

ТРАНСПОРТНЫЕ СИСТЕМЫ

2.9.1 – Транспортные и транспортно-технологические системы страны, ее регионов и городов, организация производства на транспорте;

2.9.4. – Управление процессами перевозок;

2.9.6 – Аэронавигация и эксплуатация авиационной техники;

2.9.8 – Интеллектуальные транспортные системы

УДК 658.71.08,519.87

DOI: 10.26467/2079-0619-2024-27-2-8-24

**Применение предварительно обученных нейронных сетей
для решения задачи обратного поиска рентгеновских изображений
запрещенных предметов и веществ**

А.К. Волков¹, Л.В. Миронова¹, С.Е. Потапова¹

¹Ульяновский институт гражданской авиации

имени Главного маршала авиации Б.П. Бугаева, г. Ульяновск, Россия

Аннотация: В работе рассматриваются вопросы применения предварительно обученных нейронных сетей для решения задачи обратного поиска рентгеновских изображений запрещенных предметов и веществ. Целью работы является проведение анализа и обоснование путей повышения эффективности систем распознавания рентгеновских изображений багажа и ручной клади пассажиров. Представлен анализ существующих отечественных и зарубежных работ в области распознавания рентгеновских изображений багажа и ручной клади пассажиров. Выявлено, что, несмотря на достигнутые результаты по разработке алгоритмов распознавания запрещенных предметов и веществ, они не в полной мере справляются с таким фактором сложности, как наложение предметов. Для решения данной проблемы в работе предлагается дополнительно анализировать те рентгеновские изображения, на которых уверенность в распознавании объектов невысокая. Данный этап включает в себя следующие шаги: сегментация изображений, извлечение признаков сегментированных элементов изображений; поиск схожих изображений по базе данных; принятие решения о классе сегментированных элементов изображений. В данной статье рассматриваются три последних шага. Проанализированы варианты реализации подходов к извлечению признаков из изображений, в частности основанные на применении сверточных автоэнкодеров и предварительно обученных нейронных сетей. Выбран подход, основанный на применении предварительно обученных нейронных сетей. В работе применяется нейронная сеть архитектуры ResNet-50, предварительно обученная на коллекции ImageNet. Для применения данной модели для извлечения векторов признаков изображений, был предварительно удален последний слой классификации. Все предыдущие слои модели кодируют изображение в вектор. ResNet-50 генерирует 2048-мерный вектор признаков изображений. Для понижения размерности векторов признаков изображений используется метод главных компонент. Решение о том, является ли сегментированный элемент изображения запрещенным предметом или веществом, рассматривается как задача обратного поиска с применением алгоритма K -ближайших соседей. Класс элемента рентгеновского изображения в данном случае – это класс, наиболее часто встречающийся среди k ближайших соседей. В целях апробации предложенного подхода был сформирован обучающий набор данных, включающий 4 635 изображений отдельных предметов и веществ, которые могут встречаться в багаже и ручной клади пассажиров. Представлен сравнительный анализ времени индексации и поиска изображений при различных алгоритмах и количестве признаков. Представлен сравнительный анализ точности модели. Сделан вывод, что наиболее приемлемым является алгоритм Brute force в сочетании с методом главных компонент.

Ключевые слова: досмотр, рентгеновское изображение, нейронная сеть, задача обратного поиска, эмбединг, кластеризация, метод главных компонент, алгоритм t -распределенного стохастического эмбединга соседей.

Для цитирования: Волков А.К., Миронова Л.В., Потапова С.Е. Применение предварительно обученных нейронных сетей для решения задачи обратного поиска рентгеновских изображений запрещенных предметов и веществ // Научный Вестник МГТУ ГА. 2024. Т. 27, № 2. С. 8–24. DOI: 10.26467/2079-0619-2024-27-2-8-24

The use of pretrained neural networks for solving the problem of reverse searching of X-ray images of prohibited items and substances

A.K. Volkov¹, L.V. Mironova¹, S.E. Potapova¹

¹*Ulyanovsk Civil Aviation Institute Named After Air Chief Marshal B.P. Bugaev,
Ulyanovsk, Russia*

Abstract: The paper considers the application of pretrained neural networks to solve the problem of reverse searching of X-ray images of prohibited items and substances. The purpose of the work is to conduct an analysis and substantiate ways to improve the efficiency of baggage and passenger hand luggage X-ray image recognition systems. An analysis of existing domestic and foreign works in the field of baggage and passenger hand luggage X-ray image recognition is presented. It has been revealed that, despite the achieved results in the development of algorithms for recognizing prohibited items and substances, they do not fully cope with such a complexity factor as the overlay of objects. To solve this problem, the paper proposes to additionally analyze X-ray images with low confidence in object recognition. This stage includes the following steps: image segmentation, extraction of features of segmented image elements; search for similar images in the database; decision-making on the class of segmented image elements. This article discusses the last three steps. Variants of approaches to feature extraction from images are analyzed, particularly those based on the application of convolutional autoencoders and pretrained neural networks. The approach based on the application of pretrained neural networks is chosen. The ResNet-50 architecture neural network, pretrained on the ImageNet collection, is used during the work. In order to apply this model to extract image feature vectors, the last classification layer was preliminarily removed. All the previous layers of the model encode the image into a vector. ResNet-50 generates a 2048-dimensional feature vector of images. The principal component analysis is used to reduce the dimensionality of the image feature vectors. The decision of whether the segmented image element is a prohibited item or substance is considered as a reverse search problem using the k-nearest neighbor algorithm. In this case, the class of the X-ray image element is the class most frequently encountered among the K nearest neighbors. In order to test the proposed approach, a training dataset, including 4,635 images of individual items and substances that may be encountered in baggage and passenger hand luggage, was generated. A comparative analysis of image indexing and image search under different algorithms and feature number is presented. A comparative analysis of the model accuracy is provided. It is concluded that the most acceptable is the "Brute force" algorithm in combination with the principal component analysis.

Key words: screening, X-ray image, neural network, reverse search problem, embedding, clustering, principal component analysis, t-distributed stochastic neighbor embedding algorithm.

For citation: Volkov, A.K., Mironova, L.V., Potapova, S.E. (2024). The use of pre-trained neural networks for solving the problem of reverse searching of X-ray images of prohibited items and substances. Civil Aviation High Technologies, vol. 27, no. 2, pp. 8–24. DOI: 10.26467/2079-0619-2024-27-2-8-24

Введение

Обеспечение безопасности объектов транспортной инфраструктуры включает в себя процедуру досмотра багажа с помощью технических средств обнаружения опасных или запрещенных к перевозке предметов. Несмотря на значительные достижения в области технологий и улучшения систем досмотра, исследования показывают, что человеческий фактор оказывает существенное влияние на эффективность и надежность этих систем.

Специалисты досмотровой группы играют ключевую роль в процессе обнаружения запрещенных предметов на рентгеновских изображениях, именно от их знаний и умений зависит качество досмотра. Операторы, работающие за интроскопом, также являются сла-

бым звеном в системе безопасности из-за возможности совершения ошибок. Условия монотонности процесса досмотра и большой объем информации требуют высокой скорости анализа, что может привести к снижению уровня концентрации и внимания оператора и, как следствие, к росту вероятности пропуска запрещенных предметов.

