

**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

Объект авторского права  
УДК 004.932

**ИГНАТЬЕВА**  
**Светлана Александровна**

**ПОВТОРНАЯ ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА НА ИЗОБРАЖЕНИЯХ  
СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ С ИСПОЛЬЗОВАНИЕМ  
СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ**

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата технических наук  
по специальности 05.13.01 – системный анализ,  
управление и обработка информации

Минск 2025

Работа выполнена в учреждении образования «Полоцкий государственный университет имени Евфросинии Полоцкой»

Научный руководитель –

**Богуш Рихард Петрович,**  
доктор технических наук, доцент,  
заведующий кафедрой вычислительных систем  
и сетей учреждения образования  
«Полоцкий государственный университет  
имени Евфросинии Полоцкой».

Официальные оппоненты:

**Тузиков Александр Васильевич,**  
член-корреспондент НАН Беларуси,  
доктор физико-математических наук,  
профессор, заведующий лабораторией  
математической кибернетики  
государственного научного учреждения  
«Объединенный институт проблем  
информатики НАН Беларуси»;

**Лукашевич Марина Михайловна,**  
кандидат технических наук, доцент,  
доцент кафедры информационных  
систем управления Белорусского  
государственного университета.

Оппонирующая организация – Учреждение образования  
«Брестский государственный  
технический университет».

Защита состоится 20 февраля 2025 года в 16.30 часов на заседании совета по защите диссертаций Д 02.01.02 при Белорусском государственном университете по адресу: 220030, г. Минск, ул. Ленинградская 8, корпус юридического факультета, ауд. 407.

Телефон ученого секретаря: 209 57 09; e-mail: Kochyn@bsu.by.

С диссертацией можно ознакомиться в Фундаментальной библиотеке Белорусского государственного университета.

Автореферат разослан «16» января 2025 года.

Ученый секретарь  
совета по защите диссертаций Д 02.01.02,  
кандидат технических наук, доцент



В.П.Кочин

## ВВЕДЕНИЕ

Широкое распространение систем видеонаблюдения позволяет решать множество практических задач. Среди них важной является автоматизация повторной идентификации человека (реидентификации, *англ.* Person Re-Identification). Под повторной идентификацией понимают определение по запросу присутствия идентифицированного ранее системой видеонаблюдения человека на одной и той же сцене в разное время или в различных местах видеонаблюдения. В качестве запроса, как правило, выступают изображения людей. Поэтому при практической реализации должны быть выявлены их отличительные особенности. Сравнение признаков запроса с дескрипторами множества изображений (галереи) людей позволяет выполнить реидентификацию. Причем состав набора признаков в значительной мере определяет эффективность повторной идентификации. Поиск и выделение наиболее отличительных особенностей объектов на изображениях, в том числе и людей, не формализован и обусловлен рядомотягающих факторов: неоднозначность внешнего вида при разных ракурсах съемки, степень освещенности, различия в разрешении камер, перекрытие человека другими людьми или объектами фона, схожесть внешнего вида разных людей, например, при использовании униформы. Для реидентификации человека эмпирический подход в поиске наиболее эффективных признаков требует нерационально больших вычислительных затрат, в связи с чем долгое время эта задача оставалась нерешенной. Совершенствование средств вычислительной техники и открытия в области глубокого обучения, в частности сверточных нейронных сетей (СНС), позволили автоматизировать процесс извлечения признаков на изображениях людей и обеспечить увеличение точности реидентификации.

Эффективность применения СНС для повторной идентификации человека зависит от множества факторов: архитектуры СНС, подходов к обучению и выбранных для этого гиперпараметров, размера и разнообразия обучающей выборки, применяемого алгоритма. Подходы для повышения точности реидентификации людей можно разделить на две группы: применяемые на этапе обучения СНС и на этапе постобработки извлеченных признаков. На этапе обучения повысить точность реидентификации можно путем выбора архитектуры СНС, использования наиболее эффективных значений гиперпараметров, увеличивая размер и разнообразие обучающих данных, используя различные функции потерь. На этапе постобработки извлеченных признаков на точность реидентификации будет оказывать влияние алгоритм, определяющий вид признаков (глобальные, локальные, дополнительные и т.д.) и методы ранжирования.

В развитие алгоритмов реидентификации с использованием СНС значительный вклад внесли Z. Zheng, L. Zheng, предложившие базовый алгоритм

реидентификации, а также Y. Yang, Y. Mang, J. Shen, G. Lin, T. Xiang, L. Shao. В Беларуси в области обработки изображений значительный вклад внесли С. В. Абламейко, А. В. Тузиков, В. В. Краснопрошин, В. В. Старовойтов, В. А. Головкин, А. М. Недзведь.

Однако широкое разнообразие областей применения реидентификации, таких как поиск пропавших людей или правонарушителей, сбор статистической информации о посещении людьми торговых центров, учреждений здравоохранения, образования и других социально-значимых объектов, приводит к существованию большого числа подходов и алгоритмов для ее решения, которые имеют разные качественные характеристики, причем точность их намного ниже максимально возможной. В связи с этим актуальным является разработка новых эффективных подходов для повторной идентификации человека на изображениях систем видеонаблюдения.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Связь работы с научными программами (проектами) и темам.**

ГПНИ «Цифровые и космические технологии, безопасность человека, общества и государства», задание 1.1.5 «Методы глубокого обучения и кроссплатформенные решения для интеллектуального анализа изображений», НИР «Алгоритмы и модели для обнаружения объектов и оценки их параметров по последовательностям изображений с использованием методов глубокого обучения», ГР № 20211439, 2021–2025 гг.

НИР «Алгоритмы обработки данных на основе дискретных преобразований, нечеткой логики и нейронных сетей», ГР № 20180570, 2018–2022 гг.

### **Цель и задачи исследования.**

*Целью диссертационной работы* является увеличение точности повторной идентификации человека в распределенной системе видеонаблюдения.

Для достижения данной цели необходимо решить следующие *задачи*:

- выполнить анализ существующих методов и алгоритмов, применяемых для повторной идентификации человека;
- провести экспериментальное исследование влияния активационной функции и гиперпараметров на динамику обучения СНС и эффективность формируемых дескрипторов при реидентификации человека;
- сформировать репрезентативное множество изображений людей, полученных при различных условиях видеонаблюдения для обучения и тестирования алгоритмов повторной идентификации;
- разработать метод аугментации, позволяющий повысить разнообразие обучающей выборки без увеличения ее размера;
- разработать технологию обучения сверточных нейронных сетей с использованием аугментированных изображений;

– разработать алгоритм повторной идентификации человека, позволяющий повысить точность при перекрытии изображения объектами снимаемой сцены или другими людьми;

– разработать программное обеспечение повторной идентификации человека по изображениям для интеллектуальных систем видеонаблюдения.

*Объектом исследований* являются изображения людей, полученные с камер видеонаблюдения.

*Предметом исследований* являются методы и алгоритмы повторной идентификации человека с использованием сверточных нейронных сетей.

### **Научная новизна.**

1. Установлено влияние функций активации в СНС на точность повторной идентификации человека на изображениях систем видеонаблюдения. Показано, что среди активационных функций ReLU, Leaky-ReLU, PReLU, RReLU, ELU, SELU, GELU, Swish, Mish, наибольшую точность в метриках Rank1, Rank5, Rank10 и mAP обеспечивают ReLU и GeLU, но минимальная скорость обучения СНС достигается при использовании ReLU.

2. Разработан метод аугментации изображений людей для повторной идентификации человека, в котором применяется ряд преобразований для изображений из пакетов, подаваемых на вход обучаемой СНС, позволяющий улучшить обобщающую способность СНС за счет уменьшения чувствительности к изменениям освещенности, местоположению ключевых отличительных признаков, низкому разрешению изображения, возможности регуляризации при обучении СНС.

3. Предложена технология двухэтапного обучения СНС, которая для предварительного этапа предполагает использовать наборы изображений, расширенные за счет аугментации, а для точной настройки весовых коэффициентов применять обучение на изображениях без аугментации и с меньшей скоростью обучения. Применение таких двух этапов способствует снижению переобучения СНС и уменьшению значения функции потерь.

