoLoV5 neural network are given. The accuracy of iris and pupil image segmentation was assessed by comparing the number of outliers and omissions in time series values obtained using the approaches discussed in the article. The approach to using the YoLoV5 neural network model proposed in the article is implemented in solving the problem of driver fatigue monitoring and is used in a prototype of a specialized device developed by the authors.

Введение

Пуриллометрический метод медицинской диагностики [1,2] базируется на простой физиологической особенности нервной системы человека: размер зрачков здорового человека реагирует на изменение яркости света. У здорового человека, в норме, яркое освещение приводит к сужению зрачка, а в темноте зрачок глаза человека расширяется. Различные отклонения в функционировании нервной системы приводят к отклонению нормальной физиологической реакции зрачка на изменение освещенности.

Авторский коллектив в работе [3] продемонстрировал возможность сочетания современных методов обработки изображения, компьютерного зрения, ис-

кусственного интеллекта и пуриллометрического подхода к оценке функционального состояния человека в решении задачи мониторинга состояния водителей транспортных средств. Ядро идеи заключается во влиянии переутомления человека за рулем на зрачковую реакцию, что позволяет в автоматическом режиме находить отклонения размера зрачка от физиологической нормы при измеренном уровне яркости фонового освещения.

Установлена математическая связь между размером зрачков и уровнем освещения, которая определяет норму размера зрачков при заданной яркости фона [1, 4]:

$$D_p = 5-3th(0,4lg L),$$

где D_p – диаметр зрачка глаза (в мм), L – яркость фона (в кд/м²), th – гиперболический тангенс. Ее запись в формате MS Excel может быть представлена следующим образом: $D_p=5-3*TANH(0,4*LOG10(A2))$, где $A2$ – пример ссылки на номер ячейки в таблице, отражающей яркость фона.

Таблица 1 – Расчет диаметров зрачка
Table 1 – Calculation of pupil diameters

L	D_p
10	3,86
20	3,56
50	3,22
200	2,82

Для проверки правильности вычислений составлена тестовая таблица расчета диаметров зрачка, в сокращенном виде представленная в таблице 1. Как показали проверочные вычисления, при увеличении значения L (яркость фона), имеется тенденция к уменьшению значения D_p (диаметр зрачка), что соответствует физиологической реакции глаза человека на изменение освещенности.

В графическом виде интерпретация полученных значений выглядит следующим образом (рисунок 1).

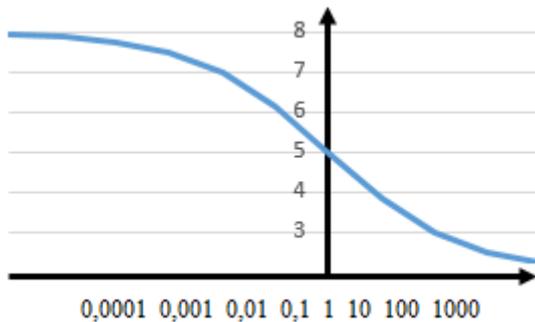


Рис. 1 – Графическая интерпретация зависимости диаметра зрачка от яркости фона в MS Excel

Fig. 1 – Graphical interpretation of the dependence of pupil diameter on background brightness in MS Excel

Представленная на рисунке 1 графическая интерпретация полностью соответствует существующим исследованиям в области пупилломоторной реакции зрачка на освещенность. Значение размера зрачка имеет тенденцию к уменьшению (физиологически – к сужению) при увеличении освещенности.

Таким образом, любое значительного отклонение размера зрачка от его известной квазиэталонной возрастной нормы при измеренном уровне фонового освещения позволяет интеллектуальной системе строить выводы о наличии отклонений функционального состояния человека за рулем [3]. К числу таких состояний также относится состояние переутомления водителя, что увеличивает риски возникновения дорожно-транспортного происшествия.

