

УДК 004.932.72'1:004.056.5

Использование нейронных сетей семейства YOLO для поиска и сокрытия чувствительной информации в видеопотоке

А.К. Писарев✉, П.В. Андреященко

Волгоградский государственный технический университет, Волгоград, Россия

В статье рассматривается использование нейронных сетей семейства YOLO для поиска и сокрытия чувствительной информации в видеопотоке. Приведено обоснование актуальности защиты персональных и конфиденциальных данных при использовании их в медицинской сфере. В качестве целевых объектов детекции определены лица людей, документы и экраны цифровых устройств. Предложен алгоритм покадровой обработки, включающий предварительную подготовку изображения, детектирование объектов, агрегацию областей интереса и их последующее сокрытие. Для обнаружения объектов применяются модели нейронных сетей семейства YOLO. Представлены результаты обучения и дообучения соответствующих моделей. Описана архитектура программной системы, а также продемонстрирован результат работы.

Ключевые слова: чувствительная информация, видеопоток, компьютерное зрение, YOLO, детектирование объектов, сокрытие информации.

Using Neural Networks from the YOLO Family to Detect and Conceal Sensitive Information in a Video Stream

A.K. Pisarev✉, P.V. Andreyaschenko

Volgograd State Technical University, Volgograd, Russia

This article examines the use of neural networks from the YOLO family to detect and mask sensitive information in a video stream. It provides a rationale for the importance of protecting personal and confidential data when used in the medical field. Human faces, documents, and digital device screens are identified as the target objects for detection. A frame-by-frame processing algorithm is proposed, which includes preliminary image preprocessing, object detection, aggregation of regions of interest, and their subsequent masking. Neural network models from the YOLO family are used for object detection. The results of training and fine-tuning the corresponding models are presented. The architecture of the software system is described, and the results of its operation are demonstrated.

Keywords: sensitive information, video stream, computer vision, YOLO, object detection, information concealment.

Актуальность

В последнее время наблюдается стремительное развитие сферы здравоохранения. Улучшается и модернизируется технологическое оснащение медицинских учреждений, совершенствуются бизнес-процессы и повышается качество оказания услуг. Основными двигателями этого прогресса являются активная цифровизация, внедрение телемедицины, а также использование современных методов машинного обучения и искусственного интеллекта (ИИ) [1]. Все это дает новые возможности для развития, но создает дополнительные риски безопасности. Так, одной из угроз является компрометация персональных данных (ПД) пациентов.

В Российской Федерации защита ПД регламентируется Федеральным законом № 152 [2]. В нем описываются основные положения о том, какие данные считаются персональными, как их правильно хранить и обрабатывать. Согласно данному закону, информация, которая относится прямо или косвенно к определенному физическому лицу, является персональной. Под обработкой ПД подразумевается любое действие или совокупность действий, совершаемых с использованием средств автоматизации. Похожие законы и регламенты есть и в других странах. В Бразилии ПД регулируются общим законом о защите данных [3], в котором медицинская информация подлежит усиленной защите, а для ее обработки необходимы явные согласия или специальные условия. В Китае закон Китайской Народной Республики (КНР) «О безопасности данных» [4] и закон КНР «О защите персональной информации» [5] классифицируют медицинские данные как конфиденциальную информацию граждан и устанавливают строгие требования к её защите при трансграничной передаче.

Медицинские организации и учреждения здравоохранения используют в своей работе различные цифровые технологии, которые в свою очередь связаны с передачей и обработкой видео. Например, к таким технологиям можно отнести системы телемедицины, дистанционное консультирование, видеофиксацию медицинских процедур, обучение персонала и мониторинг пациентов. На основе цифровых технологий формируются обучающие материалы для медицинского персонала, проводится контроль качества оказания услуг, а также они используются в автоматизации процессов, моделях машинного обучения и нейронных сетях. В случае отсутствия механизмов автоматического обнаружения и сокрытия чувствительной информации может произойти нарушение требований законодательства в области защиты персональных данных. Отсутствие механизмов также не позволяет использовать огромный массив данных для решения необходимых задач.