4. Разработан алгоритм повторной идентификации человека на изображениях, который применяет составной нейросетевой дескриптор, включающий вектора признаков всего изображения, его верхней, средней и нижней частей, которые перекрываются, и замещение признаков для скрытых областей на основе наиболее схожих изображений, определяемых по видимым частям фигуры человека. Алгоритм позволяет повысить точность повторной идентификации при скрытии части изображения человека другими объектами.

### **Положения, выносимые на защиту:**

*Метод аугментации изображений людей* для повторной идентификации, отличающийся применением циклического сдвига изображений, исключением цветности и замещением фрагментов изображения уменьшенными копиями

изображений, который повышает робастность сверточных нейронных сетей к ошибкам детектора, изменению освещенности, уменьшению разрешения изображения, местоположению ключевых отличительных признаков.

*Технология двухэтапного обучения сверточных нейронных сетей, отличающаяся* применением аугментации на предварительном этапе для расширения тренировочной выборки изображений и использованием изображений из исходного набора на этапе точной настройки весовых коэффициентов, которая позволяет повысить точность повторной идентификации при применении нейронных сетей ResNet-50 и DenseNet-121 по метрикам Rank1 до 21 %, mAP до 31 % и mINP до 59 % относительно базового алгоритма реидентификации.

*Алгоритм повторной идентификации человека, использующий составной* нейросетевой дескриптор из глобальных и локальных признаков, который *отличается* применением перекрытия локальных областей фигуры человека при формировании их признаков, анализом ключевых точек при локализации и определении видимости каждой области, вычислением дескриптора скрытой области на основе усреднения значений признаков фрагментов изображений. Применение предложенного алгоритма позволило увеличить точность реидентификации относительно базового алгоритма при использовании нейронных сетей ResNet-50 и DenseNet-121 по метрикам Rank1 до 10 %, mAP до 14 % и mINP до 49 %.

#### **Личный вклад соискателя ученой степени.**

Основные результаты, изложенные в диссертационной работе, получены автором самостоятельно. Вклад научного руководителя связан с определением темы и целей диссертационного исследования, выбором методов исследования, анализом и обобщением полученных результатов. В представленных публикациях с соавторами вклад соискателя определяется рамками излагаемых в диссертации результатов.

#### **Апробация результатов диссертации.**

Основные результаты исследований докладывались и обсуждались на следующих конференциях и семинарах: Pattern Recognition and Information Processing (PRIP'2021, PRIP'2023), Minsk; Proceedings of the IV International Workshop on Computer Modeling and Intelligent Systems CMIS–2021, Zaporizhzhia, Ukraine; XIII Junior Researchers' Conference, Novopolotsk, 2021; XII Республиканская научная конференция молодых ученых и студентов «Современные проблемы математики и вычислительной техники»: Брест, 2021; XXII Международная конференция «Информатика: проблемы, методы, технологии» IPMT–2022, Воронеж, Россия, 2022; II Международная научно-практическая конференция «Информационно-коммуникационные технологии: достижения, проблемы, инновации» ИКТ–2022, Полоцк, Беларусь, 2022; 2nd International Conference on Applied Mathematics, Modeling and Computer Simulation, AMMCS 2022, Wuhan,

China, 2022; выставка-форум IT-Академграда «Искусственный интеллект Беларуси»: Минск, ОИПИ НАН Беларуси, 2022, 2023; Международный конгресс «Информационные системы и технологии» CSIST'2022, Минск, БГУ, 2022; научно-практический семинар «Методы обработки сигналов в технической диагностике сложных систем», Минск, 2022, 2023; Международный научно-технический семинар «Технологии передачи и обработки информации», Минск, 2023; IX Международная конференция и молодежная школа «Информационные технологии и нанотехнологии» (ИТНТ–2023), Самара, Россия, 2023; Международная конференция по перспективным достижениям в информатике, информационным системам и сетям (ETISSN 2024), Баку, Азербайджан, 2024.

Имеется акт о практическом применении результатов научно-исследовательской работы в ООО «Ай Ти Ви групп», которое занимается разработкой и внедрением интеллектуальных систем видеоаналитики во многих странах мира и акт о внедрении результатов исследований в учебный процесс в учреждении образования «Полоцкий государственный университет имени Евфросинии Полоцкой».

#### **Опубликование результатов диссертации.**

По материалам диссертации опубликовано 22 научные работы: 9 статей в научных журналах в соответствии с пунктом 19 Положения ВАК о присуждении ученых степеней и присуждении ученых званий; 13 статей в материалах научных конференций.

#### **Структура и объем диссертации.**

Диссертация состоит из введения, пяти глав, заключения, списка использованных источников и приложения. Полный объем диссертации составляет 159 страниц, в том числе 49 рисунков занимают 16 страниц, 18 таблиц на 9 страницах, четыре приложения занимают 27 страниц. Список использованных источников содержит 160 наименований на 14 страницах, включая 22 собственные публикации соискателя на 3 страницах.

## **ОСНОВНАЯ ЧАСТЬ**

Во *введении* дана краткая характеристика исследуемой области и обоснована актуальность темы диссертационной работы.

*Первая глава* посвящена анализу методов и алгоритмов при повторной идентификации человека в распределенных системах видеонаблюдения. На основе проведенного анализа выполнена классификация существующих систем реидентификации человека, по взаимодействию с внешней средой, по типу анализируемых данных, по количеству и типу запроса, по времени работы системы. Описываются основные особенности каждого из рассмотренных классов систем и характерные проблемы. Приводится анализ основных проблем повторной идентификации и возможные пути снижения их влияния на точность.

Рассматриваются метрики Rank1, Rank5, Rank10, mAP и mINP, которые используются для оценки точности систем повторной идентификации.

Приводится классификация признаков людей на изображениях. Представлен анализ существующих наборов изображений и наиболее используемых СНС при реидентификации.

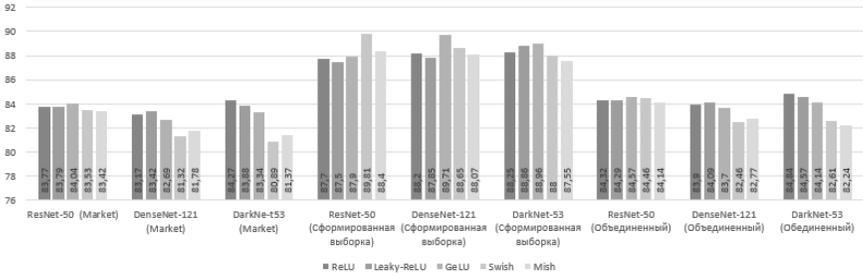
**Вторая глава** посвящена исследованию влияния на точность повторной идентификации человека типа функции активации (ФА) нейронов, гиперпараметров при обучении СНС и ранжирования извлеченных признаков.

Повторная идентификация схожа с классификацией, однако существенное отличие заключается в отсутствии строго фиксированного количества классов объектов для обучающих и тестовых выборок. В связи с этим, для ФА, показавших свою эффективность для классификации и обнаружения объектов на изображениях проведены исследования о влиянии их на точность реидентификации. При этом использовался базовый алгоритм повторной идентификации человека, формирующий глобальный вектор признаков на основе СНС, для обучения которой не применяется аугментация и другие методы, направленные на улучшение обобщающей способности. ФА ReLU, Leaky-ReLU, PReLU, RReLU, ELU, SELU, GELU, Swish, Mish оценивались по метрикам Rank1, Rank5, Rank10, mAP, а также по времени обучения модели. Первый этап экспериментов направлен на выявление наиболее эффективных ФА с использованием базового алгоритма и СНС ResNet-50, DenseNet-121 и DarkNet-53 на наборе данных Market-1501.