Возможные недостатки в работе методов компьютерного зрения и путь их устранения

Метод пупиллометрической диагностики оценки размера зрачка при изменении фонового освещения

чувствителен к качеству анализируемых данных. В различных исследованиях показано, что обработку и анализ видеоизображений таких сложных объектов как зрачок и радужка глаза человека актуально производить с применением различных методов компьютерного зрения и технологий искусственного интеллекта. Для сегментирования окружностей зрачка и радужки в ранних работах авторского коллектива допускалось применение предобученных нейронных сетей из состава фреймворка Mediapipe и простого алгоритма вычисления площади объектов – FloodFill [5,6]. Вместе с тем, высокая точность сегментации искомым окружностей данными методами не гарантировалась. На рисунке 2 показан пример ошибки сегментации радужки на изображении лица с использованием фреймворка Mediapipe.

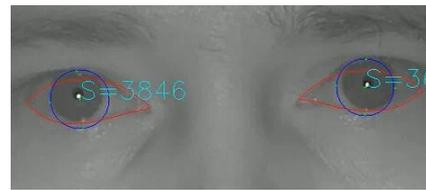


Рис. 2 – Пример ошибок в сегментации радужки
Fig. 2 – Example of errors in iris segmentation

На представленном выше рисунке показан характерный для фреймворка Mediapipe пример ошибки в определении границы радужной оболочки глаза, что ведет к неверному вычислению ее площади в пикселях.

В тоже время, при сегментации изображений имеющих низкую контрастность, а также имеющих перекрытие области глаз веками и ресницами, наличие бликов и отблесков приводит к неадекватности результата работы алгоритма вычисления площади объектов FloodFill, как показано на рисунке 3.



Рис. 3 – Пример ошибки сегментации зрачка методом FloodFill

Fig. 3 – Example of pupil segmentation error by FloodFill method

На рисунке 3 показан пример ошибки сегментации зрачка методом FloodFill, демонстрирующий случай выхода за границы окружности зрачка из-за наличия блика в искомом зрачке.

Изначальные ошибки в сегментации изображений неминуемо приводят к аномальным значениям получаемых временных рядов, как показано на рисунке 4.

На рисунке 4 показан пример наличия выбросов и пропусков данных в полученном для пупиллометрической диагностики временном ряде. Подобные аномалии приводят к искажению вычислений и

снижают точность оценки функциональных состояний в проводимых исследованиях.

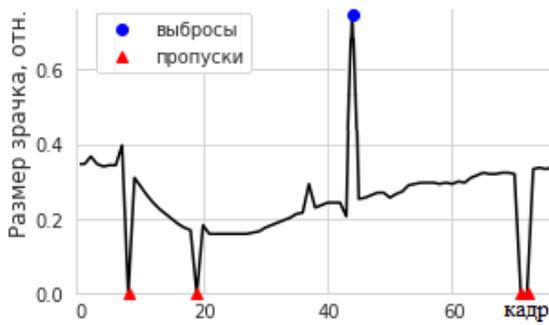


Рис. 4 – Пример выбросов и пропусков анализируемого временного ряда

Fig. 4 – Example of outliers and omissions of the analysed time series

Для решений данной проблемы авторским коллективом в работе [5] предлагалось обрабатывать временные ряды методом линейной интерполяции и сглаживающим фильтром Савицкого-Голея. В качестве обоснования для выбора фильтра Савицкого-Голея использовалось требование о недопущении искажений данных временного ряда и сохранении тенденции изменений значений размера зрачка во время исследования.

Оценка функционального состояния человека методом пупиллометрии носит кратковременный характер: оценка строится на базе 3 секундного измерения размера зрачка после светового импульсного воздействия, поэтому применение сглаживающих фильтров весьма актуально и позволяет повысить качество данных при наличии в них искажений. В тоже время, оценка состояния усталости водителя осуществляется в реальном времени и за длительный период наблюдения, а также характеризуется частым изменением уровня фоновой освещенности на всей протяженности пути автотранспортного средства [3]. При этом различия в уровнях освещенности городской транспортной среды и загородных трасс, а также частая смена экспозиции в ходе передвижения транспортного средства, не позволяет при проведении исследований опираться на наличие тенденций временного ряда, как в случаях классического пупиллографического теста. Таким образом, очевидно, что использование сглаживающих фильтров на временном ряду, не имеющим устойчивой тенденции в столь длительном периоде наблюдения, может вызвать искажение данных значительно большее, чем наличие артефактов.