В данной работе мы выделили три основных класса обнаруживаемых объектов, которые содержат чувствительную медицинскую информацию:

- лица людей;
- экраны монитора, ноутбука, телевизора;
- документы.

Выбор данных классов обусловлен тем, что, вероятнее всего, на них будет присутствовать информация, которую необходимо обязательно скрыть.

Проанализировав необходимость обеспечения конфиденциальности медицинских данных, появляется необходимость создания автоматизированного ПО для защиты информации. Текущий способ ручной обработки видеоматериала является достаточно трудоемким. При увеличении количества материала, увеличивается и время на его обработку. Не стоит также исключать человеческий фактор. Ручная обработка также не подходит для потоковой работы.

В результате исследования было принято решение создать программу автоматического поиска и сокрытия чувствительной информации. ПО будет принимать на вход видеопоток, обрабатывать его и возвращать поток со скрытыми областями. Работа программы должна соответствовать времени, близкому к реальному. Данное решение позволит снизить риск утечки данных и разглашения врачебной тайны.

Алгоритм поиска и сокрытия информации

Алгоритм программы заключается в покадровой обработке видеопотока (рис. 1). На первом этапе происходит инициализация всех моделей детектирования. В данном ПО их три, как и количества классов объектов. Решение создать отдельные детекторы для каждого класса связано с требованиями эксплуатации программы. В реальных

сценариях использования может возникнуть необходимость отключения отдельных классов от сокрытия. Далее видеопоток разбивается на кадры, происходит их нормализация. Каждая модель проходится по кадру и обнаруживает области, содержащие чувствительные объекты. После этого происходит агрегация всех найденных областей и их отсеивание по заданному порогу уверенности. Далее происходит сокрытие всех областей выбранным методом и передача обработанного кадра в выходной поток.

Обучение моделей

Выделяют различные методы машинного обучения для обнаружения информации как на изображении, так и в видеопотоке. Некоторые методы требуют низких вычислительных затрат, при этом имея ограниченную устойчивость к изменению ракурса, освещения или масштаба объектов. К таким методам можно отнести каскады Хаара [6] и гистограммы направленных градиентов [7]. Это снижает возможность их использования в медицинских видео.

Существуют также двухэтапные детекторы, которые показывают высокую точность распознавания информации, но при этом имеют высокую вычислительную сложность. К таким можно отнести Faster R-CNN [8] и его различные модификации.

В связи с этим наилучшим вариантом является использование одноэтапных детекторов, к которым относятся модели YOLO [9]. Они являются одними из лидеров, так как распознают информацию за один проход нейросети по изображению. Данные модели обеспечивают высокую точность при высокой скорости работы, что позволяет обрабатывать видеопоток при минимальной задержке. Еще одним преимуществом семейства YOLO является его масштабируемость: модели можно адаптировать для ограниченных вычислительных ресурсов без потери точности детектирования. Поэтому были выбраны модели семейства YOLO.

Модель для детектирования лиц была взята из открытых источников [10] и дообучена на небольшом датасете [11]. Обучение остановлено на 81-й эпохе из-за отсутствия улучшения целевой метрики в течение 15 эпох. Значения метрик были зафиксированы на уровне mAP50 – 0,9835, Precision – 0,9715, Recall – 0,9481 и F1-score – 0,9597 (рис. 2).