В связи с этим внедрение технологий автоматизированного распознавания рентгеновских изображений представляется важным шагом в улучшении систем досмотра. Подобные технологии могут стать вспомогательными инструментами при принятии решений операторами досмотра. Основу данных технологий составляют компьютерные алгоритмы и система искусственного интеллекта, предварительно обученная на большом количестве рентгеновских

изображений. Применение такого подхода позволяет снизить влияние человеческого фактора, повысить результативность обнаружения запрещенных предметов, увеличить скорость обработки информации.

Целью настоящей статьи является проведение анализа и обоснование путей повышения эффективности систем распознавания рентгеновских изображений багажа и ручной клади пассажиров.

Обзор существующих научных работ

Анализ литературы показал отсутствие на сегодняшний день работ российских авторов, предоставляющих достаточно полный обзор исследований, охватывающих данное направление. В свою очередь зарубежными исследователями представлен ряд публикаций, однако при быстром развитии этой отрасли некоторые алгоритмы распознавания объектов быстро устаревают.

Одной из наиболее известных моделей нейронной сети, используемой для распознавания рентгеновских изображений в условиях уже определенных классов объектов, является сверточная нейронная сеть – CNN (Convolutional Neural Network). Применение технологии сверточных сетей позволяет обнаруживать предметы разного размера и класса. Для разного уровня задач используют либо уже обученные распознавать объекты модели, либо новые наборы изображений объектов для дополнительного обучения моделей, то есть производится трансферное обучение под конкретную, частную задачу [1].

Для обнаружения и классификации объектов на рентгеновских изображениях используются различные архитектуры и алгоритмы нейронных сетей, такие как ResNet, Inception, Inception ResNet, Faster R-CNN, YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector) и др. Расширение поля задач обнаружения объектов способствовало развитию таких архитектур, как DNN (Deep Neural Network) [1]. Особенностью этой архитектуры является обнаружение объектов в режиме реального времени.

В работе [2] внимание уделено сети R-CNN (от Region Based, сверточная сеть на базе регионов), способной к распознаванию объектов в локальных регионах изображений. Сеть R-CNN позволила решать задачу поиска фрагментов на предварительно выделенных областях (регионах), где могут присутствовать искомые объекты. Для выделения областей применяли метод селекции. Авторы использовали архитектуру CaffeNet (является вариантом сети AlexNet) [3] для обучения на наборе изображений ImageNet и распознавания объектов 1000 классов. В таких условиях архитектура сети R-CNN обеспечивала высокую точность и полноту при небольшой производительности. Впоследствии это послужило причиной на основе R-CNN разработать улучшенные модели, в частности сеть Fast R-CNN, применяемую в задачах классификации объектов, ускоренную сеть Faster R-CNN, использующую вспомогательную подсеть для генерации областей объекта. В отличие от предыдущей версии сеть осуществляет извлечение признаков для всего изображения с последующим наложением рамок региона на карту признаков, а также проводит одновременное обучение процедур свертки, формирования опорных векторов и линейной регрессии [2, 4, 5]. В работе [6] авторы использовали набор данных из 32 253 рентгеновских изображений и обучили модели Faster RCNN и RetinaNet.

Сети архитектуры YOLO признаются как самые результативные по скорости и точности обработки изображений. В основе версии сети YOLOv3 лежит сеть Darknet-53 [7]. В отличие от селективных методов выбора региона, которые использовались в архитектуре сети R-CNN, в сети YOLO входное изображение разделяется на квадратные области, для которых выполняется классификация. Для каждого квадрата изображения прогнозируются три ограничивающих прямоугольника и оценивается достоверность присутствия в них объектов. В YOLOv3 возможно обнаружение объектов 80 различных классов.

Также высокую скорость и производительность показывает сеть SSD, в основе которой дискретизация выходного простран-

ства прямоугольных областей обнаружения в прямоугольнике из стандартного набора с заданными размерами для каждого местоположения на карте признаков (характерных особенностей) изображения. Архитектуры SSD [8] и YOLO используют идею якорей – выделения квадратов из изображения и прогнозирования для них покрывающих прямоугольников, для каждого из которых уточняются координаты, определяется приоритетный класс объекта и уровень достоверности отнесения к этому классу. Имея 11 627 рентгеновских снимков, 6 классов (видов распознаваемых предметов) и используя метод переноса обучения, авторы добились отличных результатов. Результаты исследований подтвердили применимость сверточных нейронных сетей для распознавания предметов на рентгеновских изображениях

Можно отметить работу [9], в которой применялась сеть Mask R-CNN, являющаяся расширенной версией Faster R-CNN и позволяющая выявлять контуры объектов разных типов и тем самым обеспечивать задачи обнаружения, классификации и сегментации, т. к. нейронная сеть выделяет отдельные объекты одного класса на изображении как разные сегменты. Для каждого класса получают свои бинарные маски [10]. Для решения задачи определения разных классов объектов одновременно использовались 13 786 изображений, для анализа были отобраны несколько высокопроизводительных комбинаций. В результате исследования было выявлено, что R-CNN с Inception ResNet V2 имеет самые высокие показатели, тогда как SSD с Inception V2 показали худшие результаты. В целом более быстрые модели менее точны, более быстрый R-CNN достиг почти такой же производительности по холодной оружию, как у ResNet V2, но более чем в три раза быстрее. Обнаружено также, что скорость одноступенчатых моделей больше подходит по частоте кадров, что неприменимо для контрольно-пропускных пунктов, так как снижается точность определения угрозы.

В работе [11] для обнаружения опасных металлических предметов также использовались сверточные нейронные сети. Была сфор-

мирована база рентгеновских изображений из 700 экземпляров малых металлических угроз (SMTs – Small Metallic Threats) различных типов и моделей. Часть полученных изображений применялась для обучения путем синтеза образцов, основанного на мультипликативной природе формирования рентгеновских просвечивающих изображений.

В исследовании Университета Дьюка в США совместно с TSA (Transportation Security Administration – Администрация транспортной безопасности) было выделено четыре категории опасных предметов: оружие, ЖАГ (жидкости, аэрозоли, газы), взрывчатые вещества (имитаторы промышленного производства), опасные предметы (колюще-режущего и ударного воздействия). При этом в каждую категорию входило от 10 до 70 предметов. По каждой категории было собрано около 3 500 изображений, что составило около 14 тысяч снимков.

Проблема формирования обучающих наборов данных в настоящее время решается несколькими способами, такими как получение реальных изображений с ограниченными классами запрещенных предметов [12], а также использованием синтезируемых и сгенерированных изображений, в том числе созданных генеративной нейронной сетью [11, 13].

Таким образом, глубокое обучение предоставило ряд возможностей для разработки и модификации алгоритмов обнаружения объектов, и на сегодняшний день сохраняет позиции стремительно развивающейся сферы знаний, в том числе в задачах обработки рентгеновских изображений.

На практике для реализации распознавания рентгеновских изображений в системах безопасности аэропортов используются различные аппаратно-программные решения. Компании Rapiscan Systems, Smiths Detection, L3 Communication Security and Detection Systems, ADANI, NeuralGuard, Hikvision предлагают оборудование, работающее на основе нейронных сетей, для сканирования и обработки рентгеновских изображений. Эти системы обеспечивают автоматическое обнаружение запрещенных предметов и применяют высокотехнологичные методы для анализа и классификации объектов на рентгеновских снимках.