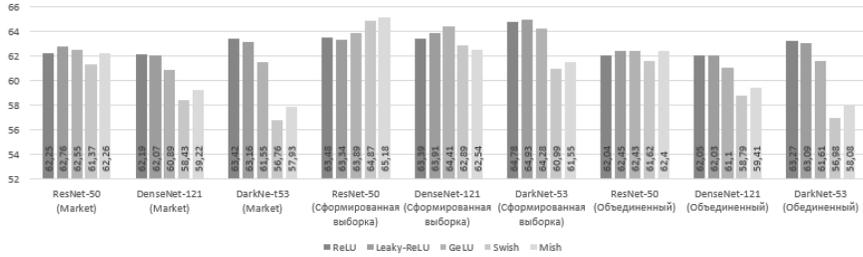
С учетом полученных результатов, на втором этапе исключены наименее эффективные ФА, такие как PReLU, RReLU, ELU, SELU. Для оценки воспроизводимости результатов точности повторной идентификации людей каждая СНС с разными ФА была обучена трижды с одинаковыми гиперпараметрами на объединенном наборе данных, состоящем из Market-1501 и сформированной выборки, включающей 16 770 изображений для 145 человек. Тестирование выполнено отдельно на данных Market-1501, созданной тестовой выборке из 15 149 изображений для 126 человек и на объединенной с последующим определением среднего арифметического значений метрик.

На рисунке 1 представлены результаты исследования. Анализ показывает, что наиболее перспективными функциями активации СНС при реидентификации являются ReLU и GeLU, при этом ReLU имеет меньшую вычислительную сложность, а, значит, обеспечивается большая скорость работы СНС, чем при использовании GeLU.

Выполнен анализ влияния гиперпараметров на точность повторной идентификации человека и определены наиболее эффективные для этой задачи значения скорости обучения СНС ResNet-50 и размера пакета изображений.



a



b

Рисунок 1 – Точность повторной идентификации человека при использовании СНС с разными ФА в метриках: a – Rank1; б – mAP

Показано, что для повышения точности реидентификации человека может быть использовано повторное ранжирование признаков изображений после первой их сортировки. Ранжированный первичный список изображений для запроса  $I_Q^{(V_0)}$  формируется с применением метрики косинусного подобия. Для каждого из  $k$  наиболее схожих с запросом изображений из полученного списка определяется  $k$ -ближайших соседей из галереи на основе метрики косинусного подобия. Применяя поочередное пересечение данных из первичного списка со списками для  $k$  наиболее схожих с запросом формируются наборы  $k_{nb}$ -взаимно-ближайших соседей, причем  $k < k_{nb}$ . Полученные таким образом изображения ранжируются с учетом их наличия в первичном списке на основе метрики Жаккара. Увеличение точности повторной идентификации с применением повторного ранжирования подтверждается результатами исследований на наборах данных Market-1501, DukeMTMC-ReID, PolReID с использованием СНС ResNet-50 и DenseNet-121 до 9% по метрике Rank1, на 10% – 25% по метрике mAP и на 43% – 112% по метрике mINP.

В *третьей главе* рассматривается сформированный набор данных для обучения СНС при повторной идентификации человека, предложены метод аугментации данных и технология двухэтапного обучения СНС.

Новый набор данных для реидентификации PolReID1077 включает 45 971 изображение для 1 077 различных людей. При формировании комплекта

изображений человека обрабатывалось до 9 видеопоследовательностей. Каждый человек зафиксирован видеокамерами с разных ракурсов в нескольких местах, при этом использовалось до 10 вариантов расстановки камер. Всего для создания PolReID1077 применялся 1 331 вариант размещения видеокамер с отличающимися характеристиками, разрешением и частотой кадров. Видеонаблюдение осуществлялось при разных погодных условиях в течение четырех сезонов, при этом зимой получено 30% изображений людей, летом – 29%, осенью и весной – 41%. Наличие у разных людей различных атрибутов в виде сумок, пакетов, рюкзаков, портфелей, продуктовых корзин, шарфов, шапок, очков, папок для бумаг, тетрадей, телефонов, наушников, еды и напитков изменяется с течением времени, что дает возможность изучать и учитывать незначительные изменения внешности. Такой подход позволяет получить более разнообразные примеры из-за большого количества вариантов фоновых помех, уровней освещения, одежды по сезону.

Разработан метод аугментации данных, который включает циклический сдвиг пикселей изображения, исключение цветности и замещение фрагмента изображения уменьшенной копией другого изображения из пакета, подаваемого на вход СНС. Циклический сдвиг и обесцвечивание применяется только к некоторым изображениям, а замещение фрагмента уменьшенной копией другого изображения применяется ко всем изображениям в пакете. Чтобы определить необходимость выполнения преобразований перед обучением СНС вводится гиперпараметр  $p$ , определяющий к какой части от общего количества изображений применять изменения. На основе  $p$  определяется пороговое значение  $t = \frac{p}{100}$ .

С использованием генератора псевдослучайных чисел (ГПСЧ) генерируется число  $r \in (0; 1)$  и, если  $r < t$ , то изображение трансформируется, в противном случае остается без изменений. По такому принципу определяются пакеты, в которых к изображениям применяется замещение фрагмента другим изображением из этого пакета. Следовательно, для каждой эпохи обучения формируются разные данные, что не позволяет СНС запоминать тренировочные примеры.

Использование циклического сдвига позволяет получить изображение человека, на котором отдельные части тела могут оказаться не на своих местах, например, голова будет расположена в нижней части изображения, ступни выше плеч, а правая рука – слева. Такой подход направлен на повышение устойчивости СНС к местоположению отличительных особенностей фигуры в кадре.

Цвет одежды является важной отличительной особенностью при повторной идентификации человека. Однако цветопередача различных видеокамер и освещенность сцены видеонаблюдения могут значительно искажать цветовые характеристики одного и то же объекта. Исключение цветности направлено на

повышение устойчивости к уровням освещения за счет увеличения влияния других признаков изображения человека.

Замещение фрагмента изображения уменьшенной копией другого из пакета позволяет СНС одновременно анализировать признаки двух людей, способствует повышению устойчивости СНС к окклюзиям и различному разрешению изображений. Данное преобразование применяется к пакету размером  $b$ , в котором изображения  $I_i^{(Y)}$  имеют метки (идентификаторы,  $ID$ )  $Y$ , где  $i \in \{1, 2, \dots, b\}$  – индекс изображения в пакете. На каждом изображении  $I_i^{(Y)}$  фрагмент замещается на уменьшенную копию  $I_{mini_j}^{(Y_{mini})}$   $j$ -го изображения  $I_j^{(Y)}$ , где  $Y_{mini}$  – идентификатор уменьшенного изображения. Определение индекса  $j$  для  $I_{mini_j}^{(Y_{mini})}$ , замещающего фрагмент на  $i$ -том изображении, осуществляется с помощью ГПСЧ, который формирует массив целых чисел размером  $b$ , заполненный значениями от 1 до  $b$  без повторений. Полученное в результате преобразований итоговое изображение  $\tilde{I}^{(\tilde{Y})}$  имеет идентификатор  $\tilde{Y}$  и описывается следующим образом:

$$\tilde{I}^{(\tilde{Y})} = M \odot I_i^{(Y)} + \Phi\left(I_{mini_j}^{(Y_{mini})}\right), \quad (1)$$

где  $M \in \{0, 1\}^{H \times W}$  – бинарная маска, определяющая местоположение фрагмента изображения, который будет исключён и заменен на уменьшенное изображение из пакета;  $H$  и  $W$  – высота и ширина изображения;  $\odot$  – поэлементное умножение;  $\Phi(\cdot)$  – операция заполнения фрагмента изображения; метка  $\tilde{Y}$  для аугментированного изображения вычисляется как:

$$\tilde{Y} = \lambda_{aug} Y + (1 - \lambda_{aug}) Y_{mini}, \quad (2)$$

$$\lambda_{aug} = 1 - \frac{h_{mini} \cdot w_{mini}}{H \cdot W}. \quad (3)$$

где  $h_{mini}$  и  $w_{mini}$  – высота и ширина уменьшенного изображения, которое замещает удаленный фрагмент, соответственно  $h_{mini} = h_{erasing}$ ,  $w_{mini} = w_{erasing}$ , где  $h_{erasing}$  – высота и  $w_{erasing}$  – ширина удаленного фрагмента. Значение  $h_{erasing}$  определяется с помощью ГПСЧ в диапазоне от  $0,25 \cdot H$  до  $0,5 \cdot H$ . Ширина удаляемого фрагмента устанавливается  $w_{erasing} = \left\lfloor \frac{h_{erasing}}{\mu} \right\rfloor$ , где  $\mu = \left\lfloor \frac{H}{W} \right\rfloor$  – соотношение сторон исходного изображения;  $\lfloor \cdot \rfloor$  – операция округления до ближайшего меньшего целого. Для  $\Phi(\cdot)$  формируется маска  $\tilde{M} = 1 - M$ , заполненная единицами, что позволяет заместить удаленную область для  $I_i^{(Y)}$  уменьшенной копией  $I_{mini_j}^{(Y_{mini})}$ . Координаты левого нижнего угла фрагмента  $(x_{erasing}; y_{erasing})$  определяются с помощью ГПСЧ.

