



<http://dx.doi.org/10.35596/1729-7648-2025-31-4-47-54>

УДК 004.89

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА КОНТРОЛЯ ДОСТУПА В УЧЕБНУЮ АУДИТОРИЮ

С. С. ИНДЮКОВ, В. А. ЧУЙКО, Е. И. КОЗЛОВА

Белорусский государственный университет (Минск, Республика Беларусь)

Аннотация. В статье представлена схема интеллектуальной системы контроля доступа, включающая камеру, направленную на вход в помещение, оператора и приложение на компьютере оператора. Сформулированы требования к приложению для системы контроля доступа по распознаванию лиц из различных видеоисточников, удобного в использовании пользовательского интерфейса, модульности для удобства расширения системы. Разработанная интеллектуальная система контроля доступа позволяет преподавателям контролировать посещаемость на занятиях, способна захватывать видео из различных видеоисточников благодаря алгоритмам распознавания и отслеживания, устойчива к низкому качеству видео.

Ключевые слова: интеллектуальная система, изображение, нейронные сети, детектор, трекер, распознаватель, графическая оболочка.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Индюков, С. С. Интеллектуальная система контроля доступа в учебную аудиторию / С. С. Индюков, В. А. Чуйко, Е. И. Козлова // Цифровая трансформация. 2025. Т. 31, № 4. С. 47–54. <http://dx.doi.org/10.35596/1729-7648-2025-31-4-47-54>.

INTELLIGENT ACCESS CONTROL SYSTEM FOR CLASSROOMS

STANISLAV S. INDZIUKOU, VLADISLAV A. CHUYKO, ALENA I. KOZLOVA

Belarusian State University (Minsk, Republic of Belarus)

Abstract. The article presents a diagram of an intelligent access control system, including a camera directed at entrance to the premises, an operator, and an application on the operator's computer. Requirements for an access control system application for facial recognition from various video sources, an easy-to-use user interface, and modularity for easy system expansion are formulated. The developed intelligent access control system allows teachers to control attendance at classes, is capable of capturing videos from various video sources thanks to recognition and tracking algorithms, and is resistant to low video quality.

Keywords: intelligent system, image, neural networks, detector, tracker, recognizer, graphical shell.

Conflict of interests. The authors declare that there is no conflict of interests.

For citation. Indziukou S. S., Chuyko V. A., Kozlova A. I. (2025) Intelligent Access Control System for Classrooms. *Digital Transformation.* 31 (4), 47–54. <http://dx.doi.org/10.35596/1729-7648-2025-31-4-47-54> (in Russian).

Введение

В современном мире с растущим темпом развития цифровизации информационные технологии все активнее входят в повседневную жизнь. Одна из наиболее актуальных сфер применения информационных технологий – кибербезопасность. При помощи различных аппаратно-программных средств информационной безопасности осуществляется контроль доступа к физическим и цифровым объектам: от различных помещений до ценных документов. Основными этапами в получении разных видов доступа являются идентификация, аутентификация и авторизация. Сегодня уже применяются физические средства контроля доступа, такие как ключ-карта, брелок, терминал для ввода пароля и др. Сравнительно недавно появились более прогрессивные системы контроля доступа, использующие для идентификации биометрические данные человека. Отпечатки пальцев, рук, радужной оболочки глаз, запись голоса, цифровая подпись и даже походка могут служить для идентификации человека. Однако у подобных систем есть недостатки. Ключ-карта подвержена копированию при наличии специальных инструментов, а листок с паролем может быть утерян. Множество исследований доказали, что отпечатки пальцев не являются уникальными [1], как и радужка глаза. Голос же с использованием современных нейросетевых технологий можно синтезировать из заранее собранного набора фраз.

В то же время разрабатываются системы контроля доступа на основе изображений человеческих лиц. Данные системы, являясь менее навязчивыми, способны предоставить хороший, а в некоторых случаях лучший способ идентификации человека для его последующей аутентификации. Примерами таких систем служат линии метрополитена и аэропорты, где на входах располагаются камеры видеонаблюдения, по изображениям которых проводится идентификация.

