

Мониторинг действий водителя на основе применения сверточных нейросетевых алгоритмов

П.В. Буранкина
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
burankina.polya@gmail.com

В.Е. Дементьев
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
dve@ulntc.ru

А.А. Сергеев
Ульяновский Государственный
Технический Университет
Ульяновск, Россия
a.sergeev@r73.center-inform.ru

Аннотация – В работе обосновывается актуальность задачи мониторинга действий водителя, формулируются требования к осуществлению такого мониторинга и предлагаются два варианта его реализации на основе применения высокопроизводительной платформы со встроенной дискретной видекартой и обычного мобильного телефона. Полученные количественные характеристики эффективности позволяют сделать вывод о практической возможности решения задачи распознавания действий водителя в режиме реального времени в том числе на низкопроизводительных платформах.

Ключевые слова — распознавание образов, поведенческий анализ, искусственные нейронные сети, мобильная разработка.

1. ВВЕДЕНИЕ

В настоящее время активно развиваются различные системы видеомониторинга и видеоаналитики для приложений, связанных с контролем поведения операторов, выполняющих трудовые функции. Такие системы наиболее актуальны и востребованы в тех областях трудовой деятельности, где ошибочные или заведомо неправильные действия человека-оператора могут привести к значительному ущербу, например, грузовые и пассажирские перевозки, конвейерное производство, обследование потенциально опасных объектов. Поведение человека в этих областях может быть описано определенными параметрами и нормами, позволяющими оценить насколько хорошо или плохо производится та или иная трудовая функция. Так, отвлечение внимания водителя большегрузного автомобиля во время движения может быть оценено как потенциально опасное событие, приводящее к возможному дорожно-транспортному происшествию. Многообразие трудовой деятельности оператора не позволяет исчерпывающим образом описать все возможные ситуации, в которые может попасть оператор, и однозначно поставить этим ситуациям определенную характеристику опасности или полезности. Однако развитие современных систем искусственного интеллекта в последние годы дает основание предполагать, что самообучаемые комплексы на основе ансамблей искусственных нейронных сетей могут оказаться способны к корректному распознаванию большинства реальных производственных ситуаций.

В настоящей работе рассматриваются возможности применения аппарата машинного

обучения для распознавания поведения водителя при осуществлении его трудовых функций.

2. ОПИСАНИЕ АЛГОРИТМОВ И РЕЗУЛЬТАТЫ.

Для этого по результатам взаимодействия с транспортными компаниями были выявлены следующие интересные особенности регистрируемой сцены, требующие распознавания. Во-первых, необходимо установить факт того, что камера направлена на рабочее место водителя и водитель в текущий момент управляет транспортным средством. Во-вторых, нужно сделать вывод о текущих действиях водителя. При этом были выделены восемь возможных вариантов этих действий:

- безопасное вождение;
- разговор по мобильному телефону;
- разговор с пассажиром;
- настройка радио;
- распитие напитков;
- набор сообщений на телефоне;
- засыпание водителя;
- отвлечение водителя от дороги.

Будем считать для определенности, что каждое отдельное изображение может содержать только одно действие. Тогда воспользуемся для распознавания указанных ситуаций стандартными методами машинного обучения [1-4] на основе применения искусственных нейронных сетей семейства YOLA 4. Такой выбор обусловлен хорошим балансом между производительностью, возможностью контроля вероятности ложной тревоги, высокой точностью распознавания и наличием специальной упрощенной версии YOLA 4 Tiny, предназначенной в том числе для работы на Android платформах. При этом для обучения будем использовать базовую обучающую выборку [5], пополненную массивом оригинальных размеченных изображений, полученных от транспортных предприятий Ульяновской области. Обучение будет производиться стандартным образом в пределах 100 эпох и частичной нормализации обучающей выборки. Для оценки качества работы использовалась 1000 изображений, не входящих в обучающую выборку. Из этого массива 100 изображений не содержали сцену с рабочим местом водителя, на 100 изображениях рабочее место было, но не было самого водителя. На оставшихся 800 изображениях водитель на своем рабочем месте был занят одним из восьми действий. При этом для случая сети Yola 4 обработка изображений велась на высокопроизводительном компьютере с использованием графических процессоров Nvidia Quadro RTX 3000, а в

случае YOLA 4 Tiny использовался обычный мобильный телефон Samsung S10. Для повышения качества обработки временной ряд результатов распознавания был обработан медианным фильтром с размером окна 5.