В результате аугментации, применяемой к пакетам изображений, обучающая выборка включает пакеты  $R$  исходных и  $\tilde{R}$  преобразованных изображений:

$$S^{train} = S_{aug\_batch}^{train} = \left\{ \left\{ R_1, R_2, \dots, R_{i\_batch} \right\}, \left\{ \tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_{i\_aug\_batch} \right\} \right\}, \quad (4)$$

где  $i\_batch$  и  $i\_aug\_batch$  – количество пакетов с исходными и преобразованными данными соответственно.

Применение аугментации к изображениям и к пакетам определяет обучающую выборку как:

$$S^{train} = S_{aug\_img\_batch}^{train} = \left\{ \left\{ R'_1, R'_2, \dots, R'_{i\_img\_aug} \right\}, \left\{ \tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_{i\_img\_batch} \right\} \right\}, \quad (5)$$

где  $R'_{i\_img\_aug}$  – пакеты, в котором аугментация применялась к отдельным изображениям;  $i\_img\_aug$  – количество таких пакетов;  $i\_img\_batch$  – число пакетов, в которых преобразования выполнялись для изображений и для пакетов. На рисунке 2 представлен пример применения разработанного метода аугментации. Изображения 8 и 12 преобразованы к полутоновым, к изображениям 6 – 8, 11, 13, 14 и 16 применен циклический сдвиг. На каждом изображении в пакете фрагмент замещается на уменьшенные копии других изображений.



**Рисунок 2 – Примеры изображений людей из набора данных PolReID1077 с применением предложенного метода аугментации**

Разработана технология двухэтапного обучения СНС, которая для предварительного этапа предполагает использовать наборы изображений, расширенные за счет аугментации, а для точной настройки весовых коэффициентов применять обучение на исходных изображениях и с меньшей скоростью обучения. Применение таких двух этапов способствует снижению переобучения СНС и уменьшению значения функции потерь.

Этап 1. Предварительное обучение.

1. Формирование пакетов  $R$  исходных и  $\tilde{R}$  аугментированных данных размером  $b$ .

2. Вычисление функции потерь  $L_R$  для каждого пакета, подаваемого на вход СНС:

2.1. Если изображения содержат замещенные фрагменты, то  $L_R = L_{\tilde{R}}$ :

$$L_{\tilde{R}} = E\left(f\left(\tilde{I}_i^{(\tilde{Y})}\right)\right) \cdot \lambda_{aug} + E\left(f\left(I_{mini_j}^{(Y_{mini})}\right)\right) \cdot (1 - \lambda_{aug}), \quad (6)$$

где  $E(\cdot)$  – кросс-энтропийная функция потерь;  $f\left(\tilde{I}_i^{(\tilde{Y})}\right)$  – вектор признаков изображения  $\tilde{I}_i^{(\tilde{Y})}$ , фрагмент которого замещен изображением  $I_{mini_j}^{(Y_{mini})}$  с меткой  $Y_{mini}$ ;  $f\left(I_{mini_j}^{(Y_{mini})}\right)$  – вектор признаков уменьшенного изображения;  $j$  – порядковый номер изображения, уменьшенная копия которого добавлена на  $\tilde{I}_i^{(\tilde{Y})}$ ;  $\lambda_{aug}$  определяется по выражению (3).

2.2. Если при формировании пакета аугментация применялась только к отдельным изображениям, то функция потерь вычисляется как:

$$L_R = E\left(f\left(I_i^{(Y)}\right)\right), \quad (7)$$

где  $f\left(I_i^{(Y)}\right)$  – дескриптор изображения  $I_i^{(Y)}$ , для которого замещение фрагмента не осуществлялось.

3. Вычисление средней величины ошибки с учетом накопления потерь для обрабатываемых пакетов в течение эпохи:

$$L_{epoch} = \frac{\sum(L_R + L_{\tilde{R}}) \cdot b}{S_{train}}, \quad (8)$$

где  $S_{train}$  – размер обучающей выборки.

4. Корректировка весовых коэффициентов. Для скрытых слоев СНС  $w_{hidden\_epoch\_num}$ , где  $epoch\_num$  – номер текущей эпохи, корректируются с меньшей скоростью  $\eta$ , чем весовые коэффициенты классификационного слоя  $w_{classifier\_epoch\_num}$ :

$$w_{classifier\_epoch\_num} = w_{classifier\_epoch\_num-1} - \eta \cdot \frac{\partial L_{epoch}}{\partial w_{classifier\_epoch\_num-1}} \quad (9)$$

$$w_{hidden\_epoch\_num} = w_{hidden\_epoch\_num-1} - 0,1 \cdot \eta \cdot \frac{\partial L_{epoch}}{\partial w_{hidden\_epoch\_num-1}} \quad (10)$$

5. Уменьшение скорости обучения  $\eta$  после 40-й эпохи.

6. Предварительный этап обучения на аугментированных данных в течение 45 эпох.

Этап 2. Точная настройка.

1. Формирование пакетов  $R$  изображений без аугментации.

2. Вычисление функции потерь по выражению (7) для пакетов  $R$ .

3. Вычисление значения функции потерь для эпохи:

$$L_{epoch} = \frac{\sum L_R \cdot b}{S_{train}}. \quad (11)$$

4. Корректировка весовых коэффициентов с использованием формул (9) и (10).

Проведённые исследования показали, что применение для тренировки СНС ResNet-50 и DenseNet-121 разработанного метода аугментации и технологии двухэтапного обучения позволяет увеличить точность реидентификации человека до 21 % по метрике Rank1, по метрике mAP до 31 %, по mINP до 59 % на наборах данных Market-1501, DukeMTMC-ReID, MSMT17 и PolReID1077.

Применение разработанных метода аугментации и технологии обучения СНС ResNet-50 и DenseNet-121 на выборке из 109 772 изображения для 6 381 человека, полученной в результате объединения созданного набора изображений PolReID1077 с существующими CUHK02, CUHK03, Market-1501, DukeMTMC-ReID, MSMT17, позволило повысить точность Rank1 до 48 %, mAP до 92 %, mINP до 292 % по сравнению с базовым-алгоритмом.

В *четвертой главе* представлен алгоритм повторной идентификации человека, который предполагает описание изображения каждого человека составным нейросетевым дескриптором. Для реализации алгоритма с помощью детектора осуществляется обнаружение изображения человека на кадре видеопоследовательности и выделение его ограничительной прямоугольной рамкой. Данная область кадра нормализуется до размеров  $256 \times 128$ . Если полученная рамка очерчивает изображения двух человек, то объектом интереса, который в дальнейшем анализируется, считается тот, площадь ограничительного прямоугольника которого больше. Затем для изображения человека  $I^{(Y)}$  формируется составной нейросетевой дескриптор, который включает глобальный вектор  $f_{global}(I^{(Y)})$  всей выделенной области и локальные вектора  $f_{p_1}(I_{p_1}^{(Y)})$ ,  $f_{p_2}(I_{p_2}^{(Y)})$ ,  $f_{p_3}(I_{p_3}^{(Y)})$  для верхней  $p_1$ , средней  $p_2$  и нижней  $p_3$  ее частей соответственно, полученные с использованием СНС. Если человек на изображении перекрывается другими людьми или объектами, то соответствующий фрагмент относится к невидимому. Такие фрагменты не используются в формировании локальных признаков.