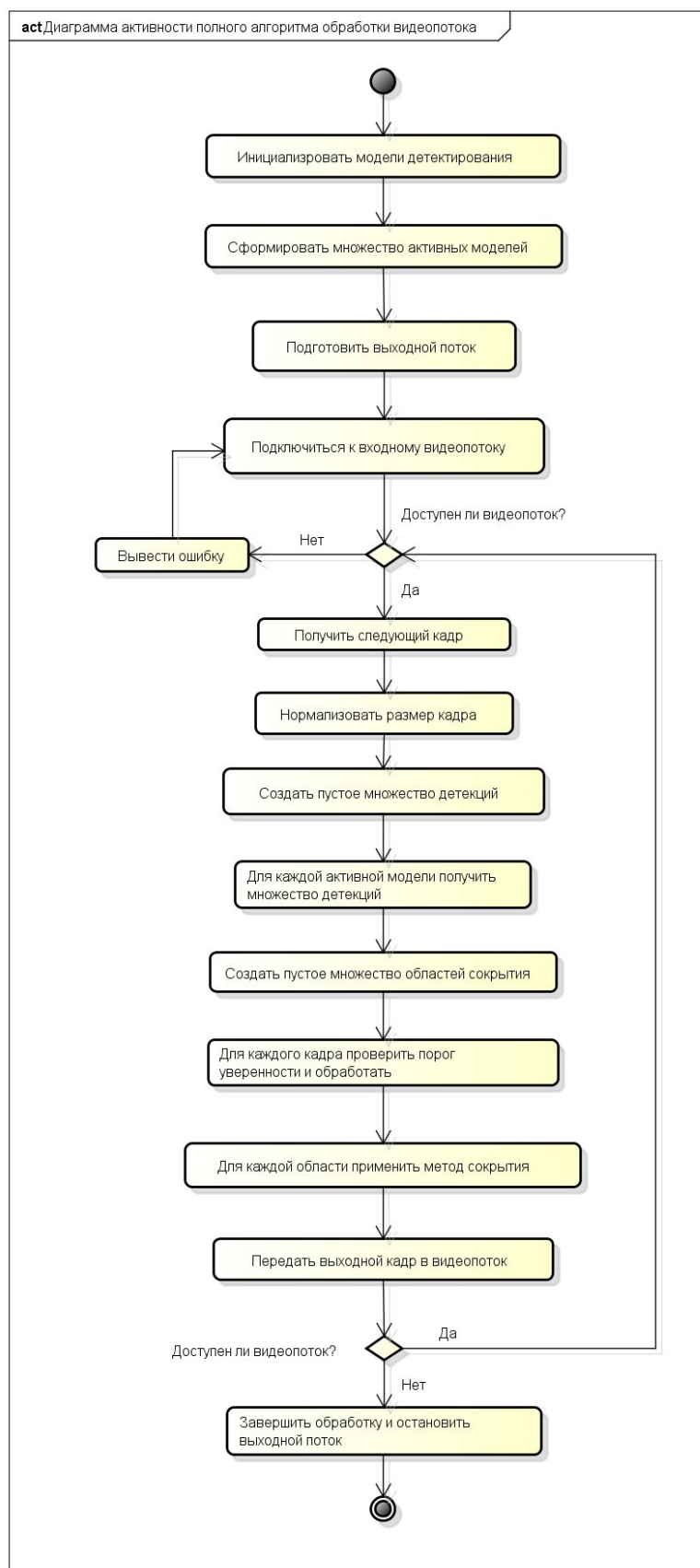


Рисунок 1. Диаграмма алгоритма автоматизированного поиска и сокрытия медицинской чувствительной информации в видеопотоке