Экспериментальная часть

Для преодоления этого недостатка и совершенствования методики разработан новый комплексный подход к обеспечению высокоточной сегментации радужки и зрачка без необходимости дополнительного сглаживания временного ряда. Результатом проведенных исследований и разработок стал программный модуль компьютерного зрения, в основе которого заложен механизм сегментации окружностей глаза с помощью предобученной нейронной

сети YoLoV5. Примеры работы программы показаны на рисунке 5.

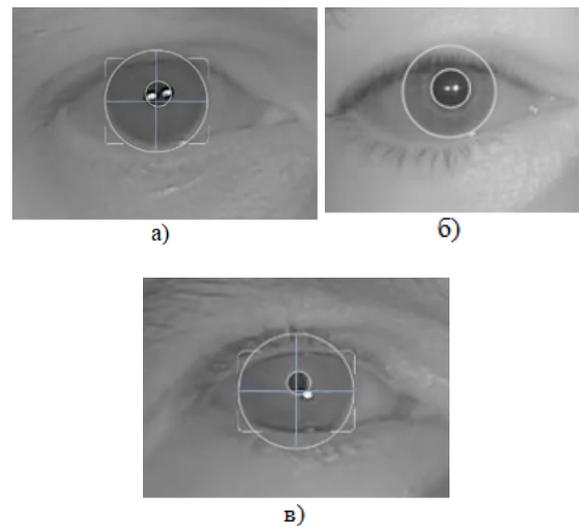


Рис. 5 – Примеры сегментации окружностей глаза нейронной сетью YoLoV5

Fig. 5 – Examples of segmentation of eye circles by the YoLoV5 neural network

Разработанный модуль компьютерного зрения показал высокую устойчивость результата сегментирования при различных искажениях изображения, например, при наличии оптических искажений, вызванных большим размером бликов на зрачке (рис. 5а), перекрытии искомого объектов другими, а именно веками и ресницами (рис. 5б), наличии в поле сегментирования наложений на окружности (рис. 5в) и в других случаях.

Оценка эффективности использования нейросетевой модели YoLoV5 проводилась в сравнении с технологией получения и анализа пупиллограмм [6], основанной на комбинации применения технологии распознавания лиц и глаз на лице человека сверточной нейронной сетью Mediapipe компании Google [7], цифровой обработки изображений и алгоритма сегментации FloodFill.

Используемая для оценки выборка данных (датасет) содержит 384 видеороликов лиц испытуемых с записью зрачковой реакции глаз человека на световое импульсное воздействие. Каждый видеоролик имеет длительность 3 секунды, по 25 кадров на каждую секунду, то есть 75 кадров изображений на каждый видеоряд. Таким образом, всего проанализировано 28 800 видеокадров. Видеокадры покадрово обрабатывались оцениваемыми методами, для чего строились графические интерпретации распознанного временного ряда зрачковой реакции в исследуемый 3 секундный период (рис.6) и текстовые файлы (рис.7), содержащие вычисленные относительные значения размеров зрачков глаза человека.

В представленном примере временного ряда можно увидеть как пропуски данных, в виде спадов графика в нулевые значения, так и выбросы визуально наблюдаемых как всплеском временного ряда с значениям близким к единичным.

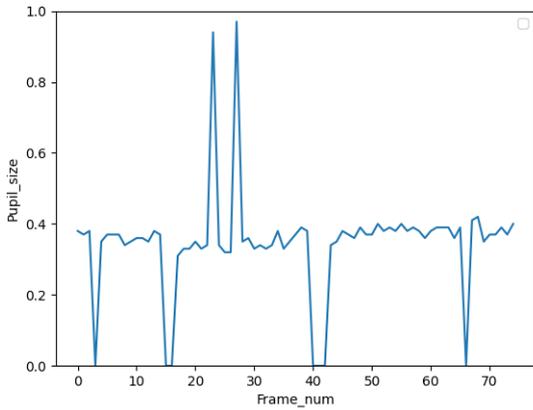


Рис. 6 – Пример распознанного временного ряда зрачковой реакции

Fig. 6 – Example of recognised pupillary response time series



Рис. 7 – Пример текстового файла

Fig. 7 – Example of a text file

На примере текстового файла показан последовательный набор табулированных данных, содержащий полученные относительные размеры зрачка глаза человека. Каждый файл содержит 75 таких значений – по количеству кадров в видеофайле. Пропуски данных в файле представлены нулевыми значениями, выбросы – значениями выше определенного порога, в данном случае, более 0.7.