Разработка интеллектуальной системы

Одним из способов идентификации пользователей (людей) является биометрическая идентификация с помощью сканирования и/или фотографии лица [1]. Данный метод универсален, поскольку снимок лица получается с помощью камеры, что позволяет интегрировать метод в уже имеющиеся системы видеонаблюдения. Кроме того, он дистанционный, поэтому можно производить идентификацию пользователей без дополнительных условий или действий с их стороны. Схема работы системы контроля посещаемости представлена на рис. 1.

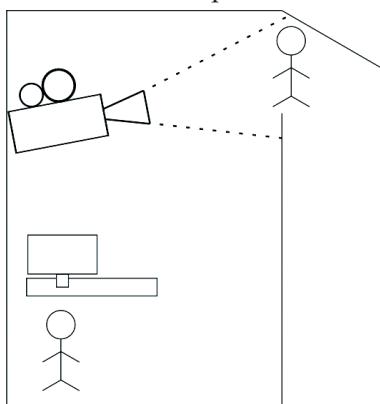


Рис. 1. Схема системы контроля посещаемости
Fig. 1. Attendance control system diagram

Предлагаемый вариант системы реализован и применяется на базе учебной аудитории факультета радиофизики и компьютерных технологий Белорусского государственного университета для контроля посещаемости учебных занятий студентами. Для получения изображения входящих студентов камера располагается перед входом в аудиторию. Полученное изображение отправляется на сервер, после чего на экране администратора отображаются имена студентов, пришедших на занятие. Данный способ контроля посещаемости является более быстрым по сравнению с классической перекличкой и позволяет преподавателю не отвлекаться от занятия.

Разработанная система контроля доступа включает в себя видеокамеры, компьютер оператора/сервер, на котором функционирует нейросетевой алгоритм распознавания и отслеживания

лиц. Серверная часть системы представляет собой компьютерную программу, состоящую из четырех частей: детектора, трекера, распознавателя, графической оболочки.

Детектор, работающий на базе нейросетей YOLO или SCRFD

Среди задач компьютерного зрения выделяют три типа распознавания изображений: классификация, детектирование и сегментация [1]. Суть классификации – определить и рассчитать вероятность присутствия искомого объекта на изображении. Детектирование представляет собой задачу классификации с дополненным выходом, где определяются не только вероятность присутствия объекта, но и координаты границ его местоположения. При сегментации осуществляется соотнесение каждого пикселя исходного изображения с искомым объектом. Таким образом на выходе модели сегментации получается маска присутствия искомого объекта на изображении.

В системе из трех перечисленных типов распознавания используется детектирование, так как в решаемой задаче необходимо знать расположение искомых объектов, лиц студентов на изображении, но нет необходимости в более точном попиксельном определении, предполагаемом сегментацией.

Сегодня лучший механизм для детектирования объектов на изображении – нейронные сети. Нейросети YOLO являются универсальными, поскольку предобучены на множество классов и имеют специализированные библиотеки для дообучения в случае, если необходимый для детектирования класс не был представлен в изначальной обучающей выборке. Также распространенное решение для детектирования человеческих лиц – нейросети семейства SCRFD [2, 3], которые специально обучены под задачу детектирования человеческих лиц.

Разработка рассматриваемой в статье системы велась в два этапа – прототипирование на стационарном компьютере и интеграция в систему видеонаблюдения учреждения образования. На этапе интеграции после проведения тестов производительности стало понятно, что нейросети SCRFD более ресурсозатратные и не позволяют добиться удовлетворительной частоты кадров на используемом оборудовании. Поэтому выбор был сделан в пользу YOLO.

Трекер как реализация алгоритма ByteTrack

Многообъектное отслеживание направлено на определение границ и идентификацию объектов в видеопотоке, поскольку система предназначена для работы в режиме реального времени. Большинство методов получают идентификационные данные путем объединения обнаруженных областей, чьи оценки превышают заданное пороговое значение. Объекты с низкими показателями уверенности, например пересекающиеся, отбрасываются, что влечет пропуски истинных объектов и фрагментацию траекторий. Для повышения точности группы методов многообъектного отслеживания используются все объединения, а не только превышающие пороговую чувствительность с отслеживанием каждой области обнаружения, а не только имеющих высокие оценки, и с применением к ним метода сходства путей (метода сходства и различия [4]), восстанавливаются истинные объекты и фильтруются обнаружения.