Английская компания NeuralGuard¹ предлагает программно-аппаратный комплекс EyeFox, который позволяет осуществить интеграцию с компьютером интроскопов. В видеоматериалах компании демонстрируется интеграция с моделями ADANI BV6045 и Rapiscan 620DV. Китайская компания Hikvision выпустила интеллектуальный анализатор интроскопа ISD-ST204D. Алгоритм способен определять до 50 категорий предметов. Присутствует возможность установить набор разрешенных и запрещенных предметов².

Из отечественных производителей, занимающихся внедрением компьютерного зрения, можно выделить компании «Диагностика-М» и ООО «Битускан» (продукция B2scan и B2secure)³. «Диагностика-М» оснастила интроскопы конвейерного типа серии ТС-СКАН искусственным интеллектом, что обеспечивает распознавание внешних контуров запрещенных предметов, а также их внутреннее содержимое и их отдельные фрагменты в разобранном виде. В работе интроскопа применяется нейронная сеть, обучающаяся на выборке из 5 млн размеченных изображений и достигающая достаточно высокого уровня локализации объектов.

Таким образом, достигнуты значительные результаты в разработке алгоритмов распознавания запрещенных предметов и веществ. Однако существующие системы распознавания не в полной мере справляются с таким фактором сложности, как наложение предметов. Данный фактор характеризует степень, в которой запрещенный предмет (или его часть) закрыт другими предметами багажа [14]. Для решения данной проблемы в су-

ществующих подходах предлагается увеличивать обучающую выборку данных, чтобы обучить модель на как можно большем количестве примеров. В данной работе предлагается подход, предполагающий дополнительный анализ тех рентгеновских изображений, на которых уверенность в распознавании объектов невысокая. Данное решение включает в себя следующие основные шаги: сегментацию изображений; извлечение признаков сегментированных элементов изображений; поиск схожих изображений по базе данных; принятие решения о классе сегментированных элементов изображений. В данной статье рассматриваются три последних шага.

Методы и методология исследования

Анализ подходов к извлечению признаков из изображений

В настоящее время существуют следующие решения для извлечения признаков из изображений.

Первое решение заключается в построении RGB-гистограмм изображений и их последующего сравнения. Подход позволяет находить похожие изображения, не имеющие существенных отличий в содержимом, при этом имеющие изменения в цвете, оттенке или балансе белого, которые затрудняют распознавание. К недостаткам данного решения можно отнести то, что с ростом обучающего набора данных возрастает вероятность ложных срабатываний.

Следующее решение состоит в том, чтобы выявить характерные визуальные признаки на границе объектов с помощью алгоритмов Scale-Invariant Feature Transform, Speeded Up Robust Features и Oriented FAST and Rotated BRIEF, а затем сравнить количество одинаковых признаков, общих для двух фотографий [15]. Данное решение применяется для оценки сходства изображений объектов с неизменными формами, имеющих мало вариаций.

В исследовании предлагается подход, основанный на нейронных сетях, который

¹ Expose hidden threats—with powerful, precise, ai-powered threat detection for security screening systems [Электронный ресурс] // NeuralGuard Automated Detection. URL: <https://www.neuralguard.com/> (дата обращения: 19.11.2023).

² ISD-ST204D [Электронный ресурс] // Hikvision. URL: https://hikvision.ru/product/isd_st204d (дата обращения: 19.11.2023).

³ «Диагностика-М» выводит на рынок уникальные «умные» интроскопы [Электронный ресурс] // ТСНК ООО «Диагностика-М». URL: https://tsnk.ru/news/2023/05/12/news_462.html (дата обращения: 19.11.2023).

предполагает кодирование изображений в вектор признаков. Под векторами признаков (эмбедингом, от англ. embedding – вложение) понимается обучаемое векторное представление изображения, то есть набор из нескольких сотен или тысяч числовых значений с плавающей запятой. Реализация данного подхода возможна на основе применения сверточных автоэнкодеров или предварительно обученных нейронных сетей.

Сверточные автоэнкодеры для извлечения признаков из изображения

Сверточный автоэнкодер принимает изображение на входе, отображает его в скрытое векторное пространство с помощью модуля кодирования и декодирует его обратно в выходное изображение с теми же размерами с помощью модуля декодирования [16]. Модули кодирования и декодирования представляют собой нейронные сети сверточного типа. Таким образом, автоэнкодер учится восстанавливать исходные данные. Дополнительно можно накладывать различные ограничения на скрытое пространство (слой) автокодировщика.

Для задачи кодирования изображения в вектор используется скрытый слой после модуля кодирования, то самое векторное представление изображения, которое в дальнейшем можно использовать для поиска похожих изображений. К недостаткам использования сверточных автоэнкодеров можно отнести то, что модели нужна более точная настройка параметров для внутренних слоев, больше обучающих данных, а также, в отличие от применения готовых моделей, требуется больше времени на обучение модели.

Использование предобученных моделей для извлечения признаков из изображения

Предварительно обученная сеть – это сохраненная сеть, прежде обученная на большом наборе данных, обычно в рамках масштабной задачи классификации изображений [15]. Для извлечения признаков из изображений применяются предобученные сверточные

нейронные сети, которые предварительно обучались для решения задачи классификации изображений. Данные сети состоят из последовательности нескольких типов слоев, таких как сверточные слои (создают вектора признаков), полносвязные слои и последний слой – классификатор. Для использования подобных моделей для извлечения признаков из изображений необходимо первоначально провести тонкую настройку (дообучение модели). Данная настройка предполагает подстройку нейронной сети для адаптации к новому обучающему набору данных. После оптимизации предобученной модели на новых обучающих данных в ней удаляют последний (классификационный) слой, в результате получается модель, кодирующая изображение в вектор признаков.

Достоинством применения предобученных моделей является то, что нет необходимости строить нейронную сеть с нуля, настраивать сверточные слои, необходимо только дообучить модель на анализируемых данных. Также данные модели, обученные на больших наборах данных, имеющие таким образом настроенные веса для извлечения необходимых признаков, лучше выделяют важные области на новых изображениях.

По результатам проведенного анализа в настоящем исследовании предлагается использовать предобученные модели для извлечения признаков из изображений.

Построение модели подобия изображений при помощи алгоритма K -ближайших соседей

После получения векторных представлений сегментированных элементов изображений необходимо решить, является ли данный объект запрещенным предметом или веществом или нет. Данная задача рассматривается как задача обратного поиска с применением алгоритма K -ближайших соседей. Данный алгоритм включает в себя три этапа:

- 1) расчет расстояния от целевого объекта до всех других объектов в обучающей выборке с использованием выбранной метрики;

2) выбор тех элементов обучающей выборки, до которых минимально расстояние по выбранной метрике от целевого объекта;

3) присвоение класса целевого объекта как класса, наиболее часто встречающегося среди k ближайших соседей.

Существует несколько основных метрик, которые можно использовать в алгоритме K -ближайших соседей: евклидово расстояние; манхэттенское расстояние; метрика Минковского; косинусное расстояние; расстояние Жаккара. Важным моментом при использовании данного алгоритма является выбор оптимального значения k .