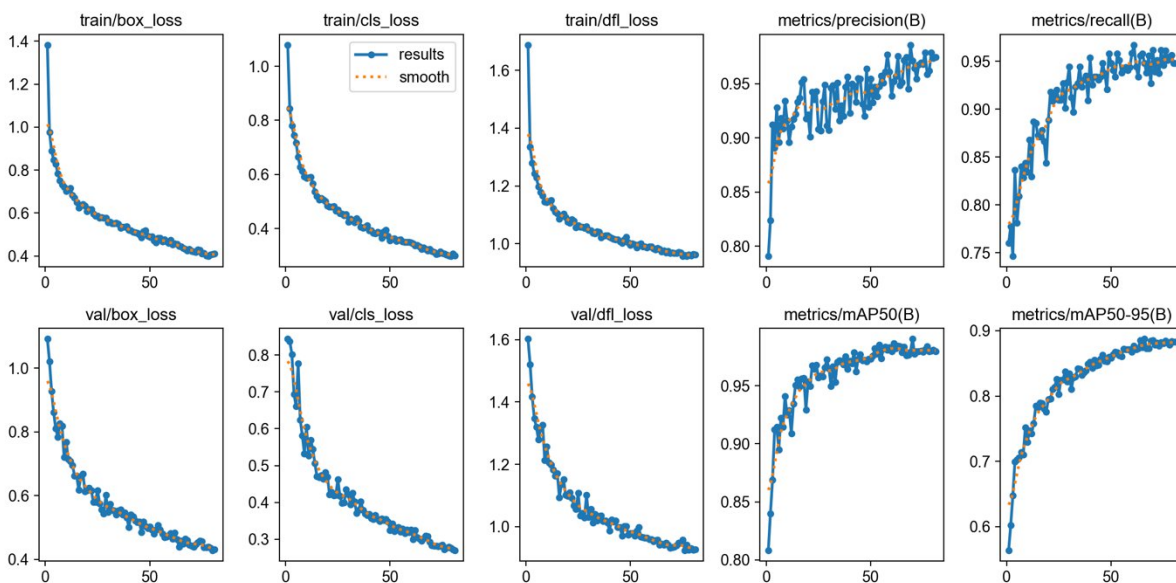


Рисунок 2. Результат дообучения модели детектирования лиц

Для детектирования экранов была использована модель YOLO26n с готовыми весами, обученными на официальном датасете COCO, включающая классы мониторов, телевизоров и ноутбуков. Результаты проведенных экспериментов приведены в официальной документации Ultralytics: значение mAP50 составляет 40,9% [12]. Данная модель обеспечивает базовое детектирование экранов на изображении.

Модель для детектирования документов была обучена на пользовательском датасете [13] для обнаружения только 1 класса. Для обучения использовалась модель YOLO26n, так как она обеспечивает достаточную точность при низких вычислительных требованиях. Обучение остановлено на 30-й эпохе, поскольку дальнейшее увеличение числа эпох не приводило к существенному росту метрик mAP50 и mAP50-95 (рис. 3). Значения метрик были зафиксированы на уровне mAP50 – 0,9937, Precision – 0,9943, Recall – 0,9914 и F1-score – 0,9929.

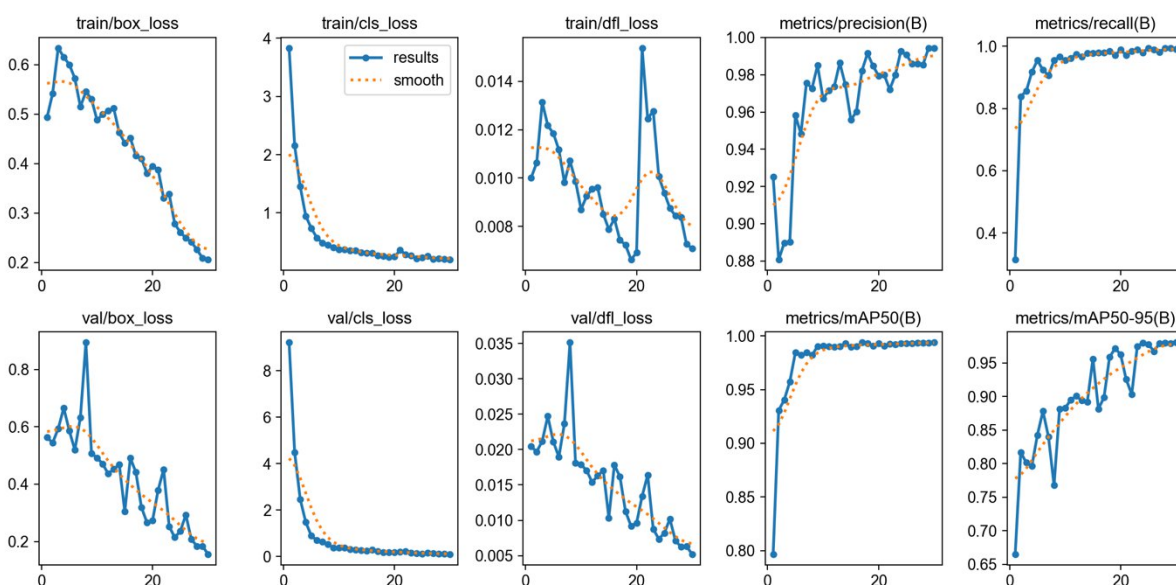


Рисунок 3. Результат обучения модели детектирования документов

Разработка программного обеспечения

Было разработано программное обеспечение с полученными на предыдущих этапах нейросетями семейства YOLO. Основной процесс обработки включает три последовательно взаимодействующих модуля (рис. 4). Модуль получения видеоданных обеспечивает подключение к источнику, чтение видеопотока и его предварительную обработку. Модуль обнаружения и сокрытия чувствительной информации выполняет поиск объектов и их маскирование. Модуль передачи обработанных данных формирует выходной видеопоток и публикует его.