Относительный размер зрачка вычислялся по следующей формуле [6]:

$$S = S_{pupil} / S_{iris},$$

где S_{pupil} – площадь зрачка в пикселях, S_{iris} – площадь радужки в пикселях.

В качестве критерия оценки точности сегментации изображений радужки и зрачка глаза использовалось количество выбросов и пропусков значений, полученных с использованием указанных выше подходов на всем наборе данных, состоящего из 28 800 видеок кадров. Результаты оценки представлены в таблице 2.

Как видно из представленных результатов, точность сегментации окружностей радужки и зрачка глаза человека на 5% больше, по сравнению с применением методов в ранних работах авторского коллектива.

Дополнительными исследованиями установлено, что большая часть аномальных значений связана с пропусками значений, вызванных невозможностью

сегментации окружностей при закрытых глазах исследуемых лиц. В то же время, негативные факторы, такие как блики, ресницы, нераскрытые полностью веки, перекрытие объектов поиска, нечетко сфокусированное изображение не оказывают влияния на сегментацию искомым окружностей при использовании нейросетевой модели YoLoV5.

Таблица 2 – Результаты оценки точности сегментации искомым окружностей

Table 2 – Results of the evaluation of the segmentation accuracy of the sought circles

Подход	Кол-во правильно распознанных видеок кадров	Количество выбросов и пропусков значений	Доля выбросов и пропусков в процентах
Сверточная нейронная сеть Medi-aPipe, цифровая обработка изображений, алгоритм FloodFill	26 423	2377	8%
Нейросетевая модель YoLoV5	27 823	976	3%

Реализация и перспективы использования предложенного подхода

Практическая реализация применения нейросетевой модели YoLoV5 для сегментации радужки и зрачка глаза человека осуществлялась в рамках государственного проекта «Платформа университетского технологического предпринимательства» и стартап-проекта по договору №94ГССС15-L/78980 от 19.08.2022 г. с Фондом содействия инноваций. Внешний вид полученного прототипа представлена на рисунке 8.



Рис. 8 – Внешний вид разработанного прототипа устройства

Fig. 8 – Appearance of the developed prototype device

Устройство устанавливается рядом с приборной панелью напротив водителя таким образом, чтобы объектив был направлен на глаза человека, а его обзору ничего не мешало. На задней стенке устройства расположен датчик освещенности, который производит замер уровня фонового освещения, падающего на глаза человека за рулем.

К настоящему времени разработанное устройство представляет собой совокупность использования микрокомпьютерных технологий, высокоуровневого программирования, методов компьютерного зрения и искусственного интеллекта в виде MVP-прототипа, доведенного до уровня технологической готовности TRL-3, что позволяет продемонстрировать его ключевые характеристики широкой публике и проводить апробацию.

Основная область применения полученного продукта – обеспечение транспортной безопасности. Ключевыми потребителями данной области являются государственные и коммерческие предприятия, а также индивидуальные предприниматели, владеющие собственным автопарком и заинтересованные в повышении безопасности при управлении транспортным средством, недопущения дорожно-транспортных происшествий, связанных с нахождением человека за рулем в состоянии, опасном для вождения, а также связанного с ними финансово-экономического и социального ущерба.

Кроме того, потребителями результатов исследований и разработки могут стать действующие производители систем контроля усталости водителей и автоконцерны, производящие современные автомобили, оснащенные различными интеллектуальными системами активной и пассивной безопасности. В данном случае существующие системы могут быть усовершенствованы путем интеграции разработанного математического, алгоритмического и программного обеспечения.

Заключение

Эффективность анализа изображений с помощью методов компьютерного зрения напрямую определяет достоверность полученной оценки функционального состояния водителя.