Для увеличения скорости поиска изображений можно рассмотреть две стратегии. Первая стратегия основана на уменьшении длины векторов признаков, так как преобученные модели генерируют большой вектор признаков (до нескольких тысяч значений). Вторая стратегия основана на применении более эффективного алгоритма поиска.

Повышение скорости поиска изображений за счет уменьшения длины векторов признаков изображений предлагается реализовать за счет использования метода главных компонент (Principal Component Analysis – PCA). PCA – это статистическая процедура, которая предполагает, что не все признаки, представляющие данные, одинаково важны и, возможно, в наборе признаков есть избыточные, которые можно удалить и получить аналогичные результаты классификации [17]. PCA считается одним из наиболее популярных методов уменьшения размерности. Данный метод не просто удаляет избыточные признаки данных, а генерирует новый набор признаков, являющийся линейной комбинацией исходных признаков. Данные признаки получаются ортогональными друг к другу и называются главными компонентами.

Чтобы определить оптимальное количество признаков, до которого необходимо сократить исходный набор данных, можно воспользоваться долей объясненной дисперсии. Данный параметр отражает ту часть общей дисперсии данных, которую объясняют первые n главных компонент. Часто устанавливают определенный порог и выбирают то

количество компонент, которые в сумме объясняют не менее этой доли (заданной порогом). Также можно воспользоваться анализом вкладов дисперсий для каждого измерения PCA, а также анализом значения кумулятивной дисперсии с каждым измерением PCA.

В качестве различных алгоритмов поиска ближайших соседей в статье будут рассмотрены следующие: алгоритм грубого перебора (Brute force); алгоритм Ball tree и алгоритм K-D tree. Алгоритм Brute force представляет собой самый простой вариант поиска соседей, основанный на попарном вычислении расстояния между точками в обучающем наборе данных. Для повышения вычислительной эффективности на больших наборах данных в алгоритмах Ball tree и K-D tree применяются древовидные структуры данных, которые ускоряют индексацию, тем самым уменьшая время вычислений.

Результаты исследования

Результаты апробации предложенного подхода

В целях апробации предложенного подхода был сформирован обучающий набор данных, включающий 4 635 изображений опасных и бытовых предметов и веществ, разбитых на 15 классов, которые могут встречаться в багаже и ручной клади пассажиров. Эти изображения составили базу данных, в которой в дальнейшем будет происходить поиск вновь сегментированных рентгеновских изображений. По очевидным соображениям для апробации в данной работе будет использоваться только набор изображений неопасных предметов и веществ (3 144 рентгеновских изображения). Разбивка по классам неопасных рентгеновских изображений представлена в табл. 1.

В данной работе для извлечения признаков использовалась архитектура нейронной сети ResNet-50, обученная на коллекции ImageNet. ResNet – это так называемая остаточная нейронная сеть, то есть сеть, построенная на остатке [18]. В архитектуре ResNet-50 применяется 50 слоев. Модель

Таблица 1
Table 1

Разбивка по классам бытовых предметов и веществ
Classification of household items and substances

№ п/п	Класс изображения	Количество
1	Гигиенические принадлежности	634
2	Еда	270
3	Инструменты и канцелярские принадлежности	443
4	Личные вещи	492
5	Одежда	207
6	Посуда	302
7	Бытовая и личная электроника	733
8	Разное	63
Всего		3 144

ResNet-50 генерирует 2048-мерный вектор признаков изображений.

Дообучение модели ResNet-50 проводилось алгоритмом стохастической оптимизации Adam. В качестве функции потерь использовалась категориальная кросс-энтропия.

В качестве параметра k в алгоритме K -ближайших соседей принято значение равное 5. В качестве метрики близости изображений использовалось евклидово расстояние. Будем считать, что целевой объект относится к тому классу, к которому относится 3 и более его ближайших соседей. Первоначально поиск соседей осуществлялся алгоритмом Brute force.

Реализация и обучение предложенных моделей были осуществлены в бесплатной облачной среде для машинного и глубокого обучения Google Colab. Использовался аппаратный ускоритель T4 GPU.

Визуализация кластеров изображений с помощью алгоритма t-SNE

В работе была обучена модель подбора изображений на основе алгоритма K -ближайших соседей, используя которую методом полного перебора будет осуществляться поиск ближайших 5 соседей к целевому объекту на основе евклидова расстояния.

Для визуализации кластеров изображений используется алгоритм t -распределенного стохастического эмбединга соседей (t -distributed Stochastic Neighbor Embedding – t-SNE). Алго-

ритм t-SNE уменьшает размерность многомерного вектора признаков до двух измерений, обеспечивая возможность высокоуровневого обзора набора данных, что помогает увидеть кластеры и близлежащие изображения [19].

Суть работы алгоритма t-SNE заключается в минимизации расхождения функции Кульбака – Лейблера между двумя распределениями вероятностей, что позволяет получить проекцию данных в сниженное пространство. При этом схожие объекты остаются близкими друг к другу, а различные объекты располагаются на некотором расстоянии друг от друга. В качестве распределений вероятностей выступают распределение, которое вычисляется для исходных данных, и распределение, которое строится для целевого (сниженного) пространства.

В качестве значений параметров алгоритма выбраны следующие: перплексия – 35; метрика близости – евклидово расстояние; количество итераций – 1 000. Параметр «перплексия» в алгоритме t-SNE характеризует количество соседей, которое будет учитываться в расчете условных вероятностей сходства. Параметр принимает значения от 5 до 50. Выбор конкретного значения зависит от целей исследования. Если требуется визуализировать более общую структуру данных, то выбирают большие значения перплексии. Для визуализации локальной структуры выбирают маленькие значения перплексии. В данной работе выбрано среднее значение равное 35.

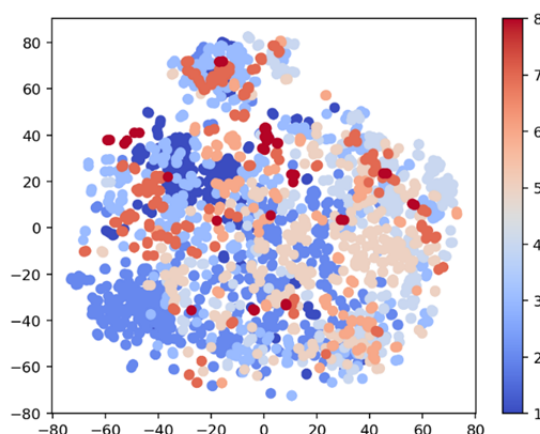


Рис. 1. Визуализация кластеров изображений с помощью алгоритма t-SNE
(до сокращения количества признаков)

Fig. 1. Visualization of image clusters using the algorithm t-SNE (before feature number reduction)