Дополнительно в состав системы входят модули конфигурирования и журналирования. Конфигурационный модуль обеспечивает управление параметрами источника, моделей и методов сокрытия информации. Модуль журналирования фиксирует события и ошибки, возникающие в процессе работы программы.

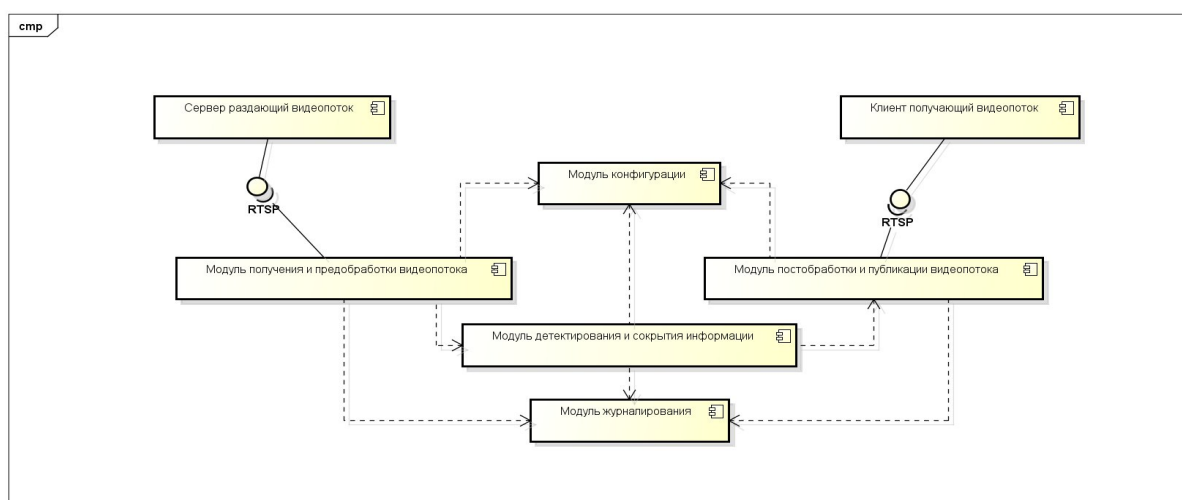


Рисунок 4. Общая архитектура программной системы в UML нотации

В результате работы система автоматически определяет области объектов, относящихся к чувствительной информации, после чего скрывает их. Результат работы системы можно увидеть на рисунке 5.



Рисунок 5. Пример работы системы

Выводы

В работе предложен подход к автоматизированному поиску и сокрытию чувствительной информации в видеопотоке. Для обнаружения целевых объектов использованы модели нейронных сетей семейства YOLO. В рамках работы перечень чувствительных объектов был ограничен тремя классами: лицами людей, экранами мониторов и документами. Для всех классов были предложены отдельные независимые детекторы, что позволит настроить отдельное детектирование для каждого класса. Дальнейшее расширение списка чувствительных объектов будет происходить проще, без переобучения одной конкретной модели.

В рамках работы был разработан алгоритм, предусматривающий разбиение видеопотока на отдельные кадры. Изображения подготавливаются и направляются в модуль детектирования. Все найденные области фильтруются по порогам уверенности, а после этого агрегируются в один результат для последующего сокрытия. Благодаря данному алгоритму, в видеопотоке скрываются только те области, которые содержат чувствительную информацию.

Было проведено обучение моделей детектирования для документов и лиц людей. Мы получили следующие значения для моделей детектирования лиц: mAP50 – 0,9835, Precision – 0,9715, Recall – 0,9481 и F1-score – 0,9597. Для документов мы получили следующие значения: mAP50 – 0,9937, Precision – 0,9943, Recall – 0,9914 и F1-score – 0,9929. Для детектирования экранов использовалась модель YOLO26n, обученная на наборе данных COCO. Результаты метрик, полученные после обучения моделей, показывают возможность использовать модели нейронных сетей семейства YOLO для решения просматриваемой задачи.