Вышеописанные подходы основаны на сохранении лишь обнаруженных сегментов с высокой оценкой. В отличие от них, применяемый в этой части системы метод ByteTrack [5] включает сегменты обнаружения как с высокой, так и с низкой оценками. Изначально блоки обнаружения с высокой оценкой ассоциируются с треклетами (путь объекта). Часть треклотов остается несопоставленной из-за окклюзии, размытости или значительного изменения размера отслеживаемого объекта. Впоследствии сегменты обнаружения с низкой оценкой используются для сопоставления с несопоставленными треклетами, позволяя восстановить объекты в сегментах при фильтрации (являющейся этапом метода ByteTrack) фонового шума.

Распознаватель на основе ArcFace

Традиционно в задачах распознавания лиц используется функция softmax как в качестве функции активации, так и функции потерь. Однако softmax как функция потерь явно не оптимизирует векторы признаков для обеспечения большего сходства для внутриклассовых объектов и разнообразия для межклассовых объектов, что приводит к снижению производительности при распоз-

навании лиц в условиях больших внутриклассовых вариаций. Для устранения этого недостатка в процессе исследований в качестве функции активации была выбрана ArcFace (Additive Angular Margin Loss) [6]. Функция потерь записывается следующим выражением:

$$W^T x_i = \|W_i\| \|x_i\| \cos \theta_j, \quad (1)$$

где $\|W_i\|$ – единичный вес, $\|W_i\|=1$, фиксируется l_2 -нормализацией; $\|x_i\|$ – выделенные признаки, фиксируются l_2 -нормализацией и приводятся к s ; s – радиус; θ_j – угол между весом W_j и признаком x_j .

На этапе нормализации признаков и весов делается предсказание, зависящее от угла между признаками и весами. Таким образом, выделенные признаки распределяются на гиперсфере радиусом s . Затем для одновременного обеспечения внутриклассового сходства и межклассового расхождения добавляется штраф за угловое расстояние между x_i и W_{yj} . В конечном виде функция потерь представляется формулой

$$L_3 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s \cos(\theta_{y_i} + m)}}{e^{s \cos(\theta_{y_i} + m)} + \sum_{j=1, j \neq y_i}^N e^{s \cos \theta_j}}, \quad (2)$$

где θ_{y_i} – целевой угол; m – штраф за угловое отклонение.

Графическая оболочка, реализованная на фреймворке Qt5

Все нейронные сети, использованные в рассматриваемом случае, – сверточные. В них слои свертки выделяют из исходного изображения вектора признаков, после чего дальнейшая работа ведется с данными меньшей размерности.

Для работы алгоритма распознавания была выбрана нейронная сеть YOLOV8 и дообучена, поскольку в базовой реализации возможность детектирования лиц отсутствует. Дообучение подразумевает использование предварительно обученной модели и корректировку ее параметров для улучшения производительности на конкретной задаче или наборе данных. Этот процесс, также известный как переобучение модели, позволяет улучшить понимание и предсказание результатов для конкретных данных, с которыми модель столкнется в разрабатываемом приложении. Для дообучения использовался датасет фотографий знаменитостей CelebA [7]. С целью удобства работы пользователя и администратора с приложением был создан интерфейс программы, вид которого на мониторе оператора с выделенной областью распознавания представлен на рис. 2.