Выделение трех частей изображения человека  $p_1$ ,  $p_2$ ,  $p_3$  основано на обнаружении и анализе семнадцати ключевых точек ( $kp_l$ , где  $l$  – порядковый номер точки,  $l \in \{1, 2, \dots, 17\}$  на изображении тела человека, включая: нос ( $kp_1$ ), глаза ( $kp_2$ ,  $kp_3$ ), уши ( $kp_4$ ,  $kp_5$ ), плечи ( $kp_6$ ,  $kp_7$ ), локти ( $kp_8$ ,  $kp_9$ ), запястья ( $kp_{10}$ ,  $kp_{11}$ ), бедра ( $kp_{12}$ ,  $kp_{13}$ ), колени ( $kp_{14}$ ,  $kp_{15}$ ) и лодыжки ( $kp_{16}$ ,  $kp_{17}$ ). Точка считается обнаруженной, если  $T_{kp_l} \geq 0,5$ , где  $T_{kp_l}$  – степень уверенности нейронной сети в достоверности ее предсказания. На основе координат ключевых точек и результатов их обнаружения вычисляются координаты фрагментов, причем фрагмент  $p_2$

включается как в  $p_1$ , так и в  $p_3$ , т.е. присутствует перекрытие областей.

Для описания локальных фрагментов на изображении применяются маски  $M_{p_1}$ ,  $M_{p_2}$  и  $M_{p_3}$ , состоящие из нулей и единиц. Единичной областью в данных масках характеризуются видимые локальные фрагменты, соответственно, нулевые значения соответствуют скрытым частям. Таким образом составной вектор признаков изображения человека представляется как:  $f_{gen}(I^{(Y)}) = \{f_{global}(I^{(Y)}), f_{p_1}(I_{p_1}^{(Y)}), f_{p_2}(I_{p_2}^{(Y)}), f_{p_3}(I_{p_3}^{(Y)})\}$ , где  $I_{p_1}^{(Y)} = I^{(Y)} \odot M_{p_1}$ ;  $I_{p_2}^{(Y)} = I^{(Y)} \odot M_{p_2}$ ;  $I_{p_3}^{(Y)} = I^{(Y)} \odot M_{p_3}$ ;  $\odot$  – поэлементное умножение.

### Замещение признаков невидимых областей

При первом проходе изображений по сети формируется составной дескриптор для всех изображений галереи. Если одна или несколько локальных областей считаются другими объектами, то соответствующий компонент составного дескриптора будет недействительным и равным нулю. Для изображений, вектор признаков которых имеет нулевые составляющие, выполняется ранжирование таблицы признаков по метрике косинусного подобия для каждого видимого локального фрагмента, для которого определяются  $k$ -ближайших соседей,  $k=8$ , из них определяются  $k_1$ -лучших,  $k_1=2$ . Описание недостающего фрагмента изображения осуществляется усреднением признаков соответствующих ему  $k_1$ -ближайших соседей. Затем таблица дескрипторов обновляется, а дальнейшее ранжирование выполняется в соответствии с запросом с использованием метрики косинусного подобия (рисунок 3).

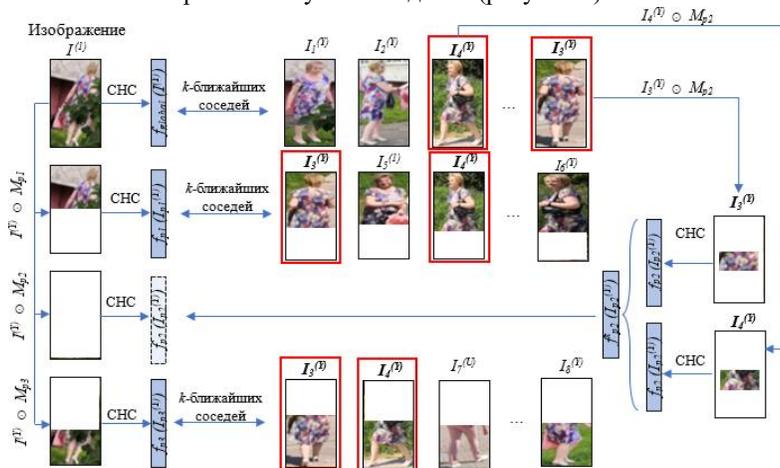


Рисунок 3 – Схема алгоритма замещения признаков для скрытых областей на основе наиболее схожих изображений, определяемых по видимым частям фигуры человека

Разработанный алгоритм предполагает, что при обучении на вход СНС

передаются пакеты  $R_p$ , которые содержат изображения и координаты масок  $M_{p1}$ ,  $M_{p2}$  и  $M_{p3}$  для выделения локальных фрагментов.

Для каждого пакета  $R_p$ , при формировании которого не применялась аугментация, вычисляется значение функции потерь  $L_{R_p}$  для:

$$L_{R_p} = \lambda_{global} \cdot S_{global} \cdot E\left(f_{global}\left(I_i^{(Y)}\right)\right) + \lambda_{p1} \cdot S_{p1_i} \cdot E\left(f_{p1}\left(I_{p1_i}^{(Y)}\right)\right) + \\ + \lambda_{p2} \cdot S_{p2_i} \cdot E\left(f_{p2}\left(I_{p2_i}^{(Y)}\right)\right) + \lambda_{p3} \cdot S_{p3_i} \cdot E\left(f_{p3}\left(I_{p3_i}^{(Y)}\right)\right), \quad (12)$$

где  $\lambda_{global}$ ,  $\lambda_{p1}$ ,  $\lambda_{p2}$ ,  $\lambda_{p3}$  – коэффициенты, определяющие степень влияния на функцию потерь каждого из фрагментов;  $S_{global}$  – площадь изображения;  $S_{p1_i}$ ,  $S_{p2_i}$ ,  $S_{p3_i}$  – площади локальных фрагментов;  $i$  – порядковый номер изображения в пакете;  $E(.)$  – кросс-энтропийная функция потерь.

Для случаев, в которых при формировании пакета  $R_{aug_p}$  применялась предложенная аугментация и изображение имеет замещенные фрагменты  $\tilde{I}_i^{(\tilde{Y})}$ , функция потерь определяется как:

$$L_{R_{aug_p}} = \lambda_{global} \cdot S_{global} \cdot \left( E\left(f_{global}\left(I_i^{(Y)}\right)\right) + E\left(f\left(\tilde{I}_i^{(\tilde{Y})}\right) \cdot (1 - \lambda_{aug})\right) \right) + \\ + \lambda_{p1} \cdot S_{p1_i} \cdot E\left(f_{p1}\left(I_{p1_i}^{(Y)}\right)\right) + \\ + \lambda_{p2} \cdot S_{p2_i} \cdot E\left(f_{p2}\left(I_{p2_i}^{(Y)}\right)\right) + \lambda_{p3} \cdot S_{p3_i} \cdot E\left(f_{p3}\left(I_{p3_i}^{(Y)}\right)\right), \quad (13)$$

где  $f\left(\tilde{I}_i^{(\tilde{Y})}\right)$  – вектор признаков изображения с фрагментом, замещенным при аугментации;  $\lambda_{aug}$  – определяется по выражению (3).

После прохождения всех изображений обучающей выборки вычисляется значение функции потерь для эпохи по аналогии с (8) и корректируются весовые коэффициенты по формулам (9) и (10).

Применение предлагаемого алгоритма позволило увеличить точность реидентификации относительно базового алгоритма при использовании СНС ResNet-50 и DenseNet-121 по метрике Rank1 до 10 %, по метрике mAP до 14%, по mINP до 49 %. Проведенные исследования показали, что применение для обучения сформированного объединенного набора изображений на основе созданного PolReID1077 и существующих, разработанных метода аугментации, технологии двухэтапного обучения и алгоритма повторной идентификации человека позволило повысить точность повторной идентификации по метрике Rank1 до 54 %, по метрике mAP до 116 %, по mINP до 400 % относительно базового алгоритма реидентификации.

В **пятой главе** выполнена программная реализация системы повторной идентификации человека на основе разработанного алгоритма реидентификации. Архитектура программной реализации данной системы состоит из трех модулей:

1. Модуль обнаружения людей, в котором осуществляется разделение

входных видео на кадры, обнаружение изображений людей и ключевых точек на них с использованием СНС YOLOv7pose. В качестве входных данных могут использоваться видео, полученные с IP-, веб-камер или хранящиеся на диске компьютера.