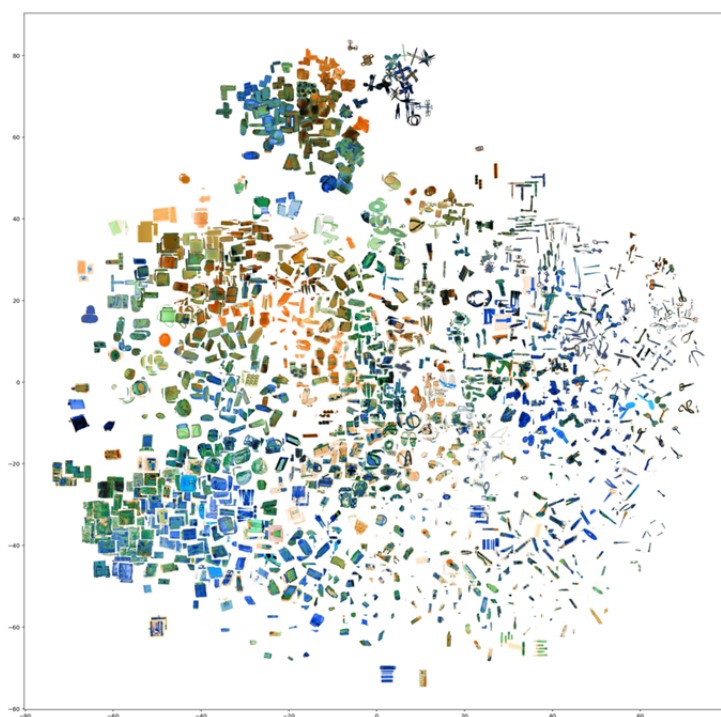


Рис. 2. Визуализация кластеров изображений с помощью алгоритма
t-SNE с добавлением самих изображений (до сокращения количества признаков)

Fig. 2. Visualization of image clusters using the algorithm
t-SNE with the addition of the images themselves (before feature number reduction)

На рис. 1 показаны полученные кластеры элементов рентгеновских изображений (каждый кластер представляет определенный класс объектов и отмечен одним цветом) и их расположение относительно друг друга.

Для повышения интерпретируемости кластеров добавим сами изображения в эти

кластеры и также визуализируем их с помощью алгоритма t-SNE (рис. 2).

Из рис. 2 видно, что в каждом кластере сгруппированы рентгеновские изображения элементов багажа или ручной клади одного класса. При этом выделенные классы бытовых предметов и веществ находятся близко

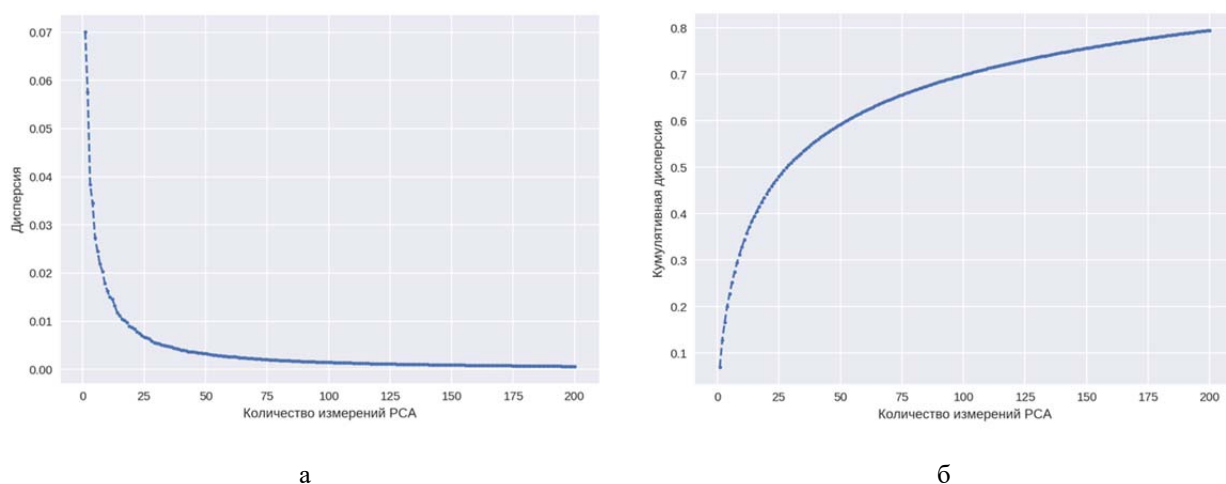


Рис. 3. Дисперсия для каждого измерения PCA (а); значения кумулятивной дисперсии с каждым измерением PCA (б)
Fig. 3. A Variance for each PCA measurement (a); cumulative variance values with each PCA measurement (б)

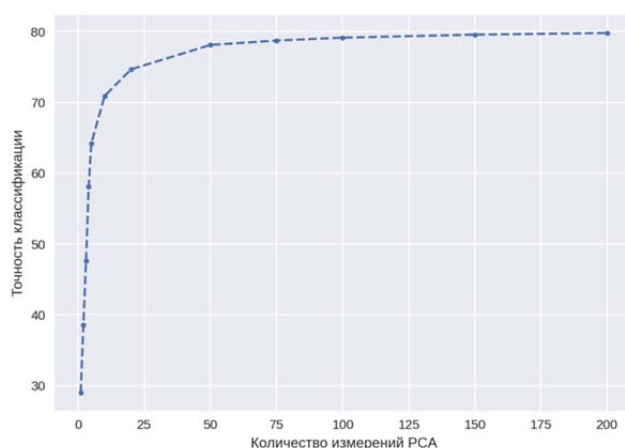


Рис. 4. Прирост точности с увеличением числа главных компонент
Fig. 4. Increment of accuracy with increasing the number of principle components

к себе подобным. К примеру, класс «Бытовая и личная электроника» расположен в левой части графика. Кластеры органических предметов (еда) находятся в центре графика.

Визуализация кластеров после уменьшения размерности векторов признаков методом главных компонент

Как было отмечено выше, PCA преобразует исходный вектор признаков изображений в новый вектор, но с меньшими размерами. Для определения баланса между полнотой описания исходных данных и желаемым количеством признаков в уменьшенном набо-

ре оценим вклад первых 200 измерений. Результаты анализа вкладов дисперсий для данного ряда измерений PCA показаны на рис. 3, а.

Исходя из анализа индивидуальных вкладов компонент (рис. 3, а) в общую дисперсию можно сделать вывод, что главные компоненты, следующие за первой сотней, почти не вносят вклада в общую дисперсию (их вклад близок к 0) и ими можно пренебречь. График накопленной объясняемой дисперсии представлен на рис. 3, б. Дополнительные 100 измерений (от 100 до 200) объясняют лишь небольшой процент от кумулятивной дисперсии. Затем график кумулятивной дисперсии

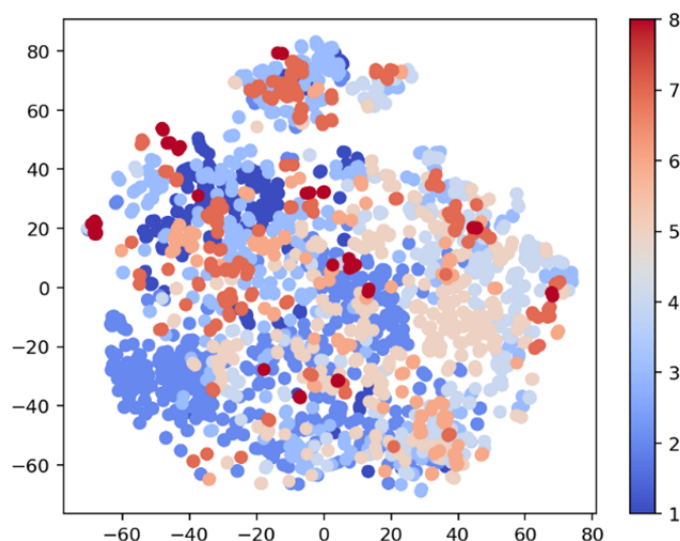


Рис. 5. Визуализация кластеров изображений с помощью алгоритма t-SNE
(после сокращения количества признаков)

Fig. 5. Visualization of image clusters using the algorithm t-SNE (after feature number reduction)

постепенно выходит на плато. Также оценим влияние количества признаков на точность PCA. Полученные результаты представлены на соответствующем графике (рис. 4).