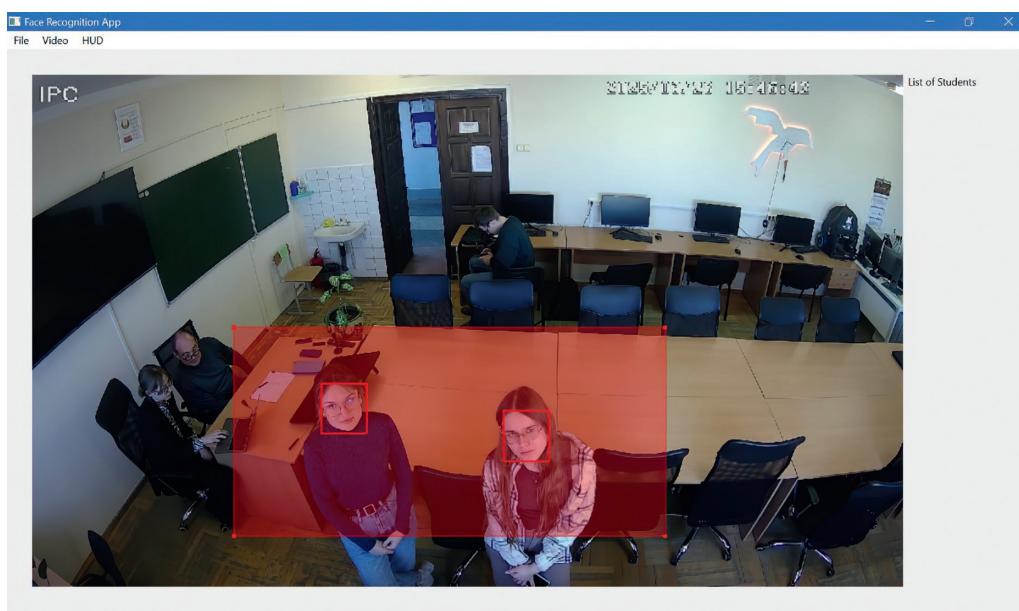


Рис. 2. Интерфейс оператора программы распознавания лиц (выделена область распознавания)

Fig. 2. Interface of the face recognition program operator (recognition area is allocated)

Конвейер работы предлагаемого приложения изображен на рис. 3.

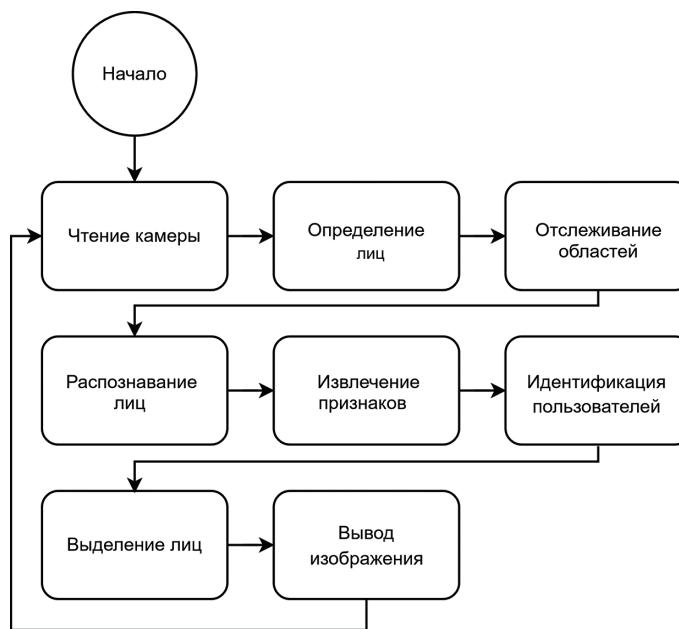


Рис. 3. Конвейер работы программы
Fig. 3. Program execution pipeline

Работа приложения состоит из следующих этапов.