Разработанный метод компьютерного зрения, основанный на нейросетевой модели YoLoV5, представленный в исследовании, показал высокую эффективность для решения поставленной задачи, что подтверждается проведенными экспериментами.

Ожидается, что сочетание высокоточной оценки состояния усталости водителя с эффективными методами ее реализации в программно-аппаратном комплексе позволит не только повысить безопасность вождения, но и снизить социально-экономические последствия, связанные с авариями по причине усталости водителей.

Литература

1. D.A. Atchison, G. Smith, *Optics of the Human Eye. Second Edition*, Taylor&Francis Group, Бока-Ратон, Флорида, США, 2023. 499 с.
2. А.М. Ахметвалеев, М.А. Подольская, А.С. Катасёв, И.В. Ахметвалеева, Р.Ф. Закиев, *Российский кардиологический журнал*, **27**, S5, 14-15 (2022).
3. И.В. Ахметвалеева, Р.Ф. Закиев, А.М. Ахметвалеев, А.С. Катасёв, *Вестник Технологического университета*, **26**, 4, 58-62 (2023).
4. Т.Н. Хацевич, *Медицинские оптические приборы. Физиологическая оптика*, СГГА, Новосибирск, 2010, 135 с.
5. А.М. Ахметвалеев, *Математические методы в технологиях и технике*, **9**, 75-78 (2022).
6. Akhmetvaleev A.M., Danilchuk M.Yu., Katasev A.S., Akhmetvaleeva I.V., *Studies in Systems, Decision and Control*, **416**, 75–86 (2022).
7. Lugaresi C., Tang, J., Nash, H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.L., Yong M.G., Lee J., et al., *Computer Science. Distributed, Parallel, and Cluster Computing*, 2019, 1906.08172.

References

1. D.A. Atchison, G. Smith, *Optics of the Human Eye. Second Edition*, Taylor&Francis Group, Boca Raton, Florida, USA, 2023. 499 p.
2. A.M. Akhmetvaleev, M.A. Podolskaya, A.S. Katasev, I.V. Akhmetvaleeva, R.F. Zakiev, *Russian Journal of Cardiology*, **27**, S5, 14-15 (2022).
3. I.V. Akhmetvaleeva, R.F. Zakiev, A.M. Akhmetvaleev, A.S. Katasev, *Herald of Technological University*, **26**, 4, 58-62 (2023).
4. T.N. Khatsevich, *Medical optical devices. Physiological optics*. SSGA, Novosibirsk, 2010, 135 p.
5. A.M. Akhmetvaleev, *Mathematical Methods in Technology and Engineering*, **9**, 75-78 (2022).
6. Akhmetvaleev A.M., Danilchuk M.Yu., Katasev A.S., Akhmetvaleeva I.V., *Studies in Systems, Decision and Control*, **416**, 75–86 (2022).
7. Lugaresi C., Tang, J., Nash, H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.L., Yong M.G., Lee J., et al., *Computer Science. Distributed, Parallel, and Cluster Computing*, 2019, 1906.08172.

© **А. М. Ахметвалеев** – к-т. техн. наук, доцент кафедры Систем Информационной Безопасности (СИБ), Казанский национальный исследовательский технический университет им. А.Н. Туполева (КНИТУ им. А.Н. Туполева), Казань, Россия, amir.akhmetvaleev@gmail.com; **И. В. Ахметвалеева** – аспирант кафедры СИБ, КНИТУ им. А.Н. Туполева, Irina.Akhmetvaleeva@gmail.com; **Р. Ф. Закиев** – ассистент кафедры СИБ, КНИТУ им. А.Н. Туполева, zrf2010@gmail.com.

© **A. M Akhmetvaleev** – PhD (Technical Sci.), Associate Professor, The Information Security Systems (ISS) Department, Kazan National Research Technical University named after A.N. Tupolev (KNRTU named after A.N. Tupolev), Kazan, Russia, amir.akhmetvaleev@gmail.com; **I. V. Akhmetvaleeva** – PhD-Student of the ISS department, KNRTU named after A.N. Tupolev, Irina.Akhmetvaleeva@gmail.com; **R. F. Zakiev** – Assistant of the ISS department, KNRTU named after A.N. Tupolev, zrf2010@gmail.com.