2. Модуль повторной идентификации человека с использованием СНС DenseNet-121. При тренировке данной СНС применялись разработанные в рамках диссертационного исследования технология двухэтапного обучения и метод аугментации в предлагаемом алгоритме реидентификации с использованием составного нейросетевого дескриптора и замещением признаков невидимых областей фигуры человека. Модуль включает реидентификацию и формирование набора изображений людей (Image), разделенных по отдельным папкам согласно присвоенным идентификаторам. Реализован с использованием языка программирования Python, фреймворков и библиотек PyTorch, Numpy, OpenCV.

3. Модуль пользовательского интерфейса, для реализации которого применялись компоненты кроссплатформенной библиотеки SFML (Simple and Fast Multimedia Library): System – для управления временем и потоками; Window – для управления окнами и потоками; Graphics – для отображения графических примитивов и изображений.

На каждом кадре  $F^K$ , где  $K$  - номер видеокamеры, с помощью детектора выполняется обнаружение всех изображений людей  $I^{(j)}$  с присвоением им идентификатора  $Y$ . Для каждого нового изображения вычисляется расстояние между признаками со всеми обнаруженными ранее. Если это значение больше установленного порога  $Thr$ , то человеку на изображении присваивается новый идентификатор, в противном случае – идентификатор человека, с которым расстояние между признаками было минимально. Изображение сохраняется в папку на жестком диске с соответствующим номером.

Библиотека, используемая для захвата видеопотока, определяется в зависимости от источника видео. При обработке видео из файла для извлечения кадров используется `moviepy` для Python. Проверка наличия подключенных веб-камер и захват видеопотока с них выполняется с помощью `camerata` для Python. Кадры с IP-камер извлекаются с применением OpenCV.

На основном экране интерфейса разработанного программного обеспечения отображается видео с одной из камер видеонаблюдения. В нижней части экрана представлены все видео, доступные системе. Обнаруженные изображения людей сортируются по группам согласно присвоенным им идентификаторам. Группы изображений людей располагаются в правой половине основного экрана. Для реидентификации по запросу от пользователя необходимо выбрать интересующего человека путем нажатия левой кнопки мыши на соответствующую ограничительную рамку человека на кадре или выбрать соответствующую группу в правой части экрана для отображения в отдельном окне. Наведение курсора на интересующее изображение человека позволяет получить информацию о нем: идентификатор, место сохранения изображения, источник видео (номер камеры), дату и время получения кадра.

## ЗАКЛЮЧЕНИЕ

### Основные научные результаты диссертации

1. На основе анализа методов и алгоритмов повторной идентификации человека предложена классификация систем реидентификации: по взаимодействию с внешней средой разделяются на открытые и закрытые; по типу анализируемых данных могут быть статическими или динамическими; по количеству запросов делятся на одиночные и множественные; по типу запросов подразделяются на однородные и неоднородные; по времени работы делятся на краткосрочные и долгосрочные системы [3–А, 5–А, 11–А].

2. Исследовано влияние функций активации ReLU, Leaky-ReLU, PReLU, RReLU, ELU, SELU, GELU, Swish, Mish в СНС ResNet-50, DenseNet-121, DarkNet-53 на точность повторной идентификации человека. Анализ результатов показал, что для данной задачи наиболее перспективными ФА являются ReLU и GeLU, при этом минимальная скорость обучения достигается при использовании ReLU [2–А, 13–А]. Определены наиболее эффективные значения гиперпараметров при обучении СНС ResNet-50 для реидентификации людей: размер пакета  $b = 16$ , скорость обучения  $\eta = 0,07$ , количество эпох обучения равно 80 [6–А, 15–А]. Установлено, что применение повторного ранжирования признаков позволяет увеличить точность реидентификации человека на наборах данных Market-1501, DukeMTMC-ReID, PolReID с использованием СНС ResNet-50 и DenseNet-121 до 9% по метрике Rank1, на 10% – 25% по метрике mAP и на 43% – 112% по метрике mINP [4–А].

3. Сформирован новый набор данных PolReID1077 для обучения и тестирования СНС при повторной идентификации человека, отличающийся от существующих использованием видеоданных, полученных во все времена года при внешнем и внутреннем видеонаблюдении в большом количестве различных мест съемки с применением видеокамер с отличающимися характеристиками. Всего набор данных содержит 45 971 изображение 1077 различных людей. Для тренировки СНС используется 26 332 изображения для 604 человек, для тестирования – 18 711 для 403 человек. В тестовую выборку в качестве дистракторов не имеющих запросов, включены 928 изображений 70 людей, полученных только с одной видеокамеры [7–А, 14–А, 16–А, 17–А].

4. Разработан метод аугментации изображений для повторной идентификации человека, включающий циклический сдвиг изображения на различное число пикселей по вертикали и горизонтали, исключение цветности и замещение фрагментов изображений уменьшенными копиями разных размеров других изображений из пакета, что повышает устойчивость СНС к ошибкам детектора, изменению освещенности, уменьшению разрешения изображения, местоположению ключевых отличительных признаков [6–А, 8–А, 18–А, 20–А].

5. Предложена технология двухэтапного обучения СНС, которая основана на концепции переноса обучения и отличается от классического подхода, предполагающего использование для предварительного обучения больших наборов

данных, применением на данном этапе обучающих выборок, расширенных за счет аугментации. Для точной настройки весовых коэффициентов применяются изображения из исходного набора. Разработанная технология позволяет увеличить точность повторной идентификации человека при использовании СНС ResNet-50 и DenseNet-121 до 21 % по метрике Rank1, по метрике mAP до 31 %, по mINP до 59 % [6–А, 7–А, 8–А, 12–А, 20–А].

6. Разработан алгоритм повторной идентификации человека, в котором для повышения точности реидентификации используется составной вектор признаков, включающий глобальные и локальные отличительные особенности фигуры человека с учетом их видимости. Вектор признаков для скрытой области строится на основе усреднения значений признаков фрагментов изображений, являющихся ближайшими для большинства видимых соответствующих локальных областей, и он включается в составной дескриптор. Применение предлагаемого алгоритма позволило увеличить точность реидентификации относительно базового алгоритма при использовании СНС ResNet-50 и DenseNet-121 по метрике Rank1 до 10 %, по метрике mAP до 14 %, по mINP до 49 %. В совокупности применение для обучения сформированного большого объединенного набора изображений на основе созданного PolReID1077 и существующих, разработанных метода аугментации, технологии двухэтапного обучения и алгоритма позволило повысить точность повторной идентификации по метрике Rank1 до 54 %, по метрике mAP до 116 %, по mINP до 400 % относительно базового алгоритма реидентификации [9–А, 21–А, 22–А].

7. Разработано программное обеспечение для повторной идентификации людей по данным систем видеонаблюдения на основе предложенных подходов [1–А, 6–А, 10–А]. Программный продукт позволяет отображать результат реидентификации человека на мониторе с указанием даты и места его обнаружения, определяемого по номеру камеры, каждое изображение обнаруженного человека, формирует набор данных с идентифицированными людьми [19–А].

### **Рекомендации по практическому использованию результатов**

Полученные результаты использованы в ООО «Ай Ти Ви групп», которое занимается разработкой и внедрением интеллектуальных систем видеоналитики, а также применяются в учебном процессе на лабораторных и лекционных занятиях для студентов и магистрантов Полоцкого государственного университета имени Евфросинии Полоцкой.

Разработанные метод аугментации и технология двухэтапного обучения СНС могут использоваться при тренировке нейронных сетей в других областях компьютерного зрения. Предложенный алгоритм повторной идентификации человека может использоваться в многокамерных системах внутреннего или наружного видеонаблюдения.

## СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ ПО ТЕМЕ ДИССЕРТАЦИИ

*Статьи в научных изданиях в соответствии с п. 19*

*Положения о присуждении ученых степеней и присвоении ученых званий*

1–А. Повышение эффективности обнаружения объектов небольших размеров на 8 К изображениях при использовании сверточных нейронных сетей / Р. П. Богуш, С. В. Абламейко, С. А. Игнатьева, Е. Р. Адамовский // Известия ГГУ им. Ф. Скорины. – 2021. – № 3. – С. 85–91.