1. Инициализация приложения, загрузка конфигурации/выбор источника видео.
 2. Детектирование лиц на изображении.
 - 2.1. Изображение из сетевой камеры через протокол RTSP (от англ. Real Time Stream Protocol – протокол стриминга (потоковой передачи) в реальном времени) поступает в программу.
 - 2.2. Нейросеть для детектирования лиц обрабатывает изображение и определяет области, содержащие человеческие лица.
 3. Отслеживание. Распознанные области передаются на вход трекера. Трекер сопоставляет ранее распознанные области и восстанавливает пути, относящиеся к ним, в случае неуспешного распознавания.
 4. Распознавание.
 - 4.1. Отслеживаемые и распознанные области сопоставляются. При сходстве более 90 % изображение помечается как актуальное.
 - 4.2. Актуальные изображения подаются на вход нейросети для выделения векторов признаков пользователей.
 5. Идентификация. С использованием метода максимального правдоподобия определяется принадлежность вектора признаков к пользователю, чьи данные были зарегистрированы.
 6. Графический интерфейс.
 - 6.1. На исходном изображении распознанные лица выделяются прямоугольными областями разного цвета, список имен пополняется новыми пользователями.
 - 6.2. Готовое изображение отображается на графическом интерфейсе приложения.
- Камера, используемая в данной системе, располагается таким образом, чтобы полностью захватывать вход в аудиторию. Система позволяет располагать камеру на различных расстояниях и углах к проходу. В случае малого разрешения камеры, например 640×480 пикселей, ее следует располагать ближе, а камеры с большим разрешением могут быть расположены дальше от входа пропорционально используемому разрешению.
- При использовании широкоугольных объективов или камер с большим полем зрения могут возникать геометрические искажения, такие как бочкообразная или подушкообразная дисторсия, что усложняет точное определение положения и размеров объектов вблизи краев кадра. Для выравнивания конечного изображения применяются алгоритмы коррекции дисторсии, например, на основе калибровочных матриц или полиномиальных моделей.

Сверточные нейронные сети позволяют нивелировать некоторые дефекты изображения благодаря своей способности выделять признаки при умеренном уровне шума или размытия. Однако при сильных искажениях или низком освещении эффективность детекции может снижаться.

Предварительно настраивается область работы алгоритма выделения лиц на изображении. Это позволяет увеличить точность процесса распознавания лиц на изображении, затратив тот же объем вычислительных ресурсов на заведомо необходимую область обнаружения. В случае отсутствия выделенных областей на детектор передается полное изображение с камеры (этапы 1, 2), после чего его размерность снижается до разрешения 112×112 пикселей. В случае присутствия выделенных областей на детектор подается не полное изображение, а его выделенные фрагменты. Таким образом, до принимаемого нейронной сетью формата 112×112 пикселей сжимается изображение меньшего разрешения, что позволяет лучше сохранить информацию при уменьшении размерности [8].

В случае наличия нескольких областей для детектирования выделение объектов в каждой из них происходит отдельно, т. е. количество проходов нейронной сети, а, следовательно, и нагрузка, увеличиваются пропорционально. В первой итерации программы распознавание, детектирование и обработка графического интерфейса осуществлялись на одном потоке, однако данная реализация не обеспечивала достаточной частоты кадров для полноценного отслеживания движущихся целей. Для повышения производительности приложения было решено разделить нагрузку на три потока: поток графического интерфейса, поток детектирования и трекинга, поток распознавания [9].

Поток графического интерфейса производит начальную инициализацию приложения: загружает файлы конфигураций, подключается к камерам, создает остальные потоки программы.

Поток детектирования и трекинга получает изображения с камер, обрабатывает выделенные области в случае их присутствия, детектирует лица на изображениях. После процесса детектирования области, в которых были обнаружены лица, передаются трекеру. Последний сопоставляет ранее обнаруженные и новые области и строит пути их следования.

Поток распознавания работает с областями, обнаруженными детектором и трекером. В случае совпадения последних областей детектирования и трекинга распознанная область считается важной, и изображение из данной области подается на нейросеть-распознаватель.

Далее, если коэффициент уверенности предсказания нейронной сети больше заданного порогового значения, распознанный пользователь заносится в список успешно распознанных. Этот коэффициент рассчитывается по формуле

$$c = \max(xX^T), \quad (3)$$

где x – вектор признаков из области; X – матрица признаков всех пользователей.

Матрица признаков получается выше порогового значения, распознавание заносится в список успешных. Список успешных распознаваний передается обратно на поток графического интерфейса, где распознанные области лиц обводятся прямоугольником, показывая, что лицо распознано, а имена распознанных людей отмечаются в списке.

Для формирования базы данных пользователей (студентов) необходимо сделать фотографии лиц пользователей с различных ракурсов. В данной реализации одно лицо фотографируется с 20 ракурсов: пять ракурсов по горизонтали, на каждый из которых делается по четыре ракурса по вертикали. Таким образом, полученные 20 фотографий составляют виртуальный 3D-скан лица пользователя. Далее все фотографии пользователя проходят через нейросеть-распознаватель для получения векторов признаков. Полученные векторы признаков заносятся в список пользователей под именем пользователя и участвуют в дальнейших сопоставлениях при распознавании. Поскольку характеристики лица, используемые в вышеописанной системе, подпадают под определение биометрических персональных данных, их применение возможно только с письменного или электронного согласия субъекта [10].