Как видно на рис. 4, после 100 измерений дальнейшее увеличение их числа не приводит к более высокой точности. Таким образом, первые 100 главных компонент содержат большую часть информации об анализируемых векторах признаков рентгеновских изображений. Следовательно, значение равное 100 будет выбрано в качестве длины вектора признаков после применения PCA. Сокращение размера вектора признаков почти в 20 раз от исходного (2 048) позволит обеспечить значительно более высокую скорость поиска при сопоставимой точности, что будет показано в дальнейшем.

После сокращения количества признаков также была обучена модель ближайших соседей с теми же параметрами, что и у предыдущей модели. На рис. 5 показаны полученные кластеры элементов рентгеновских изображений. На данном рисунке видно, что кластеры стали более обособленными.

Также для повышения интерпретируемости кластеров добавим сами изображения

в эти кластеры и также визуализируем их с помощью алгоритма t-SNE (рис. 6).

После уменьшения длины векторов признаков видно (рис. 6), что кластеры стали немного более компактными и обособленными.

Пример поиска изображения

Рассмотрим пример поиска изображения фрагмента ключа (рис. 7, а), относящийся к классу личных вещей. Пусть в результате сегментации рентгеновского изображения получен следующий фрагмент (рис. 7, б), относительно которого необходимо решить, к какому классу он относится. При помощи модели подобию, обученной с применением сокращенного вектора признаков (100-мерный вектор), найдем ближайших соседей.

Найденные ближайшие соседи, их классы и расстояния до них представлены на рис. 8.

Согласно рис. 8 из 5 найденных соседей 4 относятся к классу личных вещей, а 1 к классу гигиенических принадлежностей. Таким образом, можно считать, что целевой объект относится к классу личных вещей. Предлагаемая модель справилась с поставленной задачей.

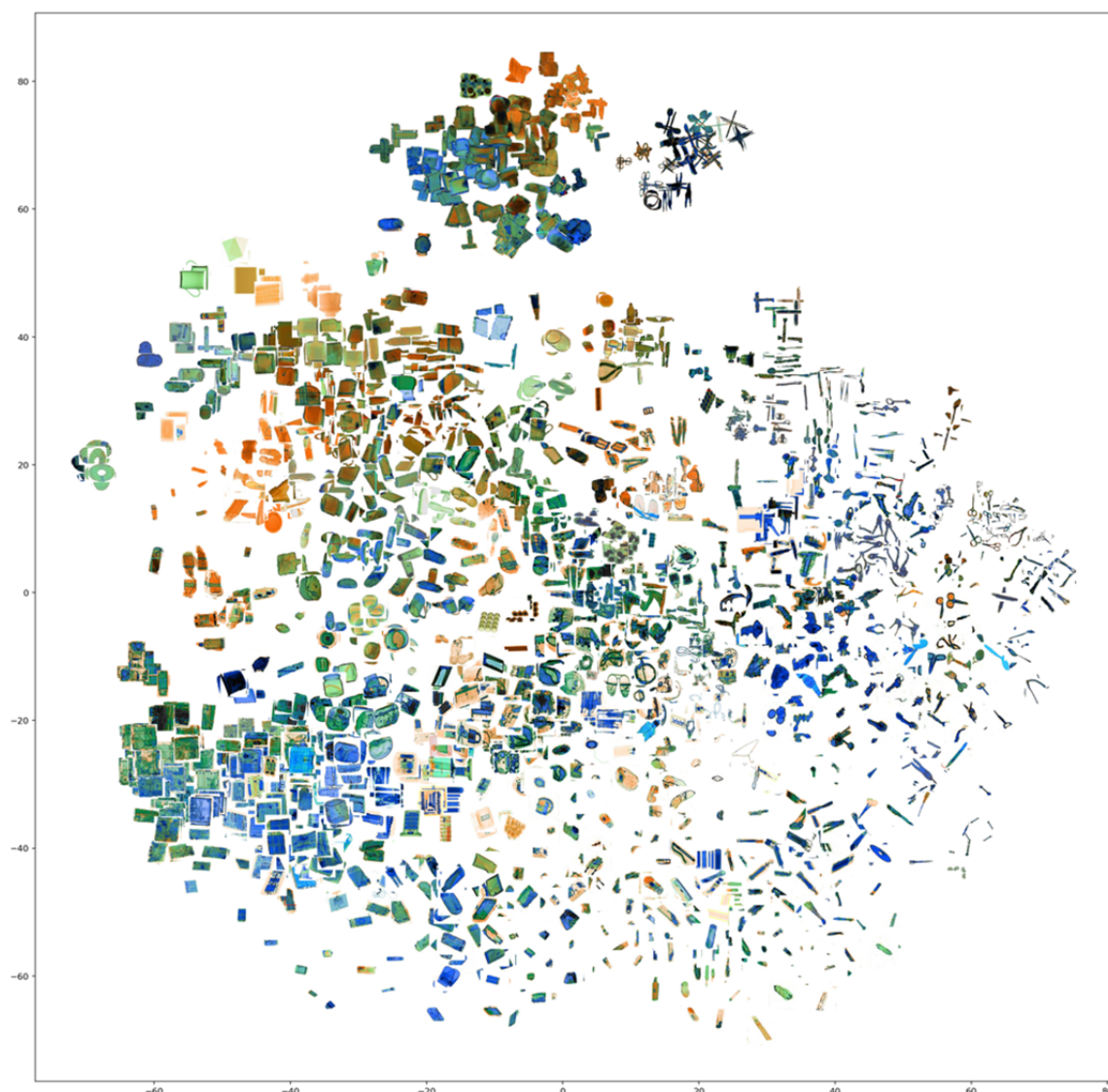


Рис. 6. Визуализация кластеров изображений с помощью алгоритма t-SNE с добавлением самих изображений (после сокращения количества признаков)

Fig. 6. Visualization of image clusters using the algorithm t-SNE with the addition of the images themselves (after feature number reduction)



а



б

Рис. 7. Исходный вид целевого изображения (а); вид сегментированного целевого изображения (б)

Fig. 7. The original view of the target image (а); the view of the segmented target image (б)