2–А. Choice of activation function in convolution neural network for person re-identification in video surveillance systems / H. Chen, S. Ihnatsyeva, R. Bohush, S. Ablameyko // Programming and computer software. – 2022. – № 5. – P. 312–321.

3–А. Игнатьева, С. А. Принципы организации и анализ подходов к повышению точности повторной идентификации людей в распределенных системах видеонаблюдения / С. А. Игнатьева // Вестник Полоцкого государственного университета. Серия С, Фундаментальные науки. – 2022. – № 4. – С. 13–25.

4–А. Bohush, R. Person re-identification accuracy improvement by training a CNN with the new large joint dataset and re-rank / R. Bohush, S. Ihnatsyeva, S. Ablameyko // Machine Graphics and Vision. – 2022. – № 31 (1/4). – P. 93–109.

5–А. Person Re-identification in Video Surveillance Systems Using Deep Learning: Analysis of the Existing Methods / H. Chen, S. A. Ihnatsyeva, R. P. Bohush, S. V. Ablameyko // Automation and Remote Control. – 2023. – Vol. 84., № 5. – P. 497–528.

6–А. Игнатьева, С. А. Увеличение точности реидентификации людей на основе двухэтапного обучения сверточных нейронных сетей и аугментации / С. А. Игнатьева, Р. П. Богуш // Информатика. – 2023. – № 1. – С. 40–54.

7–А. Игнатьева, С. А. Формирование обучающей выборки для сверточных нейронных сетей при реидентификации людей по видеоданным / С. А. Игнатьева, Р. П. Богуш // Доклады БГУИР. – 2023. – № 3. – С. 87–95.

8–А. Data augmentation and fine tuning of convolution neural network during training for person re-identification in video surveillance systems / S. Ye, R. Bohush, S. Ihnatsyeva, S. V. Ablameyko // Optical memory and Neural Network. – 2023. – № 4. – P. 233–246.

9–А. Игнатьева, С. А. Алгоритм реидентификации людей по изображениям систем видеонаблюдения с использованием нейросетевого составного дескриптора / С. А. Игнатьева, Р. П. Богуш // Системный анализ и прикладная информатика. – 2024. – № 1. – С. 12–17.

10–А. Object Detection Algorithm for High Resolution Images Based on Convolutional Neural Network and Multiscale Processing / R. Bohush, S. Ablameyko, S. Ihnatsyeva, Ya. Adamovskiy // Computer Modeling and Intelligent Systems. CMIS–2021 : proc. of the IV Intern. workshop, Zaporizhzhia, Ukraine, 27 April, 2021 / National University “Zaporizhzhia Polytechnic” ; ed. S. Subbotin [et al.]. – Zaporizhzhia, 2021. – P. 135–144.

11–А. Ihnatsyeva, S. Analysis person re-identification methods / S. Ihnatsyeva, R. Bohush // European and national dimension in research = Европейский и национальный контексты в научных исследованиях : electronic collected materials of XIII Junior Researchers’ Conference, Novopolotsk, May 17–21, 2021 / Polotsk State University ; publishing board: Y. Holubeu [et al.] : in 3 parts. - Novopolotsk : PSU, 2021. – Part 3 : Technology. – 2021. – P. 204–208. – URL: <https://elib.psu.by/handle/123456789/27658> (date of access: 13.10.2024).

12–А. Ihnatsyeva, S. Joint Dataset for CNN-based Person Re-identification / S. Ihnatsyeva, R. Bohush, S. Ablameyko // Pattern Recognition and Information Processing (PRIP'2021) : proc. of the 15th Intern. conference, Minsk, 21–24 Sept. 2021 / ed.: A. Tuzikov [et al.]. – Minsk : UИP NASB. – 2021. – P. 33–37.

13–А. Игнатъева, С. А. Сравнительный анализ функций активации и их влияние на точность ре-идентификации людей с использованием сверточных нейронных сетей / С. А. Игнатъева // Современные проблемы математики и вычислительной техники : материалы XII Республ. науч. конф. молодых ученых и студентов, 18–19 нояб. 2021 г. / Брестский государственный технический университет ; редкол.: В. А. Головка [и др.]. – Брест : БрГТУ, 2021. – С. 40–48.

14–А. Игнатъева, С. А. Состав и особенности нового набора данных для решения задачи повторной идентификации человека в системах видеонаблюдения / С. А. Игнатъева, Р. П. Богуш // Информатика: проблемы, методы, технологии : материалы XXII Междунар. науч.-практ. конф., Воронеж, 10–12 февр. 2022 г., редкол.: Д. Н. Борисова [и др.]. – Воронеж : ВГУ, 2022. – С. 592–599.

15–А. Игнатъева, С. А. Влияние выбора гиперпараметров при обучении сверточных нейронных сетей на точность повторной идентификации людей в системах видеонаблюдения / С. А. Игнатъева // Информационно-коммуникационные технологии: достижения, проблемы, инновации (ИКТ–2022) : материалы II Междунар. науч.-практ. конф., Новополоцк, 30–31 марта 2022 г. / Полоцкий государственный университет имени Евфросинии Полоцкой ; редкол.: О. А. Романов [и др.]. – Новополоцк : ПГУ, 2022. – С. 163–167.

16–А. Estimation CNN-Based Person Re-Identification Accuracy in Video Using Different Datasets / Ye. Shiping, S. Ihnatsyeva, R. Bohush, Ch. Chen,

S. Ablameyko // *Advances in Transdisciplinary Engineering*. – 13–14 August 2022. – Vol. 30: Applied Mathematics, Modeling and Computer Simulation. – P. 978–985. – DOI: 10.3233/ATDE221122.

17–А. Богуш, Р. П. Сопровождение и повторная идентификация людей в интеллектуальных системах видеонаблюдения с применением сверточных нейронных сетей / Р. П. Богуш, С. А. Игнатьева, С. В. Абламейко // Первая выставка-форум IT-академграда «Искусственный интеллект в Беларуси»: сборник докладов, Минск, 13–14 окт. 2022 г. – Минск: ОИПИ НАН Беларуси, 2022. – С. 46–53.

18–А. Игнатьева, С. А. Аугментация данных для регуляризации в нейросетевых алгоритмах повторной идентификации людей по видеоданным / С. А. Игнатьева, Р. П. Богуш, С. В. Абламейко // *Информационные системы и технологии = Information Systems and Technologies*: материалы междунар. науч. конгресса по информатике, Минск, 27–28 окт. 2022 г. / Белорус. гос. ун-т; редкол.: С. В. Абламейко [и др.]: в 3 ч. – Минск: БГУ, 2022. – Ч. 2. – С. 132–137.

19–А. Экспериментальный прототип открытой системы повторной идентификации людей при многокамерном видеонаблюдении / С. А. Игнатьева, Н. А. Томашевич, А. А. Голубенок, Р. П. Богуш // *Технологии передачи и обработки информации = Technologies of information transmission and processing*: материалы Междунар. науч.-технич. семинара, Минск, март – апрель 2023 г. – Минск: БГУИР. – 2023. – С. 117–122.

20–А. Игнатьева, С. А. Двухэтапное обучение СНС при использовании аугментации изображений для повторной идентификации человека / С. А. Игнатьева, Р. П. Богуш // *Информационные технологии и нанотехнологии (ИТНТ-2023)*: сборник трудов по материалам IX Междунар. конф. и молодежной школы, г. Самара, 17–23 апр. 2023 г. / ред.: В. В. Сергеев: в 6 т. – Самара: Самарский национальный исследовательский университет имени академика С.П. Королева, 2023. – Т. 3. – С. 30792.

21–А. Игнатьева, С. А. Реидентификация людей по данным систем видеонаблюдения с использованием машинного обучения / С. А. Игнатьева, Р. П. Богуш // Вторая выставка-форум IT-академграда «Искусственный интеллект Беларуси»: сборник докладов, Минск, 12–13 окт. 2023 г. – Минск: ОИПИ НАН Беларуси, 2023. – С. 112–118.