Заключение

1. Спроектирована, разработана и интегрирована в учебный процесс факультета радиофизики и компьютерных технологий Белорусского государственного университета интеллектуальная сис-

тема контроля посещаемости студентами занятий. Благодаря ей время контроля преподавателем посещаемости снизилось с 10 мин до 10 с. Интеллектуальная система позволяет беспрепятственно проверять студентов и преподавателей на рабочем месте, без привлечения сторонних лиц.

2. Разработанная система имеет ряд достоинств, среди которых – гибкость, интегрируемость и беспрепятственность, что позволяет применять ее в других сценариях, где невозможна или недопустима остановка каждого отдельного человека. Систему можно использовать в больницах для непрерывного отслеживания пациентов и их состояния, распознавания лиц в потоках людей в метро, аэропортах и других общественных местах для обнаружения нежелательных лиц, занесенных в реестр разыскиваемых, и т. д. [2].

3. Работа выполнена в рамках задания Государственной программы научных исследований «Фотоника и электроника для инноваций» (№ ГР 20212701).

Список литературы

1. Головатая, Е. А. Нейросетевые технологии в обработке и защите данных / Е. А. Головатая, А. В. Курочкин. Минск: Белор. гос. ун-т, 2021. С. 51–61.
2. Sample and Computation Redistribution for Efficient Face Detection [Electronic Resource] / Jia Guo [et al.]. Mode of access: <https://openreview.net/pdf?id=RhB1AdoFfGE>. Date of access: 09.03.2025.
3. Yaseen, M. What Is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector [Electronic Resource] / M. Yaseen. Mode of access: https://www.researchgate.net/publication/383494295_What_Is_YOLOv8_An_In-Depth_Exploration_of_the_Internal_Features_of_the_Next-Generation_Object_Detector. Date of access: 09.03.2025.
4. ByteTrack: Multi-Object Tracking by Associating Every Detection Box [Electronic Resource] / Y. Zhang [et al.]. Mode of access: <https://github.com/ifzhang/ByteTrack>. Date of access: 10.09.2024.
5. Милль, Д. С. Система логики силлогистической и индуктивной. Изложение принципов доказательства в связи с методами научного исследования / Д. С. Милль, пер. с англ. М.: Изд-во магазина «Книжное дело», 1900.
6. ArcFace: Additive Angular Margin Loss for Deep Face Recognition [Electronic Resource] / J. Deng [et al.]. Mode of access: <https://paperswithcode.com/paper/arcface-additive-angular-margin-loss-for-deep?ref=broutonlab.com/>. Date of access: 11.11.2024.
7. Large-Scale CelebFaces Attributes (CelebA) Dataset [Electronic Resource] / Z. Liu [et al.]. Mode of access: <https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>. Date of access: 10.01.2025.
8. О защите персональных данных: закон Республики Беларусь от 7 мая 2021 года № 99-З [Электронный ресурс]. Режим доступа: <https://pravo.by/document/?guid=12551&p0=H12100099>. Дата доступа: 09.03.2025.
9. Train with Mixed Precision [Electronic Resource] // NVIDIA Deep Learning Performance. Mode of access: <https://docs.nvidia.com/deeplearning/performance/mixed-precision-training/index.html>. Date of access: 30.03.2025.
10. Davis, J. Understanding Mixed Precision Training [Electronic Resource] / J. Davis // TDS Archive. Mode of access: <https://medium.com/data-science/understanding-mixed-precision-training-4b246679c7c4>. Date of access: 10.04.2025.