На рис. 1-3 приведены некоторые результаты распознавания.



Рис. 1. Обнаружен класс C7 - Разговор с пассажиром

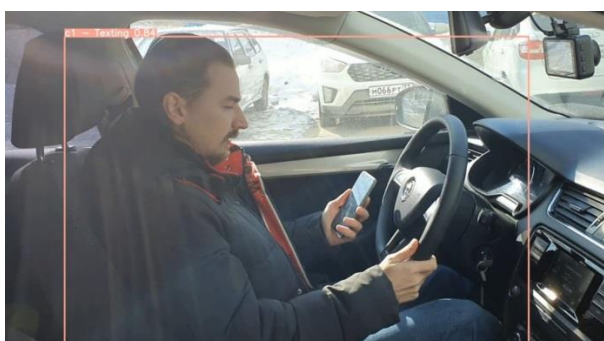


Рис. 2. Обнаружен класс C1 - Использование моб. телефона



Рис. 3. Обнаружен класс C0 - Безопасное вождение

В таблице ниже представлены количественные характеристики эффективности, соответствующие проведенным экспериментам.

Таблица I. РЕЗУЛЬТАТЫ РАСПОЗНАВАНИЯ

Распознаваемая сцена	Оценка вероятности правильного обнаружения YOLA 4	Оценка вероятности ложной тревоги YOLA 4	Оценка вероятности правильного обнаружения YOLA 4 Tiny	Оценка вероятности ложной тревоги YOLA 4 Tiny
Отсутствует рабочее место	0.97	0.03	0.95	0.04

водитель				
Отсутствует водитель	0,98	0,02	0,96	0,03
Безопасное вождение	0,91	0,06	0,89	0,06
Разговор по мобильному телефону;	0,95	0,03	0,92	0,04
Разговор с пассажиром	0,85	0,09	0,82	0,09
Настройка радио	0,92	0,03	0,9	0,05
Распитие напитков	0,94	0,03	0,92	0,04
Набор сообщений на телефоне	0,91	0,04	0,89	0,06
Засыпание водителя	0,94	0,03	0,92	0,04
Отвлечение водителя от дороги	0,88	0,04	0,86	0,05

ЗАКЛЮЧЕНИЕ

Таким образом, анализ полученных результатов показывает, что и YOLA 4 и YOLA 4 Tiny обеспечивают достаточно высокое качество распознавания сцены. При этом средняя скорость обработки YOLA 4 составило 26 кадров в секунду на графическом процессоре Nvidia Quadro RTX 3000, а YOLA 4 Tiny 3,6 кадра в секунду на борту обычного мобильного телефона Samsung S10. Указанные результаты позволяют рекомендовать рассмотренные нейросетевые процедуры для реализации реальных систем мониторинга действий водителя.

БЛАГОДАРНОСТИ

Работа поддержана Фондом содействия развитию малых предприятий в НТС в рамках договора 8ГУКодИИС12-D7/72685 от 27.12.2021 (Программа КодИИ).

ЛИТЕРАТУРА

- [1] Maturana, D. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition / D. Maturana // IEEE_RSJ International Conference on Intelligent Robots and Systems (IROS). – 2015. – P. 922-928.
- [2] Redmon, J. You Only Look Once: unified, real-time object detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – P. 779-788.
- [3] Deng, J. ImageNet: A Large-Scale Hierarchical Image Database / J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2009. – P. 248-255. DOI: 10.1109/CVPR.2009.5206848
- [4] Андриянов, Н.А. Обнаружение объектов на изображении: от критериев Байеса и Неймана-Пирсона к детекторам на базе нейронных сетей EfficientDet / Н.А. Андриянов, В.Е. Дементьев, А.Г. Ташлинский // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 139-159. DOI: 10.18287/2412-6179-CO-922.
- [5] State Farm Distracted Driver [Electronical Resource]. – Access mode: <https://www.kaggle.com/competitions/state-farm-distracted-driver-detection/code> (request date 11.10.2022).