22–А. Ihnatsyeva S. Person re-identification using compound descriptor and invisible region replacement / S. Ihnatsyeva, R. Bohush // *Pattern Recognition and Information Processing (PRIP'2023) = Распознавание образов и обработка информации (2023)*: proc. of the 16th Intern. conf., Minsk, 17–19 Oct. 2023 / ed.: A. Nedzved [et al.]. – Minsk: UIIP NASB, 2023. – P. 193–196.

## РЕЗЮМЕ

Игнатъева Светлана Александровна

### **Повторная идентификация человека на изображениях систем видеонаблюдения с использованием сверточных нейронных сетей**

**Ключевые слова:** системы видеонаблюдения, машинное обучение, аугментация изображений, алгоритм повторной идентификации, компьютерное зрение, интеллектуальная система.

**Целью диссертационной работы** является увеличение точности повторной идентификации человека в распределенной системе видеонаблюдения.

**Методы исследования и используемая аппаратура:** в работе использовались методы цифровой обработки изображений, методы машинного обучения и компьютерного зрения; персональный компьютер с графическим процессором Nvidia и поддержкой технологии CUDA.

**Полученные результаты и их новизна:** предложена классификация систем повторной идентификации человека; установлено влияние функции активации на точность повторной идентификации человека; сформировано репрезентативное множество изображений людей, полученных при различных условиях видеонаблюдения для обучения и тестирования алгоритмов повторной идентификации; предложен метод аугментации изображений людей для реидентификации, включающий циклический сдвиг пикселей изображения, исключение цветности и замещение фрагмента уменьшенной копией другого изображения из пакета; технология двухэтапного обучения СНС; алгоритм повторной идентификации человека, основанный на построении составного нейросетевого дескриптора и замещения признаков скрытых областей фигуры человека. Разработано программное обеспечение для повторной идентификации людей по данным систем видеонаблюдения на основе предложенных подходов.

**Рекомендации по использованию:** результаты, полученные в диссертационном исследовании использованы в ООО «Ай Ти Ви групп», а также применяются в учебном процессе на лабораторных и лекционных занятиях для студентов и магистрантов Полоцкого государственного университета имени Евфросинии Полоцкой. Предложенный метод аугментации и технология двухэтапного обучения СНС могут использоваться при тренировке нейронных сетей в других областях компьютерного зрения. Предложенный алгоритм реидентификации может использоваться в многокамерных системах внутреннего или наружного видеонаблюдения.

**Область применения:** интеллектуальные системы видеонаблюдения, которые могут использоваться для поиска пропавших людей или правонарушителей, сбора статистической информации о посещении людьми торговых центров, учреждений здравоохранения, образования и других социально-значимых объектов.

## РЭЗІЮМЭ

Ігнацьева Святлана Аляксандраўна

### Паўторная ідэнтыфікацыя чалавека на выявах сістэм відэаназірання з выкарыстаннем згортачных нейронавых сетак

**Ключавыя словы:** сістэмы відэаназірання, машыннае навучанне, аугментацыя малюнкаў, алгарытм паўторнай ідэнтыфікацыі, камп'ютэрны зрок, інтэлектуальная сістэма.

**Мэта працы:** павышэнне дакладнасці паўторнай ідэнтыфікацыі чалавека у размеркаванай сістэме відэаназірання.

**Метады даследавання і выкарыстаная апаратура:** у рабоце выкарыстоўваліся метады лічбавай апрацоўкі відарысаў, метады машыннага навучання і камп'ютарнага зроку; персанальны камп'ютар з графічным працэсарам Nvidia і падтрымкай тэхналогіі CUDA.

**Атрыманыя вынікі і іх навізна:** прапанавана класіфікацыя сістэм паўторнай ідэнтыфікацыі чалавека; устаноўлены ўплыў функцыі актывацыі на дакладнасць паўторнай ідэнтыфікацыі чалавека; сфарміравана рэпрэзентатыўнае мноства выяваў людзей, атрыманых пры розных умовах відэаназірання для навучання і тэсціравання алгарытмаў паўторнай ідэнтыфікацыі; прапанаваны метады аугментацыі малюнкаў людзей для рэідэнтыфікацыі, які ўключае цыклічны зрух пікселяў выявы, выключэнне каляровасці і замяшчэнне фрагмента выявы паменшанай копіяй іншага малюнка з пакета; тэхналогія двухэтапнага навучання ЗНС; алгарытм рэідэнтыфікацыі людзей, заснаваны на пабудове састаўнога нейросетавога дэскрыптарна і замяшчэння прыкмет утоеных абласцей фігуры чалавека. Распрацавана праграма забеспячэнне для рэідэнтыфікацыі людзей па дадзеных сістэм відэаназірання на аснове прапанаваных падыходаў.

**Рэкамендацыі па выкарыстанні:** вынікі, атрыманыя ў дысертацыйным даследаванні, выкарыстаны ў ТАА «Ай Ці Ві груп», а таксама прымяняюцца ў навучальным працэсе на лабараторных і лекцыйных занятках для студэнтаў і магістрантаў Полацкага дзяржаўнага ўніверсітэта імя Ефрасінні Полацкай. Прапанаваны метады аугментацыі і тэхналогія двухэтапнага навучання ЗНС могуць выкарыстоўвацца пры трэніроўцы нейронавых сетак у іншых абласцях камп'ютэрнага зроку. Прапанаваны алгарытм паўторнай ідэнтыфікацыі можа выкарыстоўвацца ў шматкамерных сістэмах унутранага ці вонкавага відэаназірання.

**Вобласць ужывання:** інтэлектуальныя сістэмы відэаназірання, якія могуць выкарыстоўвацца для пошука зніклых людзей ці правапарушальнікаў, збора статыстычнай інфармацыі аб наведванні людзьмі гандлёвых цэнтраў, устаноў аховы здароўя, адукацыі і іншых сацыяльна-значных аб'ектаў.

## SUMMARY

Ihnatsyeva Sviatlana Aleksandrovna

### **Person re-identification in video surveillance images using convolutional neural networks**

**Keywords:** video surveillance systems, machine learning, image augmentation, reidentification algorithm, computer vision, intelligent system.

**The purpose of the dissertation work** is increasing the person re-identification accuracy in a distributed video surveillance system.

**Research methods and equipment used:** the work used digital image processing methods, machine learning and computer vision methods; a personal computer with an Nvidia GPU and support for CUDA technologies.

**Equipment used:** personal computer with main characteristics: Intel Core i5 3.11 GHz, 16 Gb RAM, Nvidia GeForce RTX-3060 6 Gb.

**The results obtained and their novelty:** a classification person re-identification systems is proposed; activation function the influence on the person re-identification accuracy was established; a representative person's images dataset obtained under various video surveillance conditions was generated for training and testing re-identification algorithms; method for augmenting images for person re-identification is proposed including image pixels cyclic shift, transform to grayscale and replacement of an image fragment with a reduced copy of another image from the batch; two-stage CNN training technology; algorithm for person re-identification, based on the construction of a compound neural network descriptor and feature replacement of hidden areas of the human figure. Software has been developed for person re-identification based on video surveillance systems data based on the proposed approaches.

**Usage recommendations:** The results obtained in the dissertation research were used at LLC "ITV Group", and are also used in the educational process in laboratory and lecture classes for students and undergraduates of Euphrosyne Polotskaya State University of Polotsk. The proposed augmentation method and two-stage CNN training technology can be used when training neural networks in others areas of computer vision. The proposed re-identification algorithm can be used in multi-camera indoor or outdoor video surveillance systems.

**Application area:** intelligent video surveillance systems that can be used for searching for missing people or offenders, collecting statistical information on person's visits to shopping centers, healthcare institutions, education and other socially significant facilities.



Подписано в печать 16.01.2025. Формат 60×84 <sup>1</sup>/<sub>16</sub>. Бумага офсетная.  
Ризография. Усл. печ. л. 1,39. Уч.-изд. л. 1,5.  
Тираж 60 экз. Заказ 4.

Отпечатано с оригинала-макета заказчика  
в республиканском унитарном предприятии  
«Издательский центр Белорусского государственного университета».  
Свидетельство о государственной регистрации издателя, изготовителя,  
распространителя печатных изданий № 2/63 от 19.03.2014.  
Ул. Красноармейская, 6, 220030, Минск.