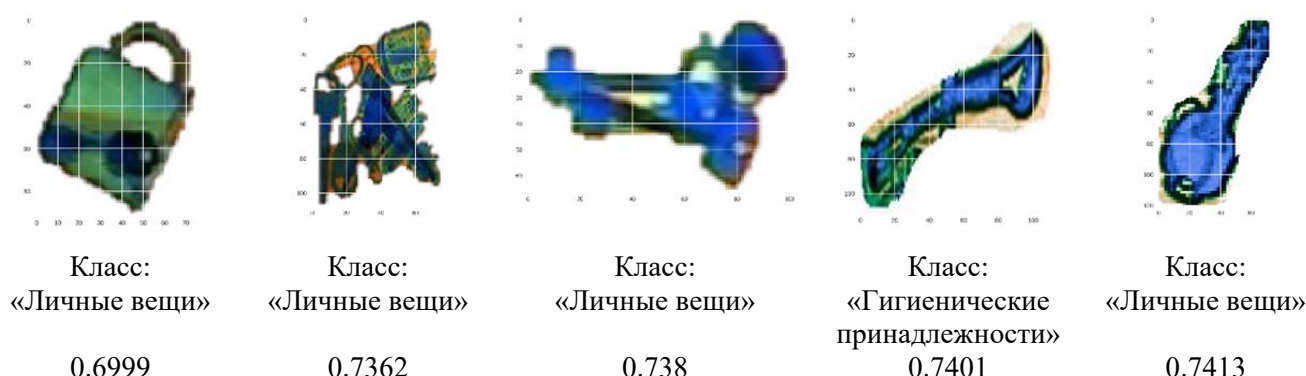


Рис. 8. Ближайшие соседи и расстояния до них
Fig. 8. Nearest neighbors and distances to them

Обсуждение полученных результатов

Оценка времени поиска и индексации изображений

В табл. 2 представлен сравнительный анализ времени индексации и поиска изображений при различных алгоритмах и количестве признаков.

Анализ табл. 3 позволяет сделать вывод, что время индексирования алгоритмов K-D tree и Ball tree значительно выше алгоритма Brute force. При этом данный алгоритм в сочетании с PCA имеет меньшее значение времени поиска 1 000 изображений. В связи с этим предлагается использовать данный алгоритм. Не исключено, что при дальнейшем увеличении обучающего набора данных эффективнее станет какой-либо из алгоритмов K-D tree или Ball tree.

Таким образом, предлагаемый подход потенциально пригоден для применения в реальных производственных условиях.

Оценка точности модели

В табл. 3 представлен сравнительный анализ точности модели.

Согласно табл. 4 видно, что при применении PCA точность несколько снижается, но

все же остается сопоставимой с моделью, использующей 2048-мерный вектор признаков изображений.

Заключение

В статье проведен анализ и представлено предложение по повышению эффективности систем распознавания рентгеновских изображений багажа и ручной клади пассажиров.

По результатам проведенного анализа отечественных и зарубежных работ выявлено, что, несмотря на достигнутые результаты по разработке алгоритмов распознавания запрещенных предметов и веществ, они не в полной мере справляются с таким фактором сложности, как наложение предметов. Для решения данной проблемы в работе предлагается подход, предполагающий дополнительный анализ тех рентгеновских изображений, на которых уверенность в распознавании объектов невысокая. В работе сформирован набор обучающих данных и представлена апробация решения задачи поиска запрещенного предмета.

Практическая значимость результатов исследования заключается в том, что применение предложенного подхода позволит повысить точность распознавания запрещенных предметов и веществ на рентгеновских изображениях багажа и ручной клади пассажиров.

Таблица 2
Table 2

Результаты анализа времени индексации и поиска изображений
Results of indexing time and image search analysis

Алгоритм	Количество признаков изображения	Время поиска 1 изображения	Время поиска 1 000 изображений	Время индексирования обучающего набора данных
Brute force	2 048	17,1 мс	295 мс	2,27 мс
K-D tree	2 048	13,8 мс	12,9 с	725 мс
Ball tree	2 048	10,7 мс	9,81 с	554 мс
PCA + Brute force	100	1,57 мс	37,3 мс	22,3 мс
PCA + K-D tree	100	1,03 мс	705 мс	44,9 мс
PCA + Ball tree	100	864 мкс	455 мс	38,3 мс

Таблица 3
Table 3

Результаты анализа точности модели
Results of model accuracy analysis

Алгоритм	Точность модели
Brute Force	80,11
PCA + Brute Force	79,26

Направления дальнейших исследований связаны с увеличением числа обучаемых данных, нахождением оптимального значения параметра k в алгоритме ближайших соседей, к примеру «методом локтя»; исследованием влияния предварительной обработки изображений на точность работы алгоритмов.

Список литературы

1. **Bozinovski S., Ante F.** The influence of pattern similarity and transfer learning upon training of a base perceptron B2 // Proceedings of Symposium Informatica. 1976. № 3. Pp. 121–126.
2. **Girshick R.** Fast R-CNN // 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile, 2015. Pp. 1440–1448. DOI: 10.1109/ICCV.2015.169
3. **Krizhevsky A., Sutskever I., Hinton G.E.** ImageNet classification with deep convolutional neural networks // Communications of

the ACM. 2017. Vol. 60, iss. 6. Pp. 84–90. DOI: 10.1145/3065386

4. **Girshick R.** Rich feature hierarchies for accurate object detection and semantic segmentation / R. Girshick, J. Donahue, T. Darrell, J. Malik // 2014 IEEE Conference on Computer Vision and Pattern Recognition, USA, Columbus, 2014. Pp. 580–587. DOI: 10.1109/CVPR.2014.81

5. **Ren S.** Faster R-CNN: Towards real-time object detection with region proposal networks / S. Ren, K. He, R. Girshick, J. Sun // Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS'15). 2015. Vol. 1. Pp. 91–99.

6. **Liu J., Leng X., Liu Y.** Deep convolutional neural network based object detector for x-ray baggage security imagery // 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). USA, Portland, 2019. Pp. 1757–1761. DOI: 10.1109/ICTAI.2019.00262

7. **Redmon J., Farhadi A.** YOLOv3: An incremental improvement [Электронный ресурс] // Computer Vision and Pattern Recognition. 2018. DOI: 10.48550/arXiv.1804.02767 (дата обращения: 19.11.2023).

8. **Liu W., Anguelov D., Erhan D. et al.** SSD: Single shot multibox detector // Proceedings 14th European Conference: Computer Vision – ECCV 2016. The Netherlands, Amsterdam, October 11–14, 2016. Part 1. Pp. 1–17. DOI: 10.1007/978-3-319-46448-0_2

9. **He K.** Mask R-CNN / К. Хе, G. Gkioxari, P. Dollár, R. Girshick [Электронный ресурс] // Computer Vision and Pattern Recognition. 2017. DOI: 10.48550/arXiv.1703.06870 (дата обращения: 19.11.2023).

10. **Андрянов Н.А., Дементьев В.Е., Ташлинский А.Г.** Обнаружение объектов на изображении: от критериев Байеса и Неймана-Пирсона к детекторам на базе нейронных сетей EfficientDet // Компьютерная оптика. 2022. Т. 46, № 1. С. 139–159. DOI: 10.18287/2412-6179-CO-922

11. **Jaccard N.** Automated detection of smuggled high-risk security threats using deep learning / N. Jaccard, T. Rogers, E. Morton, L. Griffin [Электронный ресурс] // Computer Vision and Pattern Recognition. 2016. 7 p. DOI: 10.48550/arXiv.1609.02805 (дата обращения: 19.11.2023).

12. **Liang K.J., Sigman J.B., Spell G.P. et al.** Toward automatic threat recognition for airport x-ray baggage screening with deep convolutional object detection [Электронный ресурс] // Computer Vision and Pattern Recognition. 2019. 11 p. DOI: 10.48550/arXiv.1912.06329 (дата обращения: 19.11.2023).