Для описанного алгоритма была разработана программа на языке Python с использованием библиотек OpenCV, Ultralytics, NumPy и Torch. Дальнейшее развитие исследования связано с оптимизацией работы модуля детектирования и межкадровым отслеживанием объектов.

СПИСОК ИСТОЧНИКОВ

1. Гюльбяков Н.Р. Цифровизация в здравоохранении: актуальный обзор литературы / Н.Р. Гюльбяков, Н.В. Алексеева, Х.Н. Гюльбякова // Международный научно-исследовательский журнал. – 2024. – № 11 (149). – URL: <https://research-journal.org/archive/11-149-2024-november/10.60797/IRJ.2024.149.90> (дата обращения: 28.05.2026).
2. О персональных данных: Федеральный закон от 27.07.2006 № 152-ФЗ (ред. от 24.06.2025): [принят Государственной Думой 8 июля 2006 года: одобрен Советом Федерации 14 июля 2006 года] // КонсультантПлюс. – URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (дата обращения: 28.05.2026).
3. Lei Geral de Proteção de Dados Pessoais (LGPD) // Gov.br [Электронный ресурс]. – URL: <https://www.gov.br/planalto/pt-br/vice-presidencia/acesso-a-informacao/lgpd> (дата обращения: 28.05.2026).
4. О безопасности данных: Закон КНР: [принят 29-м заседанием Постоянного комитета Всекитайского собрания народных представителей 13-го созыва 10 июня 2021 г.: вступил в силу 1 сентября 2021 г.] // Всекитайское собрание народных представителей Китайской Народной Республики. – URL: http://en.npc.gov.cn.cdurl.cn/2021-06/10/c_689311.htm (дата обращения: 20.06.2026).
5. О защите персональной информации: Закон КНР: [принят 30-м заседанием Постоянного комитета Всекитайского собрания народных представителей 13-го созыва

20 августа 2021 г.: вступил в силу 1 ноября 2021 г.] // Государственный реестр КНР. – URL: <https://www.jsmm.gov.cn/u/cms/www/202110/1317253417oi.pdf> (дата обращения: 20.06.2026).

6. Viola P.A. Robust Real-Time Face Detection / P.A. Viola, M.J. Jones // International Journal of Computer Vision. – 2004. – Vol. 57, No. 2. – P. 137–154.

7. Dalal N. Histograms of Oriented Gradients for Human Detection / N. Dalal, B. Triggs // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, USA. – IEEE, 2005. – Vol. 1. – P. 886–893.

8. Faster R-CNN: towards real-time object detection with region proposal networks / Sh. Ren, K. He, R. Girshick, J. Sun // IEEE Transactions on Pattern Analysis and Machine Intelligence. – Vol. 39, Iss. 6. – P. 1137–1149.

9. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2016. – P. 779–788.

10. YOLOv8-Face: Pre-trained YOLOv8-Face models // GitHub [Электронный ресурс]. – URL: <https://github.com/lindevs/yolov8-face> (дата обращения: 28.05.2026).

11. Face Detection Dataset, version 1 // Roboflow Universe [Электронный ресурс]. – URL: <https://universe.roboflow.com/uniform-fhonp/face-detection-quvrj/dataset/1> (дата обращения: 28.05.2026).

12. Ultralytics YOLO26 // Ultralytics Docs [Электронный ресурс]. – URL: <https://docs.ultralytics.com/models/yolo26/#performance-metrics> (дата обращения: 28.05.2026).

13. Document segmentation v2 // Roboflow Universe [Электронный ресурс]. – URL: <https://universe.roboflow.com/lung-x8el1/document-segmentation-v2-gt86h/dataset/2> (дата обращения: 28.05.2026).

ИНФОРМАЦИЯ ОБ АВТОРАХ

Писарев Андрей Константинович, студент, Волгоградский государственный технический университет, Волгоград, Россия.

e-mail: andrey-pisarev-0@mail.ru

Андреященко Павел Владимирович, студент, Волгоградский государственный технический университет, Волгоград, Россия.