Поступила 16.05.2025

Принята в печать 31.07.2025

Доступна на сайте 12.01.2026

References

1. Golovataya E. A., Kurochkin A. V. (2021) *Neural Network Technologies in Data Processing and Protection*. Minsk, Belarusian State University. 51–61 (in Russian).
2. Guo J., Deng J., Lattas A., Zafeiriou S. (2025) *Sample and Computation Redistribution for Efficient Face Detection*. Available: <https://openreview.net/pdf?id=RhB1AdoFfGE> (Accessed 9 March 2025).
3. Yaseen M. (2025) *What Is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector*. Available: https://www.researchgate.net/publication/383494295_What_is_YOLOv8_An_In-Depth_Exploration_of_the_Internal_Features_of_the_Next-Generation_Object_Detector (Accessed 9 March 2025).
4. Zhang Y., Sun P., Jiang Y., Yu D., Weng F., Yuan Z., et al. (2024) *ByteTrack: Multi-Object Tracking by Associating Every Detection Box*. Available: <https://github.com/ifzhang/ByteTrack> (Accessed 10 September 2024).
5. Mill J. S. (1900) *A System of Syllogistic and Inductive Logic. An Exposition of the Principles of Proof in Relation to the Methods of Scientific Research*. Moscow, Publishing House of the “Knizhnoe Delo” (Book Business) Store (in Russian).

6. Deng J., Guo J., Yang J., Xue N., Kotsia I., Zafeiriou S. (2024) *ArcFace: Additive Angular Margin Loss for Deep Face Recognition*. Available: <https://paperswithcode.com/paper/arcface-additive-angular-margin-loss-for-deep?ref=broutonlab.com/> (Accessed 11 November 2024).
7. Liu Z., Luo P., Wang X., Tang X. (2025) *Large-Scale CelebFaces Attributes (CelebA) Dataset*. Available: <https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html> (Accessed 10 January 2025).
8. *On the Protection of Personal Data, Law of the Republic of Belarus of May 7, 2021 No 99-Z*. Available: <https://pravo.by/document/?guid=12551&p0=H12100099> (Accessed 9 March 2025) (in Russian).
9. Train with Mixed Precision. *NVIDIA Deep Learning Performance*. Available: <https://docs.nvidia.com/deeplearning/performance/mixed-precision-training/index.html> (Accessed 30 March 2025).
10. Davis J. (2025) Understanding Mixed Precision Training. *TDS Archive*. Available: <https://medium.com/datascience/understanding-mixed-precision-training-4b246679c7c4> (Accessed 10 April 2025).

Received: 16 May 2025

Accepted: 31 July 2025

Available on the website: 12 January 2026

Вклад авторов

Индюков С. С. осуществил разработку модели интеллектуальной системы и подготовил рукопись статьи.

Чуйко В. А. определил концепцию работы, осуществил критический анализ модели интеллектуальной системы и текста статьи.

Козлова Е. И. провела критический анализ содержания статьи, подготовила окончательный вариант рукописи для публикации.

Authors' contribution

Indziukou S. S. developed the model of the intelligent system and prepared the manuscript of the article.

Chuyko V. A. defined the concept of the work, carried out a critical analysis of the model of the intelligent system and the text of the article.

Kozlova E. I. conducted a critical analysis of the content of the article and prepared the final version of the manuscript for publication.

Сведения об авторах

Индюков С. С., студ., Белорусский государственный университет (БГУ)

Чуйко В. А., магистр физ.-мат. наук, ст. преп. каф. интеллектуальных систем, БГУ

Козлова Е. И., канд. физ.-мат. наук, доц., зав. каф. интеллектуальных систем, БГУ

Information about the authors

Indziukou S. S., Student, Belarusian State University (BSU)

Chuyko V. A., Mr. Sci. (Phys. and Math.), Senior Lecturer at the Department of Intelligent Systems, BSU

Kozlova A. I., Cand. Sci. (Phys. and Math.), Associate Professor, Head of the Department of Intelligent Systems, BSU

Адрес для корреспонденции

220064, Республика Беларусь,
Минск, ул. Курчатова, 5
Белорусский государственный университет
Тел.: +375 29 853-07-96
E-mail: Vchuyko@bsu.by
Чуйко Владислав Александрович

Address for correspondence

220064, Republic of Belarus,
Minsk, Kurchatova St., 5
Belarusian State University
Tel.: +375 29 853-07-96
E-mail: Vchuyko@bsu.by
Chuyko Vladislav Aleksandrovich