13. **Kolte S., Bhowmik N., Dhiraj.** Threat Object-based anomaly detection in X-ray images using GAN-based ensembles [Электронный ресурс] // Neural Computing and Applications. 2023. Vol. 35. Pp. 23025–23040. DOI: 10.1007/s00521-022-08029-z (дата обращения: 19.11.2023).

14. **Michel S.** Increasing X-ray image interpretation competency of cargo security screeners / S. Michel, M. Mendes, J.C. de Ruiter, C.M. GerKoomen, A. Schwaninger // International Journal of Industrial Ergonomics. 2014.

Vol. 44, iss. 4. Pp. 551–560. DOI: 10.1016/j.ergon.2014.03.007

15. **Коул А., Ганджу С., Казам М.** Искусственный интеллект и компьютерное зрение. Реальные проекты на Python, Keras и TensorFlow. СПб.: Питер, 2023. 608 с.

16. **Шолле Ф.** Глубокое обучение на Python. СПб.: Питер, 2018. 400 с.

17. **Айвазян С.А.** Прикладная статистика. Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. М.: Финансы и статистика, 1989. 607 с.

18. **He K.** Deep residual learning for image recognition / К. Хе, X. Zhang, S. Ren, J. Sun // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). USA, Las Vegas, 2016. Pp. 770–778. DOI: 10.1109/CVPR.2016.90

19. **Van der Maaten L.J.P., Hinton G.E.** Visualizing high-dimensional data using t-SNE // Journal of Machine Learning Research. 2008. Vol. 9. Pp. 2579–2605.

References

1. **Bozinovski, S., Ante, F.** (1976). The influence of pattern similarity and transfer learning upon training of a base perceptron B2. In: *Proceedings of Symposium Informatica*, no. 3, pp. 121–126.

2. **Girshick, R.** (2015). Fast R-CNN. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile, pp. 1440–1448. DOI: 10.1109/ICCV.2015.169

3. **Krizhevsky, A., Sutskever, I., Hinton, G.E.** (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, vol. 60, issue 6, pp. 84–90. DOI: 10.1145/3065386

4. **Girshick, R., Donahue, J., Darrell, T., Malik, J.** (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, pp. 580–587. DOI: 10.1109/CVPR.2014.81

5. **Ren, S., He, K., Girshick, R., Sun, J.** (2015). Faster R-CNN: Towards real-time object

detection with region proposal networks. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS'15)*, vol. 1, pp. 91–99.

6. Liu, J., Leng, X., Liu, Y. (2019). Deep convolutional neural network based object detector for x-ray baggage security imagery. In: *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. USA, Portland, pp. 1757–1761. DOI: 10.1109/ICTAI.2019.00262

7. Redmon, J., Farhadi, A. (2018). YOLOv3: An incremental improvement. *Computer Vision and Pattern Recognition*. DOI: 10.48550/arXiv.1804.02767 (accessed: 19.11.2023).

8. Liu, W., Anguelov, D., Erhan, D. et al. (2016) SSD: Single shot multibox detector. In: *Proceedings 14th European Conference: Computer Vision – ECCV 2016*. The Netherlands, Amsterdam, October 11–14, part 1, pp. 1–17. DOI: 10.1007/978-3-319-46448-0_2

9. He, K., Gkioxari, G., Dollár, P., Girshick, R. (2017). Mask R-CNN. *Computer Vision and Pattern Recognition*. DOI: 10.48550/arXiv.1703.06870 (accessed: 19.11.2023).

10. Andriyanov, N.A., Dementiev, V.E., Tashlinskiy, A.G. (2022). Detection of objects in the images: from likelihood relationships towards scalable and efficient neural networks. *Computer Optics*, vol. 46, no 1, pp. 139–159. DOI: 10.18287/2412-6179-CO-922 (in Russian)

11. Jaccardm, N., Rogers, T., Morton, E., Griffin, L. (2016). Automated detection of smuggled high-risk security threats using deep learning. *Computer Vision and Pattern Recognition*, 7 p. DOI: 10.48550/arXiv.1609.02805 (accessed: 19.11.2023).

12. Liang, K.J., Sigman, J.B., Spell, G.P., (2019). Toward automatic threat recognition for

airport x-ray baggage screening with deep convolutional object detection. *Computer Vision and Pattern Recognition*, 11 p. DOI: 10.48550/arXiv.1912.06329 (accessed: 19.11.2023).

13. Kolte, S., Bhowmik, N., Dhiraj. (2023). Threat Object-based anomaly detection in X-ray images using GAN-based ensembles. *Neural Computing and Applications*, vol. 35, pp. 23025–23040. DOI: 10.1007/s00521-022-08029-z (accessed 19.11.2023).

14. Michel, S., Mendes, M., de Ruiter, J.C., GerKoomen, C.M., Schwaninger, A. (2014). Increasing X-ray image interpretation competency of cargo security screeners. *International Journal of Industrial Ergonomics*, vol. 44, issue 4, pp. 551–560. DOI: 10.1016/j.ergon.2014.03.007

15. Koul, A., Gandzhu, S., Kazam, M. (2023). Artificial intelligence and computer vision. Real projects in Python, Keras and TensorFlow. St. Petersburg: Piter, 608 p. (in Russian)

16. Sholle, F. (2018). Deep learning in Python. St. Petersburg: Piter, 400 p. (in Russian)

17. Ayvazyan, S.A., Buxshtaber, V.M., Enyukov, I.S., Meshalkin, L.D. (1989). Applied statistics. Classification and dimensionality reduction. Moscow: Finansy i statistika, 607 p. (in Russian)

18. He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA, Las Vegas, pp. 770–778. DOI: 10.1109/CVPR.2016.90

19. Van der Maaten, L.J.P., Hinton, G.E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605.

Сведения об авторах

Волков Александр Константинович, кандидат технических наук, доцент, доцент кафедры обеспечения авиационной безопасности ФГБОУ ВО «Ульяновский институт гражданской авиации имени Главного маршала авиации Б.П. Бугаева», volkovalex8@rambler.ru.

Миронова Лидия Владимировна, соискатель кафедры обеспечения авиационной безопасности ФГБОУ ВО «Ульяновский институт гражданской авиации имени Главного маршала авиации Б.П. Бугаева», lida737@yandex.ru.

Потапова Светлана Евгеньевна, старший преподаватель кафедры обеспечения авиационной безопасности ФГБОУ ВО «Ульяновский институт гражданской авиации имени Главного маршала авиации Б.П. Бугаева», potapovase@yandex.ru.

Information about the authors

Alexander K. Volkov, Candidate of Technical Sciences, Associate Professor, Associate Professor at the Chair of Aviation Security, Ulyanovsk Civil Aviation Institute named after Air Chief Marshal B.P. Bugaev, volkovalex8@rambler.ru.

Lidiya V. Mironova, Applicant at the Chair of Aviation Security, Ulyanovsk Civil Aviation Institute named after Air Chief Marshal B.P. Bugaev, lida737@yandex.ru.

Svetlana E. Potapova, Senior Lecturer at the Chair of Aviation Security, Ulyanovsk Civil Aviation Institute named after Air Chief Marshal B.P. Bugaev, potapovase@yandex.ru.

Поступила в редакцию	07.12.2023	Received	07.12.2023
Одобрена после рецензирования	16.01.2024	Approved after reviewing	16.01.2024
Принята в печать	21.03.2024	Accepted for publication	21.03.2024